

AD-A136 673

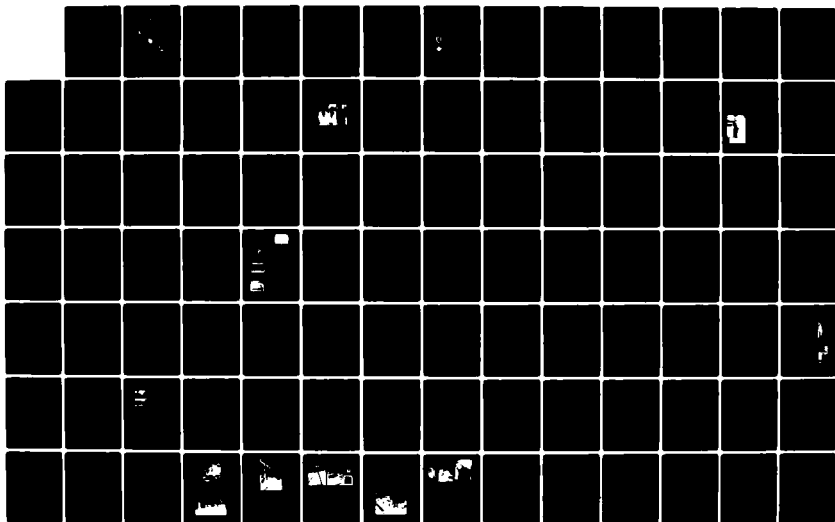
PROCEEDINGS OF THE ANNUAL SYMPOSIUM ON FREQUENCY
CONTROL (37TH) 1-3 JUNE 1983 (U) ARMY ELECTRONICS
RESEARCH AND DEVELOPMENT COMMAND FORT MONMOUTH, 1983

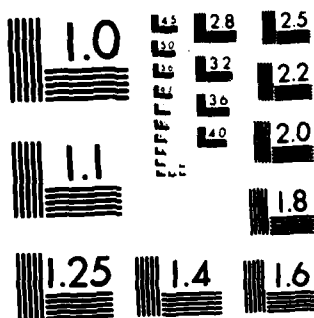
1/6

UNCLASSIFIED

F/G 9/1

NL



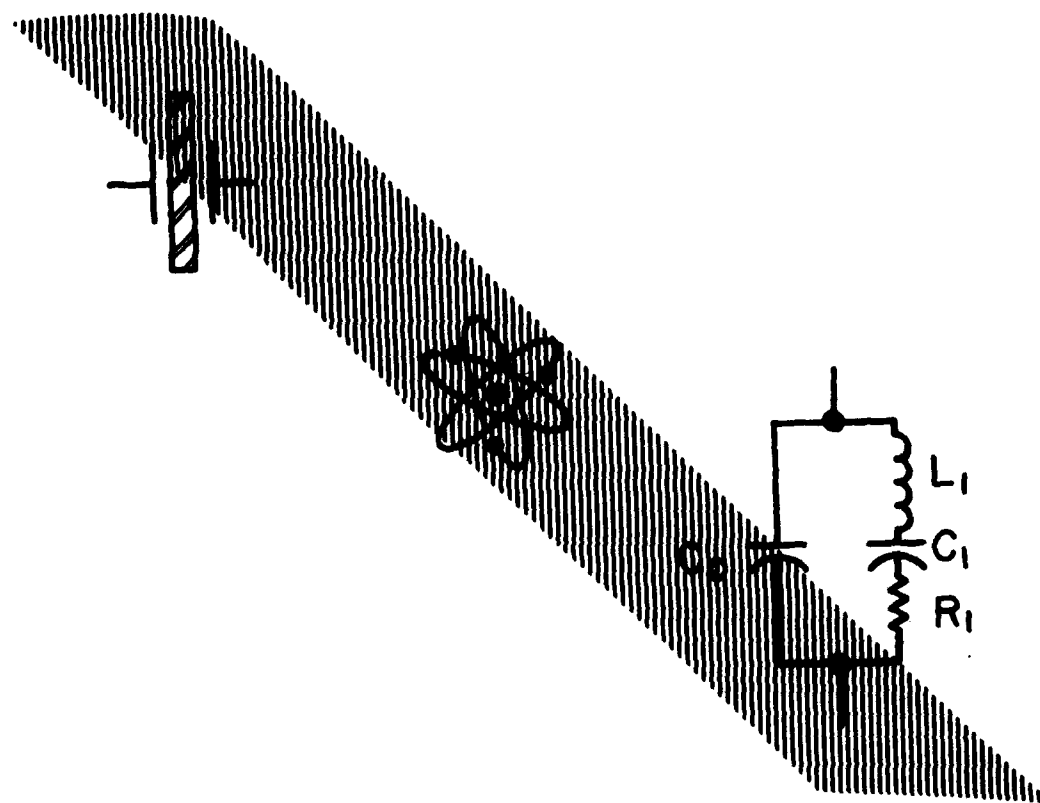


MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

(1)

**PROCEEDINGS
OF THE
37TH ANNUAL SYMPOSIUM ON FREQUENCY CONTROL
1983**

A136673



1 - 3 JUNE 1983

This document has been approved
for public release and sale; its
distribution is unlimited.

83CH1967-0

58-60781

S
A

NOV 4 1983

83 11 03 106

DTC FILE COPY

COMPONENT PART NOTICE

THIS PAPER IS A COMPONENT PART OF THE FOLLOWING COMPILATION REPORT:

(TITLE): Proceedings of the Annual Symposium on Frequency Control (37th),
1-3 June 1983, Marriott Hotel, Philadelphia, Pennsylvania.

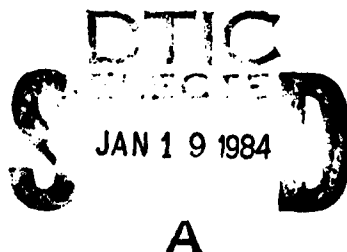
(SOURCE): Army Electronics Research and Development Command, Fort Monmouth, NJ.
Electronics Technology/Devices Lab.

TO ORDER THE COMPLETE COMPILATION REPORT USE AD-A136 673.

THE COMPONENT PART IS PROVIDED HERE TO ALLOW USERS ACCESS TO INDIVIDUALLY AUTHORED SECTIONS OF PROCEEDINGS, ANNALS, SYMPOSIA, ETC. HOWEVER, THE COMPONENT SHOULD BE CONSIDERED WITHIN THE CONTEXT OF THE OVERALL COMPILATION REPORT AND NOT AS A STAND-ALONE TECHNICAL REPORT.

THE FOLLOWING COMPONENT PART NUMBERS COMPRISE THE COMPILATION REPORT:

AD#:	TITLE:
AD-P002 444	Application of Metal Hydrides for Gas Handling in Hydrogen Masers.
AD-P002 445	Compact Rectangular Cavity for Rubidium Vapor Cell Frequency Standards.
AD-P002 446	Improved Vibration Performance in Passive Atomic Frequency Standards by Servo-Loop Control.
AD-P002 447	Ultra Stable Laser Clock, Second Generation.
AD-P002 448	Microprocessor for the NR Series of Hydrogen Maser Frequency Standards.
AD-P002 449	Mercury-199 Trapped Ion Frequency Standard: Recent Theoretical Progress and Experimental Results.
AD-P002 450	Laser Cooled $^9\text{Be}^+$ Accurate Clock.
AD-P002 451	Performance of Compact Hydrogen Masers.
AD-P002 452	Cold Hydrogen Maser Research at SAO and Related Developments.
AD-P002 453	Time Dissemination and Transfer.
AD-P002 454	An International Time Transfer Experiment.
AD-P002 455	Precision Timekeeping at the Observatory Lustuehel, Graz, Austria.
AD-P002 456	Time Dissemination from the National Research Council of Canada.
AD-P002 457	Signal Processing Techniques.
AD-P002 458	Low Phase Noise Multiple Frequency Microwave Source.
AD-P002 459	Synchronization and Tracking with Synchronous Oscillators.
AD-P002 460	Frequency (Standard) Combiner Selector.
AD-P002 461	Properties of Quartz and Other Crystals.



This document has been approved
for public release and sale; its
distribution is unlimited.

COMPONENT PART NOTICE (CON'T)

AD#:	TITLE:
AD-P002 462	Tensile Fracture Strength of ST Cut Quartz.
AD-P002 463	Effect of Alkali Ions on Electrical Conductivity and Dielectric Loss of Quartz Crystals.
AD-P002 464	Radiation Induced Transient Acoustic Loss in Quartz Crystals.
AD-P002 465	The Bulk Acoustic Wave Properties of Lithium Tetraborate.
AD-P002 466	Sputtered c-Axis Inclined Piezoelectric Films and Shear Wave Resonators.
AD-P002 467	Recalibration of Q Capability Indications from Infrared Measurements on Cultured Quartz.
AD-P002 468	Pressure Volume Temperature Behavior in the System $H_2O-NaOH-SiO_2$ and its Relationship to the Hydrothermal Growth of Quartz.
AD-P002 469	Standard Characterization Methods for the Determination of the Quality of Hydrothermally Grown Quartz.
AD-P002 470	Electrodifusion of Charge-Compensating Ions in Alpha-Quartz.
AD-P002 471	Aluminum and Hydroxide Defect Centers in Vacuum Swept Quartz.
AD-P002 472	Effect of Seed Treatment on Quartz Dislocations.
AD-P002 473	Thermoluminescence from Different Growth Sectors in Synthesis Quartz Crystals.
AD-P002 474	Characterization of Quartz Crystals by Cathodoluminescence.
AD-P002 475	Improvements of Laser Interferometric Measurement System of Vibration Displacements.
AD-P002 476	The Stress Coefficient of Frequency of Quartz Plate Resonators.
AD-P002 477	Temperature Derivatives of Elastic Stiffness Derived from the Frequency Temperature Behavior of Quartz Plates.
AD-P002 478	Transient Thermally Induced Frequency Excursions in AT- and SC-Cut Quartz Trapped Energy Resonators.
AD-P002 479	Excess Noise in Quartz Crystal Resonators.
AD-P002 480	Thickness Modes in Circular AT-Cut Quartz Plates with Circular Electrodes.
AD-P002 481	Properties of AT Quartz Resonators on Wedgy Plates.
AD-P002 482	Acoustically Coupled Resonators: Filters and Pressure Transducers.
AD-P002 483	Force Sensor Using Double-Ended Tuning Fork Quartz Crystals.
AD-P002 484	Practical Considerations for Miniature Quartz Resonator Force Transducers.
AD-P002 485	Chemically Polished High Frequency Resonators.
AD-P002 486	Further Studies on the Acceleration Sensitivity of Quartz Resonators.
AD-P002 487	Determination of Acceleration Sensitivity of Bulk Mode Resonator Plates.
AD-P002 488	An Instrument for Automated Measurement of the Angles of Cut of Doubly Rotated Quartz Crystals.
AD-P002 489	A Measurement Technique for Determination of Frequency vs Acceleration Characteristics of Quartz Crystal Units.
AD-P002 490	Evaluation of Crystal Measurement Systems.

COMPONENT PART NOTICE (CON'T)

AD#:	TITLE:
AD-P002 491	A New Frequency for Piezoelectric Resonator Measurement.
AD-P002 492	Measuring Method of Equivalent Series Capacitance and Negative Resistance of Quartz Crystal Oscillator Circuits.
AD-P002 493	S.Y. Parameters Method for Accurate Measurement of Bulk Wave Crystals Resonators at Frequencies up to 2GHz.
AD-P002 494	Spurious Resonances and Modelling of Composite Resonators.
AD-P002 495	Analysis of Composite Resonator Geometries.
AD-P002 496	An Analysis of Thickness Extensional Trapped Energy Resonators with Rectangular Electrodes in the Zinc-Oxide Thin Film on Silicon Configuration.
AD-P002 497	Minature LiTaO ₃ X-Cut Strip Resonator.
AD-P002 498	Miniaturized LiTaO ₃ Strip Resonator.
AD-P002 499	Current Trends in Crystal Filters.
AD-P002 500	Surface Acoustic Wave Bandpass Filters.
AD-P002 501	Digital Filters: An Overview.
AD-P002 502	Microelectronic Analog Active Filters.
AD-P002 503	Timing Tank Mechanical Filter for Digital Subscriber Transmission System.
AD-P002 504	Stripline Filters - An Overview.
AD-P002 505	Precision L-Band SAW Oscillator for Satellite Application.
AD-P002 506	SAW Resonator Stabilized Oscillator.
AD-P002 507	Very Long Period Random Frequency Fluctuations in SAW Oscillators.
AD-P002 508	Stress Induced Effects on the Propagation of Surface Waves.
AD-P002 509	Surface Wave Resonators on Silicon.
AD-P002 510	High-Frequency Shallow Bulk Acoustic Wave Quartz Frequency Sources.
AD-P002 511	A New Digital TCXO Circuit Using a Capacitor Switch Array.
AD-P002 512	An Improved Method of Temperature Compensation of Crystal Oscillators.
AD-P002 513	A New Frequency Temperature Compensation Method for Oscillator.
AD-P002 514	An Ultra-Stable Crystal Oscillator for Beacons (Program "SARSAT").
AD-P002 515	Modeling of TTL Inverter Amplifiers as Applied to Overtone Crystal Clock Oscillators.
AD-P002 516	MSSW Delay Line Based Oscillators.
AD-P002 517	The Status of Magnetostatic Wave Oscillators.
AD-P002 518	An X-Band GaAs FET Oscillator Using a Dielectric Resonator.
AD-P002 519	An Analysis of a Microprocessor Controlled Disciplined Frequency Standard.
AD-P002 520	Measurement Assurance.
AD-P002 521	Temperature Compensated Crystal Oscillator Survey and Test Results.
AD-P002 522	A New Frequency Calibration Service Offered by the National Bureau of Standards.

COMPONENT PART NOTICE (CON'T)

AD# :

AD-P002 523

AD-P002 524

AD-P002 525

TITLE:

Specifying Performance for Atomic Standards.

Platform Distribution System Specifications.

Modernization of the Military Specification for Quartz Crystal Units.

[illegible]

PROCEEDINGS
OF THE
37th ANNUAL
FREQUENCY CONTROL SYMPOSIUM
1983

Co-Sponsored by the



U.S. ARMY ELECTRONICS RESEARCH
AND DEVELOPMENT COMMAND
ELECTRONICS TECHNOLOGY AND
DEVICES LABORATORY

and



THE INSTITUTE OF ELECTRICAL AND
ELECTRONICS ENGINEERS, INC.
SONICS AND ULTRASONICS GROUP

IEEE Catalog No. 83CH1957-0

Library of Congress No. 58-60781

1-3 June 1983

Marriott Hotel

Philadelphia, Pennsylvania

Abstracting is permitted with 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 noncommercial 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 © 1983 by The Institute of Electrical and Electronics Engineers, Inc.

Co-Sponsored by

U.S. ARMY ELECTRONICS RESEARCH AND DEVELOPMENT COMMAND
ELECTRONICS TECHNOLOGY AND DEVICES LABORATORY

AND

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.
SONICS AND ULTRASONICS GROUP

General Chairman.....Dr. John R. Vig, U.S. Army ERADCOM
Technical Program Chairman.....Dr. Samuel R. Stein, National Bureau of Standards

Dr. A. Ballato U.S. Army, ERADCOM	Dr. T. Lukaszek U.S. Army, ERADCOM
Mr. R. Beehler National Bureau of Standards	Dr. T. Meeker Bell Laboratories
Mr. M. Bloch Frequency Electronics, Inc.	Mr. V. Rosati U.S. Army, ERADCOM
Dr. L. Claiborne Texas Instruments, Inc.	Mr. S. Schodowski U.S. Army, ERADCOM
Mr. L. Conlee Motorola, Inc.	Dr. E. Staples Rockwell Science Center
Dr. L. Cutler Hewlett-Packard Co.	Dr. S. Stein National Bureau of Standards
Dr. E. EerNisse Quartex, Inc.	Dr. J. Vig U.S. Army, ERADCOM
Mr. M. Frerking Rockwell International	Mr. R. Ward Quartztronic, Inc.
Dr. H. Hellwig Frequency and Time Systems, Inc.	Mr. C. Wardrip NASA - GSFC
Dr. W. Horton Piezo Technology, Inc.	Mr. J. White Naval Research Laboratory
Dr. R. Kagiwada TRW - DSSG	Dr. G. Winkler U.S. Naval Observatory
Mr. D. Kemper Quartztek, Inc.	Dr. N. Yannoni U.S. Air Force, RADC
Mr. J. Kusters Hewlett-Packard Co.	

Association For 1961 1962 1963		<input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
\$ 59.00 <u>NON Member</u> 29.50 <u>Member</u> NEW FEE NEW 1961-1962		
Dist	Spec	
A-1	21	



TECHNICAL SESSION CHAIRMEN

CRYSTAL RESONATOR THEORY I

E.P. EerNisse, Quartex, Inc.

MEASUREMENT ASSURANCE

W. Klepczynski, US Naval Observatory

CRYSTAL RESONATOR THEORY II

A. Ballato, US Army ERADCOM

ATOMIC FREQUENCY STANDARDS

V. Reinhardt, Bendix Corporation

COMPOSITE & OTHER RESONATORS

T. Parker, Raytheon Company

TIME DISSEMINATION AND TRANSFER

J. Oaks, US Naval Observatory

POST-DEADLINE PAPERS

N. Yannoni, US Air Force, RADC

PROPERTIES OF QUARTZ CRYSTALS

R. Laudise, Bell Laboratories

SIGNAL PROCESSING TECHNIQUES

F. Walls, National Bureau of Standards

MILITARY SPECIFICATIONS

V. Rosati, US Army ERADCOM

FILTERS

L. Conlee, Motorola, Inc.

OSCILLATORS

R. Kinsman, Motorola, Inc.

SAW DEVICES

E. Staples, Rockwell International

RESONATOR PROCESSING

R. Ward, Quartztronic, Inc.

CRYSTAL RESONATOR MEASUREMENTS

D. Kemper, Quartztek, Inc.

QUARTZ MATERIAL & RESONATOR MEASUREMENTS

A. Kahan, US Air Force, RADC

TABLE OF CONTENTS:

In Memoriam	1
Awards	3
Plenary Session	
History of Atomic Frequency Standards	
- N.F. Ramsey, Mount Holyoke College and Harvard University	5
Pulsars: Nature's Most Precise Clocks	
- J.H. Taylor, Princeton University	6
Atomic Frequency Standards	
Application of Metal Hydrides for Gas Handling in Hydrogen Masers	AD-P002 444
- H.T.M. Wang, Hughes Research Laboratory	7
Compact Rectangular Cavity for Rubidium Vapor Cell Frequency Standards	AD-P002 445
- H.E. Williams, T.M. Kwon and T. McClelland, Litton Guidance and Control Systems	12
Improved Vibration Performance in Passive Atomic Frequency Standards by Servo-Loop Control	AD-P002 446
- T.M. Kwon and T. Hahn, Litton Guidance and Control Systems	18
Ultra-Stable Laser Clock, Second Generation	AD-P002 447
- R.L. Facklam, Wright-Patterson Air Force Base	21
Microprocessor for the NR Series of Hydrogen Maser Frequency Standards	AD-P002 448
- E.E. Mengel, Applied Physics Laboratory and D.W. Stover, Bendix Field Engineering Corporation	27
Mercury-199 Trapped Ion Frequency Standard: Recent Theoretical Progress and Experimental Results	AD-P002 449
- L.S. Cutler, R.P. Giffard and M.D. McGuire, Hewlett-Packard	32
Laser Cooled $^9\text{Be}^+$ Accurate Clock	AD-P002 450
- J.J. Bollinger, W.M. Itano and D.J. Wineland, National Bureau of Standards	37
Performance of Compact Hydrogen Masers	AD-P002 451
- A. Kirk, Jet Propulsion Laboratory	42
Cold Hydrogen Maser Research at SAO and Related Developments	AD-P002 452
- R.F.C. Vessot, E.M. Mattison and E. Imbier, Smithsonian Astrophysical Observatory and Harvard College Observatory	49

TABLE OF CONTENTS (Continued)

Time Dissemination and Transfer

AD-P002 453

National and International Time and Frequency Comparisons

- D.W. Allan, National Bureau of Standards.55

An International Time Transfer Experiment

AD-P002 454

- C. Wardrip, NASA, J. Buisson, O. Oaks, M. Lister and S. Stebbins, NRL, B. Guinot, M. Granveaud, G. Freon and B. Dubois, Bureau International de L'Heure, W. Schluter and K. Nottarp, Institut Fur Angewandte Geodasie, V. Reinhardt, R. Kruger, P. Dachel, E. Detoma, J. Ingold, R. Roloff and T. Stalder, Bendix Field Engineering Corporation61

Precision Timekeeping at the Observatory Lustuehl, Graz, Austria

AD-P002 455

- D. Kirchner, Technical University Graz.67

Time Dissemination from the National Research Council of Canada

AD-P002 456

- C.C. Costain, H. Daams, J.-S. Boulanger and R.J. Douglas, NRC of Canada78

Signal Processing Techniques

AD-P002 457

Superconductive Tapped Delay Line for Low-Insertion-Loss Wideband Analog Signal-Processing Filters

- R.S. Withers, Lincoln Laboratory, MIT, and P.V. Wright, RF Monolithics81

Low Phase Noise Multiple Frequency Microwave Source

AD-P002 458

- J.T. Haynes, H. Salvo and R.A. Moore, Systems Development Division, and B.R. McAvoy, R&D Laboratories, Westinghouse.87

Synchronization and Tracking with Synchronous Oscillators

AD-P002 459

- V. Uzunoglu, Fairchild Communications and Electronics Company91

Frequency (Standard) Combiner Selector

AD-P002 460

- V.S. Reinhardt, Bendix Field Engineering Corporation, and R.J. Costlow, Applied Physics Laboratory97

Properties of Quartz and Other Crystals

AD-P002 462

Non Destructive Observation of Random Electrical Twinning in Cultured Quartz

- H. Merigoux, J.F. Darces and P. Zecchini, Laboratoire de Cristallographie et Syntheses Minerales, and J. Lamboley, Q.E. Alcatel111

Tensile Fracture Strength of ST Cut Quartz

AD-P002 462

- H.L. Chao and T.E. Parker, Raytheon.116

Effect of Alkali Ions on Electrical Conductivity and Dielectric Loss of Quartz Crystals

- J. Toulouse, E.R. Green and A.S. Nowick, Columbia University.125

AD-P002 463

TABLE OF CONTENTS (Continued)

<i>AD-P002 477</i>	Temperature Derivatives of Elastic Stiffnesses Derived from the Frequency-Temperature Behavior of Quartz Plates P.C.Y. Lee and Y.K. Yong, Princeton University.....	200
<i>AD-P002 478</i>	Transient Thermally Induced Frequency Excursions in AT- and SC-Cut Quartz Trapped Energy Resonators D.S. Stevens and H.F. Tiersten, Rensselaer Polytechnic Institute.....	208
<i>AD-P002 479</i>	Excess Noise in Quartz Crystal Resonators J.J. Gagnepain and M. Olivier, Universite de Franche and F.L. Walls, National Bureau of Standards.....	218
<i>AD-P002 480</i>	Thickness Modes in Circular AT-Cut Quartz Plates with Circular Electrodes W.D. Beaver, Comtec Economation, Inc.....	226
<i>AD-P002 481</i>	Properties of AT Quartz Resonators on Wedgy Plates J. Dworsky, Motorola, Inc.....	232
<i>AD-P002 482</i>	Acoustically Coupled Resonators: Filters and Pressure Transducers J. Detaint, H. Carru, P. Amstutz and J. Schwartzel, Centre National D'Etudes des Telecommunications.....	239
<i>AD-P002 483</i>	Force Sensor Using Double-Ended Tuning Fork Quartz Crystals S.S. Chuang, Statek Corporation.....	248
<i>AD-P002 484</i>	Practical Considerations for Miniature Quartz Resonator Force Transducers E.P. EerNisse, Quartex, Inc. and J.M. Paros, Paroscientific, Inc.....	255
<i>AD-P002 485</i>	Chemically Polished High Frequency Resonators W.P. Hanson, Piezo Crystal Company.....	261
<i>AD-P002 486</i>	Further Studies on the Acceleration Sensitivity of Quartz Resonators R.L. Filler, J.A. Kosinski and J.R. Vig, US Army ERADCOM.....	265
<i>AD-P002 487</i>	Determination of Acceleration Sensitivity of Bulk Mode Resonator Plates H. Rossman and J.T. Haynes, Westinghouse Corporation.....	272
<i>AD-P002 488</i>	Crystal Resonator Measurements An Instrument for Automated Measurement of the Angles of Cut of Doubly Rotated Quartz Crystals J.L. Chambers, Advanced Research and Applications Corporation.....	275
<i>AD-P002 489</i>	A Measurement Technique for Determination of Frequency vs Acceleration Characteristics of Quartz Crystal Units D.J. Healey III, H. Hahn and S. Powell, Westinghouse Electric Corporation.....	284

TABLE OF CONTENTS (Continued)

Radiation Induced Transient Acoustic Loss in Quartz Crystals - D.R. Koehler and J.J. Martin, Sandia Laboratories	AD-PC02 464 130
The Bulk Acoustic Wave Properties of Lithium Tetraborate - C.D.J. Emin and J.F. Werner, GEC Research Laboratories	AD-PC02 465 136
Sputtered c-Axis Inclined Piezoelectric Films and Shear Wave Resonators - J.S. Wang, K.M. Lakin and A.R. Landin, Iowa State University	AD-PC02 466 144
Recalibration of Q Capability Indications from Infrared Measurements on Cultured Quartz - B. Sawyer, Sawyer Research Products, Inc.	AD-PC02 467 151
Pressure-Volume-Temperature Behavior in the System $H_2O-NaOH-SiO_2$ and its Relationship to the Hydrothermal Growth of Quartz - E.D. Kolb, P.L. Key and R.A. Laudise, Bell Laboratories and E.E. Simpson, Western Electric	AD-PC02 468 153
Standard Characterization Methods for the Determination of the Quality of Hydrothermally Grown Quartz - J.F. Balascio and N.C. Lias, Motorola, Inc.	AD-PC02 469 157
Electrodiffusion of Charge-Compensating Ions in Alpha-Quartz - J.J. Martin, R.B. Bossoli, L.E. Halliburton, B. Subramaniam and J.D. West, Oklahoma State University	AD-PC02 470 164
Aluminum and Hydroxide Defect Centers in Vacuum Swept Quartz - H.G. Lipson, A. Kahan and J. O'Connor, US Air Force, RADC	AD-PC02 471 169
Effect of Seed Treatment on Quartz Dislocations - A.F. Armington, J.J. Larkin, J. O'Conner, J.E. Cormier and J.A. Horrigan, US Air Force, RADC	AD-PC02 472 177
Thermoluminescence from Different Growth Sectors in Synthetic Quartz Crystals - S. Katz, A. Halperin and M. Ronen, Hebrew University of Jerusalem	AD-PC02 473 181
Characterization of Quartz Crystals by Cathodoluminescence - S. Katz, A. Halperin and M. Schieber, Hebrew University of Jerusalem	AD-PC02 474 185
Improvements of Laser Interferometric Measurement System of Vibration Displacements - T. Adachi, M. Okazaki and Y. Tsuzuki, Yokohama National University	AD-PC02 475 187
Crystal Resonator Theory, Design and Processing	AD-PC02 476
The Stress Coefficient of Frequency of Quartz Plate Resonators - M. Mizan and A. Ballato, US Army, ERADCOM	194

TABLE OF CONTENTS (Continued)

AD-P002 490	Evaluation of Crystal Measurement Systems	
	- R.C. Smythe and W.H. Horton, Piezo Technology, Inc.	290
AD-P002 491	A New Frequency for Piezoelectric Resonator Measurement	
	- W.H. Horton and R.C. Smythe, Piezo Technology, Inc.	297
AD-P002 492	Measuring Method of Equivalent Series Capacitance and Negative Resistance of Quartz Crystal Oscillator Circuits	
	- M. Toki and Y. Tsuzuki, Yokohama National University and T. Mitsuoka, Toyo Communication Equipment Co., Ltd.	300
AD-P002 493	S.Y. Parameters Method for Accurate Measurement of Bulk Wave Crystal Resonators at Frequencies up to 2GHz	
	- J.P. Aubry, E. Gerard and S. Lechopier, C.E.P.E.	306
	Composite & Other Resonators	
AD-P002 494	Spurious Resonances and Modelling of Composite Resonators	
	- K.B. Yoo, H. Uberall, D. Ashrafi and S. Ashrafi, Catholic University	317
AD-P002 495	Analysis of Composite Resonator Geometries	
	- K.M. Lakin, Ames Laboratory-USDOE	320
AD-P002 496	An Analysis of Thickness-Extensional Trapped Energy Resonators with Rectangular Electrodes in the Zinc-Oxide Thin Film on Silicon Configuration	
	- H.F. Tierston and D.S. Stevens, Rensselaer Polytechnic Institute	325
AD-P002 497	Miniature LiTaO ₃ X-Cut Strip Resonator	
	- M. Okzakj and S. Watanabe, Nihon Dempa Kogya Co., Ltd.	337
AD-P002 498	Miniaturized LiTaO ₃ Strip Resonator	
	- Y. Fujiwara, S. Yamada and N. Wakatsuki, Fijitsu Limited.	343
	Filters	
	AD-P002 499	
	Current Trends in Crystal Filters	
	- R.C. Smythe and M.D. Howard, Piezo Technology, Inc.	349
	AD-P002 500	
	Surface Acoustic Wave Bandpass Filters	
	- C. Hartmann and S. Wilkus, RF Monolithics, Inc.	354
	EHF Waveguide Filters	
	- J.E. Raue, TRW Electronic Systems	361
	AD-P002 501	
	Digital Filters: An Overview	
	- L. Sheats, Raytheon Company	362

TABLE OF CONTENTS (Continued)

AD-P003 503

Microelectronic Analog Active Filters
R. Schaumann, University of Illinois. 371

AD-P003 503

Timing Tank Mechanical Filter for Digital Subscriber Transmission System
T. Gounji, T. Kawatsu, Y. Kasai and T. Takeuchi, Fujitsu, Ltd. and Y. Tomikawa and
M. Konno, Yamagata University 376

AD-P003 504

Stripline Filters: An Overview
L. Dworsky, Motorola, Inc. 387

AD-P003 505

SAW Devices
Precision L Band SAW Oscillator for Satellite Application
T. O'Shea, V. Sullivan and R. Kindell, Sawtek, Inc. 394

AD-P003 506

SAW Resonator Stabilized Oscillator
C. Bennett, RF Monolithics, Inc. 405

AD-P003 507

Very Long Period Random Frequency Fluctuations in SAW Oscillators
T. E. Parker, Raytheon Research Division 410

AD-P003 508

Stress Induced Effects on the Propagation of Surface Waves
B. K. Sinha and W. J. Tanski, Schlumberger Doll Research and T. Lukaszek and A. Ballato,
US Army ERADCOM 415

AD-P003 509

Surface Wave Resonators on Silicon
S. J. Martin, R. L. Gunshor, T. J. Miller, S. Datta and R. F. Pierret, Purdue University and
M. R. Melloch, Texas Instruments, Inc. 423

AD-P003 510

High-Frequency Shallow Bulk Acoustic Wave Quartz Frequency Sources
K. V. Rousseau, K. H. Yen, K. F. Lau and A. M. Kong, TRW 428

Oscillators

AD-P003 511

A New Digital TCXO Circuit Using a Capacitor-Switch Array
T. Uno and Y. Shimoda, Nippon Telegraph and Telephone Public Corporation 434

AD-P003 512

An Improved Method of Temperature Compensation of Crystal Oscillators
J. S. Wilson, STC Components 442

AD-P003 513

A New Frequency Temperature Compensation Method for Oscillators
T. Kudo, S. Fujii and S. Nanamatsu, NEC Corporation 448

An Ultra-Stable Crystal Oscillator for Beacons (Program "SARSAT")

A. Stahl, Quartz & Electronique Alcatel and M. Brunet, C.N.R.S. 454

AD-P003 514

TABLE OF CONTENTS (Continued)

<i>AD-P002 515</i>	Modeling of TTL Inverter Amplifiers as Applied to Overtone Crystal Clock Oscillators	
<i>AD-P002 516</i>	D. Nehring, CTS Knights	459
<i>AD-P002 517</i>	MSSW Delay Line Based Oscillators	
<i>AD-P002 518</i>	V. Landes, Airtron Corporation and J.P. Parekh, State University of New York	473
<i>AD-P002 519</i>	The Status of Magnetostatic Wave Oscillators	
<i>AD-P002 520</i>	R.L. Carter and J.M. Owens, University of Texas	477
<i>AD-P002 521</i>	An X-Band GaAs FET Oscillator Using a Dielectric Resonator	
<i>AD-P002 522</i>	K. Waking, T. Nishikawa, S. Tamura and H. Tamura, Murata Mfg. Co.	481
<i>AD-P002 523</i>	An Analysis of a Microprocessor Controlled Disciplined Frequency Standard	
<i>AD-P002 524</i>	B. Bourke and B. Penrod, Austron, Inc.	485
<i>AD-P002 525</i>	Measurement Assurance	
<i>AD-P002 526</i>	BVA Quartz Crystal Resonator and Oscillator Production A Statistical Review	
<i>AD-P002 527</i>	E.P. Graf and U.R. Peier, Oscilloquartz, SA	492
<i>AD-P002 528</i>	Temperature Compensated Crystal Oscillator Survey and Test Results	
<i>AD-P002 529</i>	V. Rosati, S. Schodowski and R. Filler, US Army ERADCOM	501
<i>AD-P002 530</i>	A New Frequency Calibration Service Offered by the National Bureau of Standards	
<i>AD-P002 531</i>	G. Kamas and J.L. Jespersen, National Bureau of Standards	506
<i>AD-P002 532</i>	Military Specifications	
<i>AD-P002 533</i>	Specifying Performance for Atomic Standards	
<i>AD-P002 534</i>	J. White, US Naval Research Laboratory	513
<i>AD-P002 535</i>	Platform Distribution System Specifications	
<i>AD-P002 536</i>	J.A. Murray, US Naval Research Laboratory	516
<i>AD-P002 537</i>	Modernization of the Military Specification for Quartz Crystal Units	
<i>AD-P002 538</i>	R.L. Filler, US Army ERADCOM	519
<i>AD-P002 539</i>	Revision of the Military Specification for Quartz Crystal Oscillators	
<i>AD-P002 540</i>	V. Rosati and S. Schodowski, US Army ERADCOM	524
<i>AD-P002 541</i>	Review of New Military Specification on Surface Acoustic Wave Devices	
<i>AD-P002 542</i>	E.A. Mariani, US Army ERADCOM	525

IN MEMORIAM

Andrew Chi died suddenly on June 19, 1982. During a three decade career, he made important contributions to diverse areas of the time and frequency field - quartz crystal technology, atomic frequency standards and time dissemination via radio broadcast and satellite. A brief record of his career follows.

Andrew R. Chi was born in Tientsin, China on September 12, 1920. He received the B.S. degree in physics and mathematics from Western Maryland College, in 1944 and the M.A. degree in physics from Columbia University, in 1946. From 1946 to 1953 he was an Instructor in physics at the Cooper Union School of Engineering in New York City.

In 1953 he joined the U.S. Army Signal Research and Development Laboratories, Fort Monmouth, as a physicist doing research on the fundamental properties of quartz. His findings contributed to the understanding and development of a method for determining the angle of cut of quartz resonator plates for various operating temperatures. He correlated the change of the resonator properties of quartz to the content of impurities introduced to the crystal structure of quartz during growth. For his findings, he was given a patent on growing synthetic quartz with controlled impurities. The crystals he made found applications as filters in communication equipment.

In 1957 he joined the Naval Research Laboratory, where he was engaged in studies of atomic resonance devices and precision frequency standards. He collaborated with physicists from the National Bureau of Standards in developing the atomic gas cell frequency standard using optical pumping techniques. He and his colleagues were the first to measure the hyperfine transition frequencies of rubidium-87 and cesium-133 atoms. They were also the first to measure the frequency shift due to pressure and temperature of various buffer gases in a gas cell frequency standard. While at the Naval Research Laboratory he also maintained the proper operation of the cesium beam atomic frequency standard whose output frequency was measured in terms of Universal Time by astronomers and scientists at the U.S. Naval Observatory, Washington, D.C., and the National Physical Laboratory in Teddington, England.

In 1963 he joined NASA's Goddard Space Flight Center and became Head of the Timing Systems Section. In this position, he planned and conducted research and development of advanced timing systems for the manned and unmanned satellite tracking networks. Mr. Chi was the first to see the need for and became the prime mover in organizing a symposium for the purpose of defining the short term frequency stability. This symposium was organized and held at Goddard Space Flight Center in November 1964. Because of the wide interest, the Institute of Electrical and Electronics Engineers (IEEE) joined the National Aeronautics and Space Administration (NASA) in sponsoring the Symposium on Short Term Frequency Stability. Mr. Chi was the Chairman of both the symposium and the Technical Program Committee. He was also the editor of the symposium proceedings.

During the planning stage of the Apollo program he served as a member of the Apollo Navigation Working Group and designed the timing system for the Manned Space Flight Network. He served as a timing expert in the Geodetic Satellite Program; as a tracking scientist for the OMEGA Position Location Experiment, as Chairman of the Timing Standards Subcommittee of the Data Systems Requirements Committee; and was NASA's representative on the Timing Committee of DOD's Inter Range Instrumentation Group. His technical contributions during this period included NASA's time code standards and a time synchronization technique using dual very low frequency transmissions.

From 1971 to his death, Mr. Chi was a senior staff engineer of the Network Engineering Division. In this position he was responsible for planning and coordinating all programs pertaining to frequency and time for use by NASA's worldwide satellite tracking and data networks. He was the principal investigator of a joint experiment between NASA and the Federal Aviation Administration on precise time transfer via a synchronous satellite. He used the two-way time transfer technique to measure the round trip propagation delay and demonstrated a system capability of transferring time to an accuracy of 50 nanoseconds.

He was the manager of a research technology operating plan on network timing and synchronization technology. As such he planned, developed, and coordinated time synchronization techniques for NASA users. He represented NASA in meetings with other Government agencies on matters pertaining to frequency control, atomic oscillators, time synchronization, and time maintenance. He also represented the United States at international meetings for example in Study Group 7 of the International Radio Consultative Committee of the International Telecommunication Union.

He was a Fellow of the Institute of Electrical and Electronics Engineers, a member of the American Physical Society and a past member of the American Association of University Professors, American Association of Physics Teachers, Research Society of America, and Philosophical Society of Washington, DC. In addition to the Distinguished Service Award, which he received in 1965, Mr. Chi was awarded the Scroll of Appreciation of the Secretary of State for his contributions to CCIR (1967), the Apollo Achievement Award (1969), the Skylab Achievement Award (1974), and the 1978 Moe I. Schneebaum Memorial Award for his significant contributions to the establishment of an international measure of frequency stability. He was given the 1980 Achievement Award by the Chinese Engineers and Scientists Association of Southern California for his contribution to international standards of frequency and time measurement.

Awards

Prior to 1983, only one award, the C.B. Sawyer Memorial Award, had been presented at the Frequency Control Symposium. According to the Sawyer Award announcement, the Award is "to consist of \$100.00 and a plaque, shall be made on an annual basis to the person, or the group of persons, who, in the opinion of an independent three-man judging committee, has made the most outstanding recent contribution to advancement in the field of quartz crystals and devices. No award will be made in a year in which the committee determined that no award is warranted. Presentations will be made at the Annual Frequency Control Symposium."

Aside from permitting the presentation of the Sawyer Award at the Symposium Banquet, the Symposium's management has had no involvement with the Sawyer Award. The Award is sponsored by Sawyer Research Products, Inc., the judging committee is selected by that company. The first Sawyer Award was presented in 1983. It has been presented each year since then, except in 1982, when no suitable award nominations were received.

For several years, sentiment had been expressed at Symposium program committee meetings for the creation of awards that can recognize outstanding contributions in all fields covered by the Symposium, not just in the field of quartz crystals and devices. Therefore, in early 1983, the program committee voted to create two new awards. One, the Cady Award, named after Walter Gulton Cady, is to recognize outstanding contributions related to piezoelectric frequency control devices. The other, the Rabi Award, named after Prof. I.I. Rabi, is to recognize outstanding contributions related to fields such as atomic and molecular frequency standards, and time transfer and dissemination. Each award consists of \$100.00, and a limited edition bronze plaque, print and certificate in a leather binder. The awards are presented to the recipients at the Symposium Banquet.



AWARD WINNERS: Dr. E. Hafner, Prof. I.I. Rabi, and Dr. E. P. EerNisse

THE RABI AWARD

The first recipient of the Rabi Award was Prof. I.I. Rabi: "For theoretical and experimental contributions to atomic beam resonance spectroscopy leading to the development of practical atomic frequency standards."

THE CADY AWARD

The first recipient of the Cady Award was Dr. Errol P. EerNisse: "For his theoretical prediction of planar stress compensation in doubly rotated quartz plate resonators leading to the realization of the SC-cut."

THE SAWYER AWARD

The Sawyer Award recipient was Dr. Erich Hafner: "For technical contributions and leadership in the fields of quartz resonator research, technology and measurement, and high precision frequency control."

PREVIOUS SAWYER AWARD WINNERS:

1966 - Dr. Warren P. Mason: "For outstanding contributions in quartz crystal devices, particularly in the field of frequency selection" and

Dr. Rudolf Bechman: "For outstanding contributions in quartz crystal devices, particularly in the field of frequency control."

1967 - Professor Raymond D. Mindlin: "For fundamental contributions to the theory of vibration in piezoelectric resonators leading directly to advancements in the art."

1968 - Daniel R. Curran: "For original and imaginative design of multielectrode piezoelectric resonators, contributing significantly to the rapid advance of the quartz filter art in the past few years" and

Dr. David B. Fraser: "For contributions to the knowledge of the mechanisms of acoustic loss in crystalline quartz, and the evaluation of this acoustic loss by optical methods."

1969 - Arthur W. Warner, Jr.: "Contributions to the development of high frequency thickness shear quartz resonators for precise frequency control and as an aid to the measurement of the intrinsic Q of quartz material."

1970 - Dr. Issac Koga: "Theoretical and experimental investigations of quartz and tutorial leadership in the field of piezoelectric crystals."

1971 - Donald L. Hammond: "For development and applications of crystal devices to highly precise frequency control, and temperature and pressure instrumentation."

1972 - Dr. W.J. Spencer: "For advances in the theory and development of piezoelectric crystal devices."

1973 - Dr. James C. King: "For major contributions to the understanding of the fundamental properties of quartz crystals, and methods for improvement of these properties in synthetic quartz."

1974 - Dr. Robert A. Laudise, Robert A. Ballman and David W. Rudd: "For outstanding contributions to the synthesis of crystalline quartz with special properties for resonator applications."

1975 - Dr. Morio Onoe: "For theoretical and practical contributions in the field of frequency control and selection, as well as leadership in national and international committees on piezoelectric devices."

1976 - Warren L. Smith: "For outstanding contributions in the field of precision crystal controlled oscillators of high spectral purity and monolithic crystal filters."

1977 - Dr. Virgil E. Bottom: "In recognition of theoretical and practical contributions to the Quartz Crystal Industry, and inspiration to his students to choose this field of endeavor."

1978 - Dr. Arthur D. Ballato: "For contributions in the field of piezoelectric crystals such as; stacked filters, electric circuit analogues and stress effects in doubly rotated plates."

1979 - Dr. Harry F. Tiersten: "For contributions to the theory of piezoelectric resonators."

1980 - Dr. Peter Chung-Yi Lee: "For contributions to the theory of vibrations in quartz crystal plates."

1981 - Dr. Eduard A. Gerber: "For pioneering research in VHF and UHF precision oscillators and filter crystals and international leadership in the field of frequency control" and

Roger A. Sykes: "For outstanding contributions in the development and application of quartz crystals in the frequency control industry."

1982 - No award given in this year.

HISTORY OF ATOMIC FREQUENCY STANDARDS

Norman F. Ramsey

Mount Holyoke College
South Hadley, MA

and

Harvard University
Cambridge, MA

The history of atomic frequency standards was reviewed with special emphasis on the methods that are currently the most stable and most widely used. The subjects discussed included the molecular beam magnetic resonance methods, atomic beam frequency standards, molecular masers, atomic masers, lasers, trapped ions, superconducting cavities, doppler broadening and future prospects. The full article will be published in the September 1983 issue of the National Bureau of Standards Journal of Research.

PULSARS: NATURE'S MOST PRECISE CLOCKS

Joseph H. Taylor
Joseph Henry Laboratories
Physics Department, Princeton University
Princeton, New Jersey 08544

When pulsars were discovered by radio astronomers in 1967, it was soon evident that these objects are natural clocks of remarkably high precision. The signal received from a typical pulsar consists of bursts of radio frequency noise, recurring at periodic intervals of half a second or so, and with pulse widths equal to a few percent of the period. If the signal is averaged synchronously for several hundred periods, thereby producing a stable waveform with enhanced signal-to-noise ratio, the phase of the pulsar "clock" can be measured to a precision of $\sim 10^{-3}$ periods. Thus, typical pulsars can be timed with measurement uncertainties of $\lesssim 500 \mu\text{s}$ after a few minutes of observation. The use of very large radio telescopes, together with sophisticated signal processing techniques and longer averaging times, can improve accuracies even further.

Pulsar timing observations have now been accumulated for more than a decade, and the data show that:

- (1) Nearly all pulsars exhibit timing accuracies (after correction for a constant period derivative) of a few milliseconds or better, over timescales exceeding ten years.
- (2) Pulsars with the most regular timing behavior tend to be those with the smallest period derivatives. In some cases the pulse timing residuals over a twelve year span are entirely consistent with random measurement errors at the level of $200 \mu\text{s}$ or less.

At these levels of accuracy the pulsars are already impressive clocks -- but they do not yet challenge the accuracies of civil timekeeping standards. However, the discovery in late 1982 of an unusually fast pulsar (period = 1.558 ms) may radically change this situation. Because the measurement uncertainties tend to be a fixed fraction of a period, accuracies of $\sim 1 \mu\text{s}$ are achievable for this object in a few minutes of observing time. Furthermore, this pulsar has the smallest period derivative yet measured for any pulsar, approximately $10^{-19} \text{ s s}^{-1}$, so if the well-established generalization (2) applies, the millisecond pulsar is likely to be a superb clock over very long timescales. After four months of observations, post-fit timing residuals (pulsar clock minus atomic time) amount to less than $2 \mu\text{s}$ rms, and are presently dominated by non-random instrumental errors. We believe that these errors can be reduced by at least a factor of 10, and thus -- over timescales of a year or longer -- that this pulsar may indeed challenge the accuracy of the best man-made timekeeping and time distribution systems. The potential applications of pulsars as time standards are many, and include observational cosmology, the astrophysics of neutron star interiors, basic gravitational physics, solar system dynamics, and, finally, very practical independent checks on the accuracy of terrestrial atomic clocks.

Harry T. M. Wang
Hughes Research Laboratories
Malibu, CA 90265

Summary

Reversible interaction of hydrogen with certain metals and alloys have been employed to effectively satisfy the hydrogen supply and flow regulation, as well as vacuum maintenance requirements of a hydrogen maser. In addition to providing significant reductions in maser size, weight and power consumption, the hydride components also lead to improved system reliability. A hydrogen supply for seven years of operation of an oscillating compact maser can be stored in 50 grams of LaNi_5 or $\text{LaNi}_{4.7}\text{Al}_{0.3}$. The source occupies a volume of less than 50 cc, including an electrically controlled palladium-silver alloy flow regulator. A combination getter-ion vacuum pump system for the maser was developed using a zirconium graphite getter. Operational experience indicates that for reliable maser operation, a clean, bakeable vacuum system design is a critical requirement.

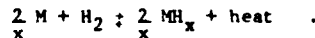
Introduction

The operation of the hydrogen maser requires a steady supply of state selected hydrogen atoms. Furthermore, signal-to-noise ratio and atomic transition linewidth considerations dictate that the atomic resonance be observed in a high vacuum environment. In conventional maser designs atomic hydrogen is obtained by dissociating hydrogen molecules in a radio frequency discharge. The supply of molecular hydrogen is typically stored in a vessel under high pressures. A sputtering ion pump is used to pump away the spent gas and other outgassing products to maintain the required high vacuum for maser operation. Both the pressure vessel (and the associated flow regulator) and the ion pump having the capacity to handle the gas load alone are bulky components which constitute a significant portion of the size and weight of the maser. For field operable devices, and especially devices for spaceborne applications, size, weight and power consumption are important considerations. Moreover, the ion pump has been trouble prone, causing reliability problems. Indeed, most maser failures are traceable to ion pump malfunctions. It is therefore desirable to eliminate the ion pump, or at least minimize its role in maser vacuum maintenance.

Many metals and alloys react reversibly with hydrogen to form compounds termed hydrides. The great affinity for hydrogen and other properties of selected metals and alloys can be exploited to provide effective remedies for gas handling and vacuum maintenance requirements in the atomic hydrogen maser. In this paper, we will discuss a hydrogen supply and a vacuum pump for the maser employing hydride interaction.

Hydrogen Storage Using Hydrides

The technique of hydrogen storage in hydrides makes use of the reversible reaction of a solid metal, M, with gaseous hydrogen, H_2 , to form a solid metal hydride, MH_x , according to the equation,



The reversible reaction means that hydrogen can be stored or discharged at will, analogous to a water sponge or a rechargeable electric battery. At a given temperature, T, the equilibrium dissociation pressure, P, for a charged hydride is relatively constant, at least in theory.¹ In practical storage media, the dissociation pressure may depend on the hydrogen content, as shown by the hydrogen absorption and desorption characteristics of a 50 gram sample of $\text{LaNi}_{4.7}\text{Al}_{0.3}$ at room temperature in Figures 1 and 2. On the other hand, the dissociation pressure is a sensitive function of temperature and is described by the van't Hoff equation,

$$\frac{d(\ln P)}{d(1/T)} = -\frac{\Delta H}{R}$$

where ΔH is the heat of reaction and R is the gas constant. For a specific application, the storage medium is selected to satisfy working temperature and pressure requirements.

Recent metallurgical advances had made a variety of media suitable for hydrogen storage readily available. Some hydrides and their hydrogen storage characteristics are shown in Table I. Hydrogen density in a gas at a pressure of 100 atmospheres and in liquid hydrogen are also shown for comparison. It is

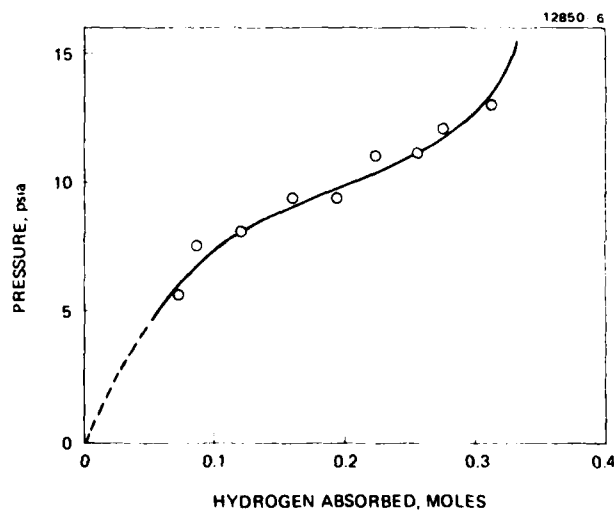


Figure 1. Hydrogen storage in $\text{LaNi}_{4.7}\text{Al}_{0.3}$: Absorption (50 grams sample).

*This work has been supported by the Naval Research Laboratory under contracts N00014-82-C-2016 and N00014-83-C-2023.

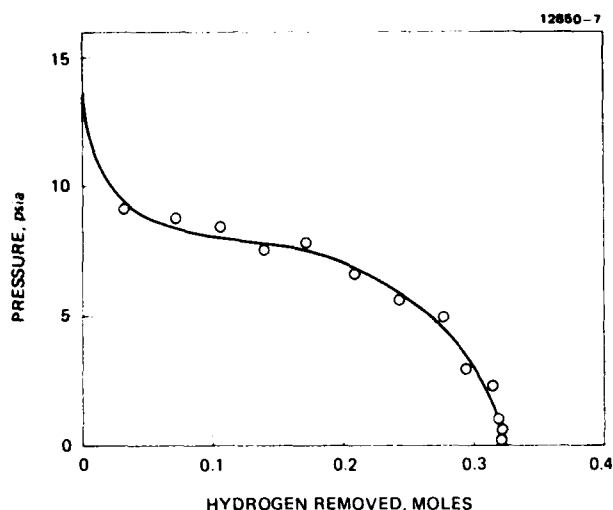


Figure 2. Hydrogen storage in $\text{LaNi}_{4.7}\text{Al}_{0.3}$: Desorption (50 grams sample).

TABLE 1. Hydrogen Storage in Metal Hydrides

MEDIUM	HYDROGEN CONTENT		DISSOCIATION PRESSURE AT 25 °C, atm	HEAT OF REACTION H, kcal/mole-H
	wt% H	H-atom/mL $\times 10^{-22}$		
U - UH_3	1.3	8.3	2.5×10^{-9}	-20.2
Mg - MgH_2	7.6	6.7	5.5×10^{-7}	-18.5
LaNi_5 - LaNi_5H_6	1.5	7.6	1.65	-7.4
$\text{LaNi}_{4.7}\text{Al}_{0.3}$ - $\text{LaNi}_{4.7}\text{Al}_{0.3}\text{H}_5$	1.4	7.2	0.42	-8.1
H_2 GAS AT 100 atm	100	0.5	-	-
H_2 LIQUID	100	4.2	-	-

interesting to note that volume for volume, the solid hydrides have hydrogen contents greater than that in liquid hydrogen. Needless to say, the hydride storage is more energy efficient and incomparably safer than liquid hydrogen. Uranium hydride has been used to provide spectroscopically pure hydrogen samples. However, since both uranium and uranium hydride are pyrophoric in finely divided form, they present handling problems. Besides, to obtain hydrogen at reasonable pressures, the hydride has to be heated to high temperatures. The requirement of elevated temperatures also applies to magnesium and other similar binary hydrides which otherwise provide a very favorable hydrogen density to weight ratio.

Several hydrides have been used to provide a hydrogen supply for the maser.² We have chosen to investigate mischmetal hydrides as possible maser hydrogen storage media. The technique has a number of desirable features:

(1) high density hydrogen storage, leading to a compact and lightweight package; (2) automatic purification of hydrogen during the charging and discharging cycle; and (3) dissociation occurring at room temperature at a pressure of the order of one atmosphere. Thus, neither a bulky mechanical regulator (for high gas pressure) nor a power consuming heater (for hydrides with too low a dissociation pressure) is needed; and (4) the discharge is endothermic, making hydride storage inherently safer than pressure vessel storage. In addition, due to the low hydrogen flow rate in the maser, no external heater is required.

Our experimental investigations centered on LaNi_5 (HY-STOR 205) and $\text{LaNi}_{4.7}\text{Al}_{0.3}$ (HY-STOR 207).³ The capacity of these hydrides is such that the hydrogen supply for 7 years operation of an oscillating compact hydrogen maser can be stored in about 50 grams of the materials, occupying a volume of about 7 cc. Hydrogen storage characteristics for a 50 gram sample of $\text{LaNi}_{4.7}\text{Al}_{0.3}$ are shown in Figures 1 and 2. The data were taken at room temperature. A prototype hydrogen supply for the maser using LaNi_5 was successfully tested on an operating maser. Even though the dissociation pressure of LaNi_5H_6 is only 1.6 atmosphere at 25°C, the exponential dependence on temperature according to the van't Hoff equation means that it will increase to 20 atmosphere at 100°C. To simplify container design, a hydride with lower dissociation pressure would be more desirable. HY-STOR 207 ($\text{LaNi}_{4.7}\text{Al}_{0.3}$) was found to satisfy the requirements. The hydride can be charged at relatively low pressures (about 1 to 2 atmospheres at room temperatures). The dissociation pressure of the hydride at 25°C is only about 0.4 atmospheres, increasing to about 7 atmospheres at 100°C. A hydrogen supply system for the maser including an integrated palladium-silver alloy flow regulator is shown in Figure 3.

The Pd-Ag flow regulator is another example of the application of hydriding interaction to maser gas handling requirements. Hydrogen permeates the wall of the Pd-Ag tube. The flux, J, per unit area, A, is proportional to the concentration gradient, dc/dr of hydrogen in the alloy.

$$J/A = -Ddc/dr$$

The diffusion coefficient, D, obeys Arrhenius relation,

$$D = D_0 \exp(-E_a/kT)$$

For Pd-Ag (80/20) alloy, $D_0 = 3.4 \times 10^{-3} \text{ cm}^2\text{s}^{-1}$ (Ref. 1), and the activation energy $E_a = 5.35 \text{ kcal/gm-atom}^{-1}$. The exponential dependence on temperature of the hydrogen permeation rate is conveniently used to provide an electrically controlled hydrogen flow for the maser.

Getter Pump for the Maser

Chemical getters have been widely used in high vacuum systems and electronic tubes. Until recently, they have not been seriously considered for use in the maser. This is probably due to the fact that the more familiar evaporable getters are inconvenient to use and have very limited capacity.

During the last several years, we have investigated the suitability of non-evaporable getters for maser applications. The aim of our experimental investigations is to gather data on getter pump characteristics and to determine maser vacuum system design criteria.

The non-evaporable zirconium graphite (ST-171)⁴ getter we shall discuss here has very attractive features: (1) tremendous capacity and pumping speed for hydrogen; and (2) hydrogen gettering (hydriding), after activation, occurs at room temperature without any power consumption. These features will lead to a compact and lightweight vacuum pump. On the other hand, the pumping speed of the getter for other outgassing products in the system at room temperature is very limited. The requirement of the maser is best met by a combination vacuum pump system consisting of a getter and a small ion pump. The

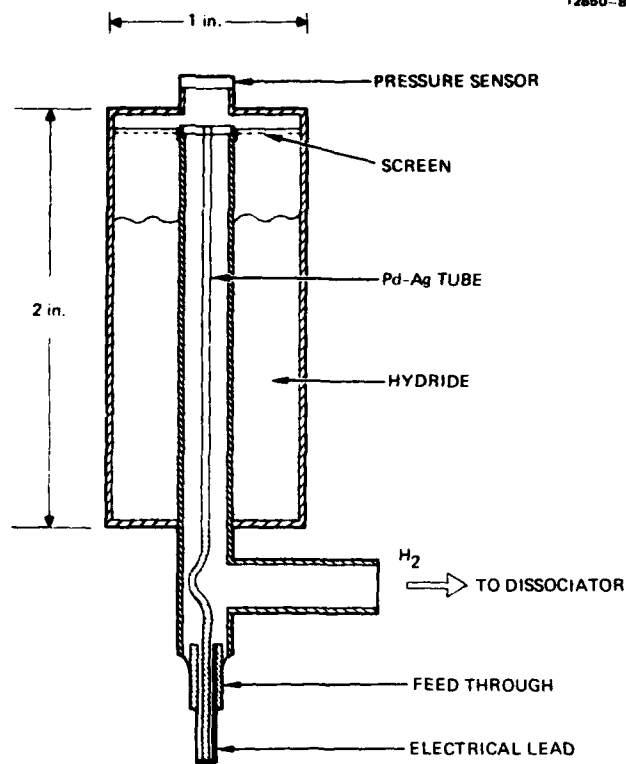


Figure 3. Schematic of a maser hydrogen supply with an electrically controlled Pd-Ag flow regulator.

getter handles the dominant hydrogen gas load while the ion pump evacuates the non-getterable contaminants.

A schematic of the getter pump developmental test station is shown in Figure 4. Our zirconium graphite getter test sample is in the form of a hollow cylinder, 40 mm O.D. x 25 mm I.D. x 46 mm long and weighing 100 grams. It is housed in a quartz container attached to an all-metal vacuum system using a flanged moly-quartz transition. The transparent container allows visual observation of the physical condition of the getter. The test station is provided with an ionization gauge for pressure monitoring, a quadrupole mass spectrometer for residual gas analysis and a 2 liter/sec appendage ion pump for an evaluation of the ion pump behavior under large hydrogen flow conditions. Molecular hydrogen is admitted to the test chamber through a voltage controlled Pd-Ag alloy flow regulator. Hydrogen flow rate is determined from pressure changes in a reservoir of known volume. A capacitance manometer with a digital readout is used for precise reservoir pressure reading. A liquid nitrogen roughing pump is used for initial pump down through a bakable metal sealed valve. Prior to getter activation and data collection, the system is baked out under high vacuum. A combination ion and titanium sublimation pump is used during high vacuum bake out and subsequent getter activation.

An effective activation procedure was established as follows. Using an external oven, the getter is heated to about 925°C under high vacuum. The temperature is raised gradually so that getter outgassing does not overload the ion-titanium sublimation pump system. After reaching 925°C, the temperature is held constant

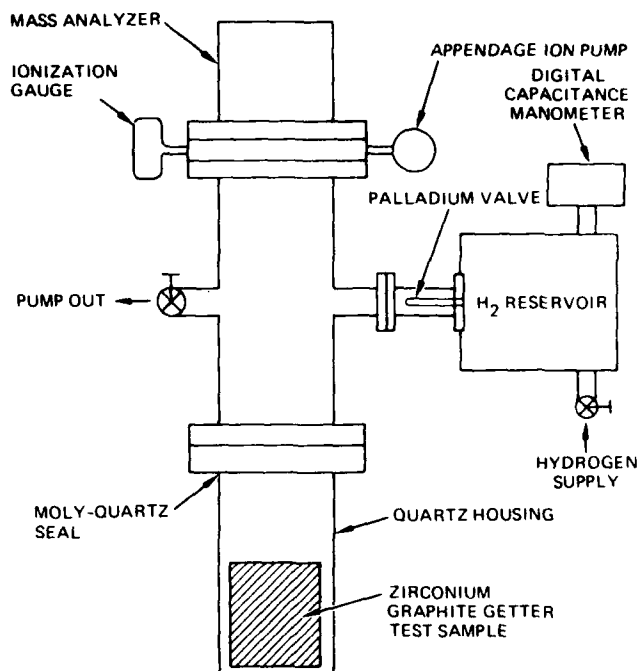


Figure 4. Schematic of a getter pump test station.

until residual pressure in the test chamber is about 10^{-6} Torr or less. This usually requires one working day to complete. By comparison, the manufacturer's recommended procedure could be completed in a fraction of an hour after attaining the 900°C activation temperature. Our experience is that the latter procedure does not provide optimum pumping speed and capacity but is probably adequate if repeated activation is desired.

In order to obtain the desired capacity data within a reasonable interval, hydrogen throughput is accelerated to about 30 times the normal flow rate in an oscillating compact maser of about 2.3×10^{-5} Torr-liter/sec. Data for a 100 gram zirconium graphite getter sample is shown in Figure 5. Some interesting observations are:

- (1) The getter has tremendous capacity for hydrogen. The 100 gram sample can pump more than 10^4 Torr-liters of hydrogen. This is to be compared with an estimated hydrogen consumption of about 5×10^3 Torr-liters in seven years of normal oscillating compact maser operation.
- (2) The initial pumping speed is very high. Even though it falls off with the amount of hydrogen pumped, the pump can be designed to easily meet maser requirements.
- (3) The deviation from exponential decrease in pumping speed after pumping about 1500 Torr-liters of hydrogen is probably due to enlarged surface area of the getter due to cracking and flaking. Powder formation was observed after the getter absorbed a few hundred Torr-liters of hydrogen. A photograph of the getter after absorbing about 5000 Torr-liters of hydrogen is shown in Figure 6. Powder confinement is a critical requirement for maser getter pump design.

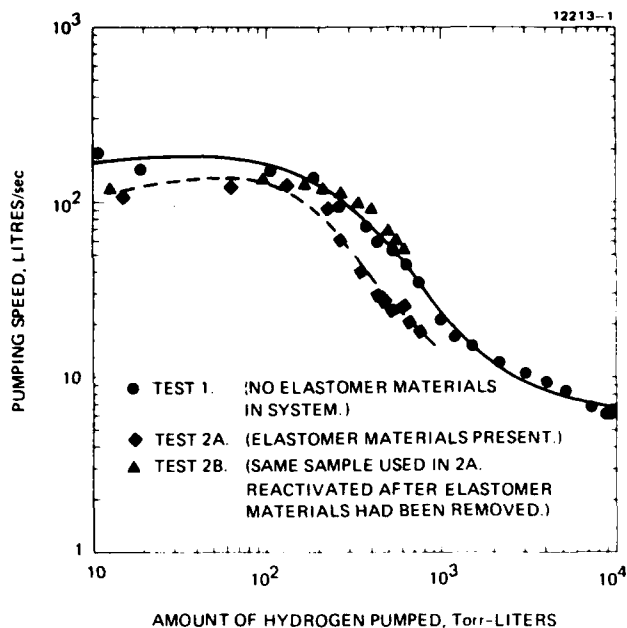


Figure 5. Zirconium-graphite getter pump characteristics, with or without the presence of elastomers in the system.



Figure 6. Photograph of a 100 gram zirconium-graphite getter sample after pumping about 5000 Torr-liters of hydrogen.

(4) Elastomers have a deleterious effect on getter pump operation, as can be seen from a comparison of data for Tests 2A and 2B shown in Figure 5. Test 2A used a fresh getter sample. A viton O-ring and a small sheet of reflen were purposely inserted in the vacuum chamber. These elastomers were chosen because they are used in conventional maser designs. Data were taken only after bake out at about 100°C for one day followed by a few days observation, with no hydrogen gas load, to ascertain that the system has completed the initial high rate of outgassing. The lower pumping speed of the getter in the presence of elastomers is obvious from the data shown in Figure 5. To confirm this result, the elastomers were removed from the test chamber. The same getter sample was reactivated by the normal procedure and its pumping characteristics measured. The results obtained denoted as Test 2B are in excellent agreement with those obtained in Test 1.

(5) Despite the large hydrogen throughput, the 2 liter/sec ion pump operated normally in the clean test station. That is, there were no significant pump current instabilities, as can be seen in Figure 7 which shows the ion pump current and ion gauge readings during the course of a test. Since both readings are proportional to the system pressure, differing by a calibration factor, the tracking of the two sets of readings, and therefore the stability of the ion pump current, is indicated by their difference shown in the bottom curve of Figure 7. (The ion pump was not operating during Test 2A and 2B to avoid masking the deleterious effect of the elastomers we were trying to observe. As we will see below, the elastomers are probably detrimental to ion pump operation as well.)

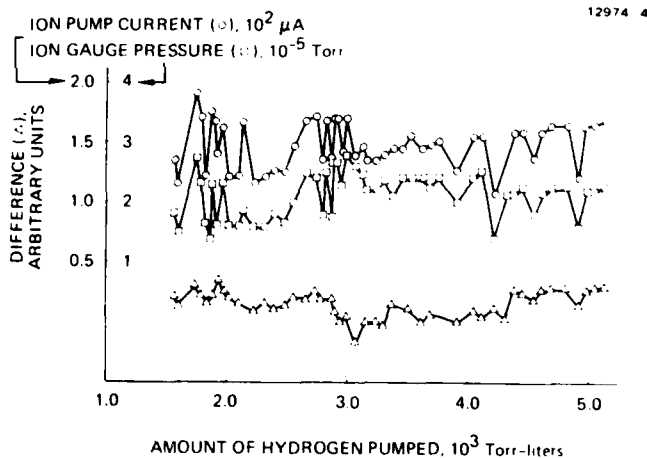


Figure 7. Ion pump current and ion gauge readings during a getter capacity test. The good tracking of the two sets of readings reflects stable ion pump operation in a clean system in the presence of large hydrogen throughput.

Combination Getter-Ion Pump System Operational Experience

Combination zirconium graphite getter and ion pump systems were used on our prototype oscillating compact masers, CHYMNS-I and II,⁵ with varying degree of success. The getter pumps were activated in situ after high vacuum had been established in the maser using an external pump. Immediately after getter activation, the background pressure was at its lowest point and rose gradually in time, as indicated by an increase in

the ion pump current. We first attempted to explain this behavior as being due to degradation in the pumping speed of the getter pump. Since CHYMNS-I and II employed viton O-rings and had not been baked to high temperatures, the degradation could presumably be accelerated. However, this does not explain the observed instability in the ion pump, which occurred in 3- to 6-month intervals, and the technique we employed to rejuvenate the system. For example, when ion pump instability occurs, the system can be brought back to normal stable operation without the use of an external pump. The procedure consists of operating the ion pump power supply in the start mode and letting a controlled amount of hydrogen into the maser to cause a thermal run-away in the ion pump. The hydrogen supply is then shut off while the ion pump undergoes a self bake out. The high temperature will force most of the hydrogen absorbed in the electrodes of the ion pump to be released. Thus, the system pressure would rise. In a few hours, the background pressure would begin to fall, and in due course, the ion pump would begin to function, leading to a stable operating high vacuum again. This would be good for another 3- to 6-month cycle. Note that without another pump to evacuate the gas released from the hot ion pump, the ion pump could not start to function. Thus, the getter pump must still be functioning to absorb the hydrogen and reduce the background pressure to enable the ion pump to function again. This technique had been successfully repeated.

As noted earlier, we have not observed any ion pump instabilities in our clean getter pump development test station, although the hydrogen throughput is at least an order of magnitude higher than in the maser. The main difference is the absence of elastomers in the getter test station and the fact that it has been baked out at high temperatures under high vacuum conditions. This suggests that elastomers may be deleterious not only to getter pump operation but also to ion pumps as well. This is due to the fact that hydrogen pumping in ion pumps is predominantly by diffusion into the cathode rather than by chemical combination at the anode, such as occurred for heavier molecular species. (The light hydrogen ions produce comparatively little sputtering.) Therefore, surface conditions and permeability of the cathode are critical factors in hydrogen pumping. The former is, of course, a sensitive function of gas composition in the system. This discussion points out the importance of eliminating, or at least minimizing the presence of, elastomers from the maser vacuum system. Cleanliness and high temperature bake-out capability are critical requirements of the maser vacuum system design.

Conclusion

Reversible interaction of hydrogen with metals and alloys can be exploited to effectively solve the gas handling problems in hydrogen masers. A prototype of a maser hydrogen source employing hydride storage techniques has been successfully tested. With regard to application of getter pumps in the maser, several design problems such as powder confinement and getter poisoning remain to be solved. If size and weight are not critical considerations, zirconium aluminum alloy (ST-101)⁴ may be a more suitable getter. The large surface area resulting from spraying the powdered getter on a large mechanical support makes the getter more resistant to poisoning. Knowing the design criteria, the problems are not insurmountable. We believe that getter pumps will contribute to greatly improved reliability, along with smaller size, lower weight and power consumption in the maser.

Acknowledgments

The contributions of Jack Lewis, Ted Calderone, and Lenny McNulty in these investigations are gratefully acknowledged.

References

1. G. Alefeld and J. Volkl, Editors, Hydrogen in Metals II (Springer-Verlag, New York, 1978).
2. R.F.E. Vessol et al., Proc. 8th Annual FTI: (1976); H.E. Peters, Proc. 34th Annual Freq. Control Symp. (1980), p. 360.
3. "HY-STOR" is a trade name of Ergonics, Wyckoff, N.J.
4. "ST-171" and "ST-101" are trade names of SAES Getters, Colorado Springs, CO.
5. H.T.M. Wang, "Characteristics of Oscillating Compact Masers," Proc. 36th Ann. Symp. on Freq. Control (1982), pp. 249-254.

COMPACT RECTANGULAR CAVITY FOR RUBIDIUM VAPOR CELL FREQUENCY STANDARDS

H.F. Williams, T.M. Kwon and T. McClelland

Litton Guidance and Control Systems
5500 Canoga Avenue
Woodland Hills, California 91365

Abstract

A new 6.8 GHz microwave cavity design for use in rubidium vapor cell frequency standards is discussed. The design consists of a rectangular cavity operating in the TE₁₀₁ mode, partially loaded with a low loss dielectric slab. This configuration combines the desirable H-field uniformity of conventional cylindrical TE₀₁₁ cavities with the reduced size advantage of cylindrical TE₁₁₁ cavities. Careful selection of dimensions and dielectric material actually result in a resonator substantially smaller than even the cylindrical TE₁₁₁ cavity, yet without significant degradation of field uniformity. Test results for a cavity based on this design in an operational Rb frequency standard are presented.

1. INTRODUCTION

Traditionally, the microwave cavity for passive rubidium frequency standards has been a cylindrical resonator operating in the TE₀₁₁ mode. Such a cavity produces a fairly uniform magnetic field near the center along the cylindrical axis, however, in order to resonate at the 6.8 GHz rubidium hyperfine frequency, it must have a diameter larger than 5 cm. (See Figure 1.) This represents a severe constraint on the minimum size of any frequency standard utilizing this cavity design, and in order to eliminate this constraint other cavity configurations have been introduced.

A cylindrical cavity operating in the TE₁₁₁ mode offers one alternative. For a given frequency and length to diameter ratio the cylindrical TE₁₁₁ cavity has a diameter about 1/3 that of the TE₀₁₁ cavity. Unfortunately the magnetic field configuration is less desirable: being maximum at the cavity walls and falling to zero at the center of the cavity. (See Figure 2.)

Additional size reduction for a given cavity is possible by partially loading the cavity with low loss dielectric material with a high dielectric constant. Although dielectric material in the cavity reduces the Q-factor because of losses, the field lines are compressed in the dielectric material and thus substantial size reduction is possible. Since high Q's are not necessary in passive devices this tradeoff is usually acceptable. Dielectric-loaded cylindrical cavities of both the TE₀₁₁ and TE₁₁₁ configuration have been successfully used in miniaturized frequency standards^{1,2}.

In this paper we report the first use of a rectangular cavity partially loaded with dielectric material in a passive rubidium frequency standard. It will be shown that such a design makes possible a cavity as small or smaller than dielectrically loaded cylindrical cavities, while maintaining high magnetic field uniformity. A rectangular cavity of this type is the logical choice for miniaturized frequency standards which are packaged in containers of rectangular cross-section.

II. FIELD EQUATIONS FOR DIELECTRICALLY LOADED TE₁₀₁ CAVITY

We consider a rectangular cavity operating in the TE₁₀₁ mode. The field configurations in such a cavity are (see Figure 3):

$$H_x = -H_0 \frac{k_z}{k_x} \sin k_x X \cos k_z Z \quad (1a)$$

$$H_y = 0 \quad (1b)$$

$$H_z = H_0 \cos k_x X \sin k_z Z \quad (1c)$$

$$E_x = 0 \quad (1d)$$

$$E_y = -j H_0 \left(\frac{\mu}{\epsilon} \right)^{1/2} \left(\frac{k}{k_x} \right) \sin k_x X \sin k_z Z \quad (1e)$$

$$E_z = 0 \quad (1f)$$

$$\text{where } k_x^2 = \omega^2 \mu \epsilon - \left(\frac{2\pi}{\lambda_g} \right)^2 \quad (2a)$$

$$k_z = \frac{\pi}{d} = \frac{2\pi}{\lambda_g} \quad (2b)$$

$$k^2 = k_x^2 + k_z^2 = \omega^2 \mu \epsilon \quad (2c)$$

If a dielectric slab is inserted such that its face is in the y-z plane, as in Figure 4, then two sets of equations replace (1), with the constraint that the two sets of equations satisfy the appropriate boundary conditions at the dielectric boundary. These boundary conditions can be expressed as

$$H_z^{(1)} \cos(k_x^{(1)} S) = -A \cos k_x^{(2)} (S-a) \quad (3a)$$

$$E_y^{(1)} j \frac{1}{k_x^{(1)}} \sin k_x^{(1)} S = -jA \frac{1}{k_x^{(2)}} \sin k_x^{(2)} (S-a) \quad (3b)$$

Dividing Equation (3b) by Equation (3a), an expression is obtained which allows a and S to be determined for a given d, ω , ϵ_1 , ϵ_2 .

$$k_x^{(2)} \tan k_x^{(1)} S = k_x^{(1)} \tan k_x^{(2)} (S-a) \quad (4)$$

The configuration of interest results for the special case

$$k_x^{(1)} = 0, \text{ or } \omega^2 \mu \epsilon = \frac{2\pi}{\lambda_g} \quad (5)$$

Then Equation (4) becomes,

$$\tan k_x^{(2)} (S-a) = k_x^{(2)} S \quad (6)$$

and the fields become,

$$H_z^{(1)} = H_0 \sin(k_z Z) \quad H_z^{(2)} = -A [H_0 \sin(k_z Z)] \cos k_x^{(2)} (x-a) \quad (7a)$$

$$H_x^{(1)} = -[H_0 \cos(k_z Z)] k_z X \quad H_x^{(2)} = +A [H_0 \cos(k_z Z)] \left(\frac{k_z}{k_x^{(2)}} \right) \sin k_x^{(2)} (x-a) \quad (7b)$$

$$E_y^{(1)} = -j \omega \mu_1 [H_0 \sin k_z Z] X \quad E_y^{(2)} = jA \frac{\omega \mu_2}{k_x^{(2)}} [H_0 \sin k_z Z] \sin k_x^{(2)} (x-a) \quad (7c)$$

where

$$A = \sqrt{1 + [k_x^{(2)} S]^2} \quad (8)$$

As can be seen from Equation (7a) the longitudinal magnetic field in the air space, $H_z^{(1)}$, is a function only of the z-coordinate, so a

relatively homogeneous field is obtained in the region where the Rb cell is located. The fields are plotted in Figures 5, 6 and 7.

The current flow in the walls of a cavity is always perpendicular to the direction of the tangential H-field at the walls. In this cavity, for $z = 0, d$, $\vec{H} = H_x \hat{x}$, and thus current flows only along the y-axis in the end faces. Furthermore, at $X = X_0 = a + \frac{\pi}{2k\sqrt{2}}$, for any value of Z , $\vec{H} = H_x \hat{x}$, and thus current flows along the line $X = X_0$ in the top and bottom faces. A cavity of this type can thus be constructed from two pieces joined together in the y-z plane at $X = X_0$, without significant degradation of cavity Q due to poor electrical contact.

The overall dimensions of this cavity are not completely determined by the above choice of parameters. The "b" dimension does not enter the expressions, so it remains a free parameter. The "a" dimension depends on the width of the air space, using Equation (6). "a" can be calculated as a function of S for any desired value of the dielectric constant, ϵ_2 . The d dimension is fixed for a rectangular cavity, being just $\lambda_g/2$, in this case this reduces to 1/2 the free space wavelength, from Equation (5).

III. EXPERIMENTAL CAVITY PERFORMANCE

We have designed a cavity based on the principles of Section II, using a machinable glass-ceramic, as the dielectric material ($\epsilon = 6$, $\tan \delta = 0.007$). The cavity dimensions are $a = 0.7$ in., $b = 0.5$ in., $d = 0.8$ in., and the width of the dielectric slab is 0.2 in.

The cavity is constructed from two pieces of metal joined in a plane parallel to the surface current flow. Cavity tuning is accomplished with a tuning screw inserted along the inside edge of the dielectric. The cavity resonance is excited by the 57th harmonic of a

120 MHz input signal, generated in a step recovery diode, and coupled to the cavity via a loop. In order to monitor the Rb resonance, holes 0.320" in diameter are cut in the ends of the cavity.

The theoretical value of Q for this cavity (ignoring the end holes, and losses in the Rb vapor cell) is 160. The measured Q, with a resonance cell inside, is ≈ 50 . This value is much smaller than the theoretical value because of losses through the end holes and in the Rb resonance cell.

This cavity design has been incorporated into a working passive Rb frequency standard. A typical Rb resonance dispersion curve for such a device is shown in Figure 8. From this curve the Allan variance for a 1 second averaging time is estimated to be $\sigma_y(\tau = 1 \text{ sec}) = 4.7 \times 10^{-12}$.

References

1. D.A. Howe, F.L. Walls, H.E. Bell, and H. Hellwig, "A Small, Passively Operated Hydrogen Maser," Proc. 33rd Ann. Sym. on Freq. Control, 554 (1979).
2. W.J. Riley, "A Rubidium Clock for Seek Talk," Proc. 14th Ann. PTTI Applications and Planning Meeting, 141 (1982).
3. Montgomery, Dicke, Purcell, MIT Radiation Laboratory Series, Vol. 8, McGraw Hill, New York (1948).

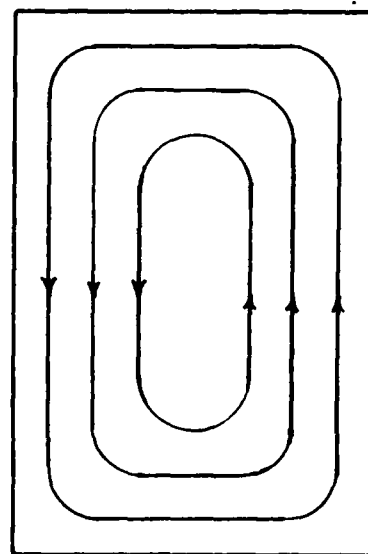
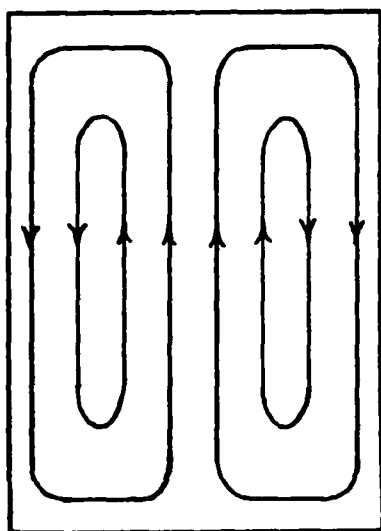
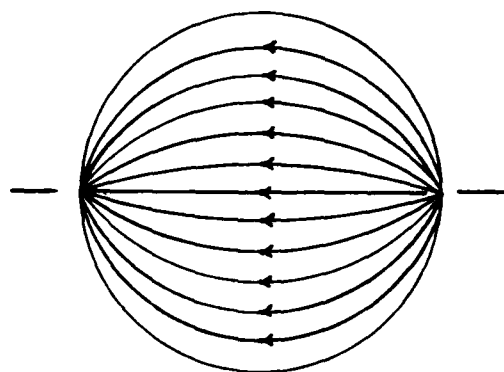
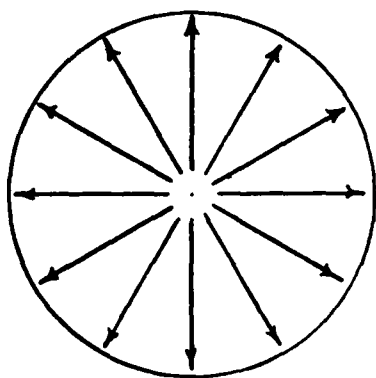


Figure 1. Cylindrical TE_{011} Cavity Showing H-Field Distribution

Figure 2. Cylindrical TE_{111} Cavity Showing H-Field Distribution

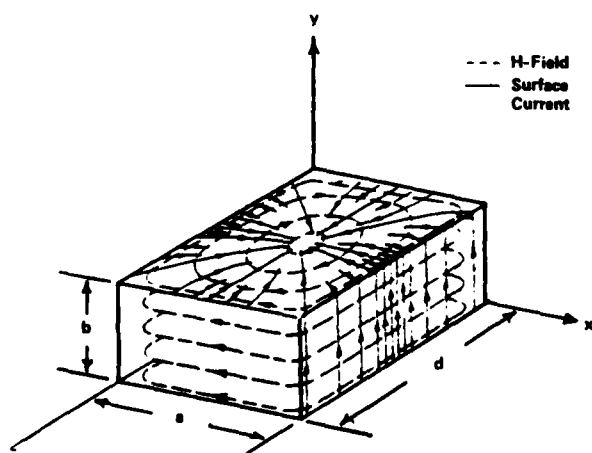


Figure 3. TE_{101} Cavity

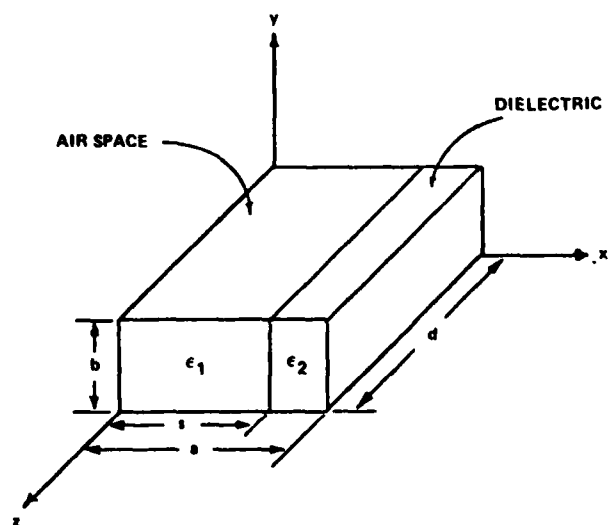


Figure 4. TE_{101} Cavity with Dielectric Slab

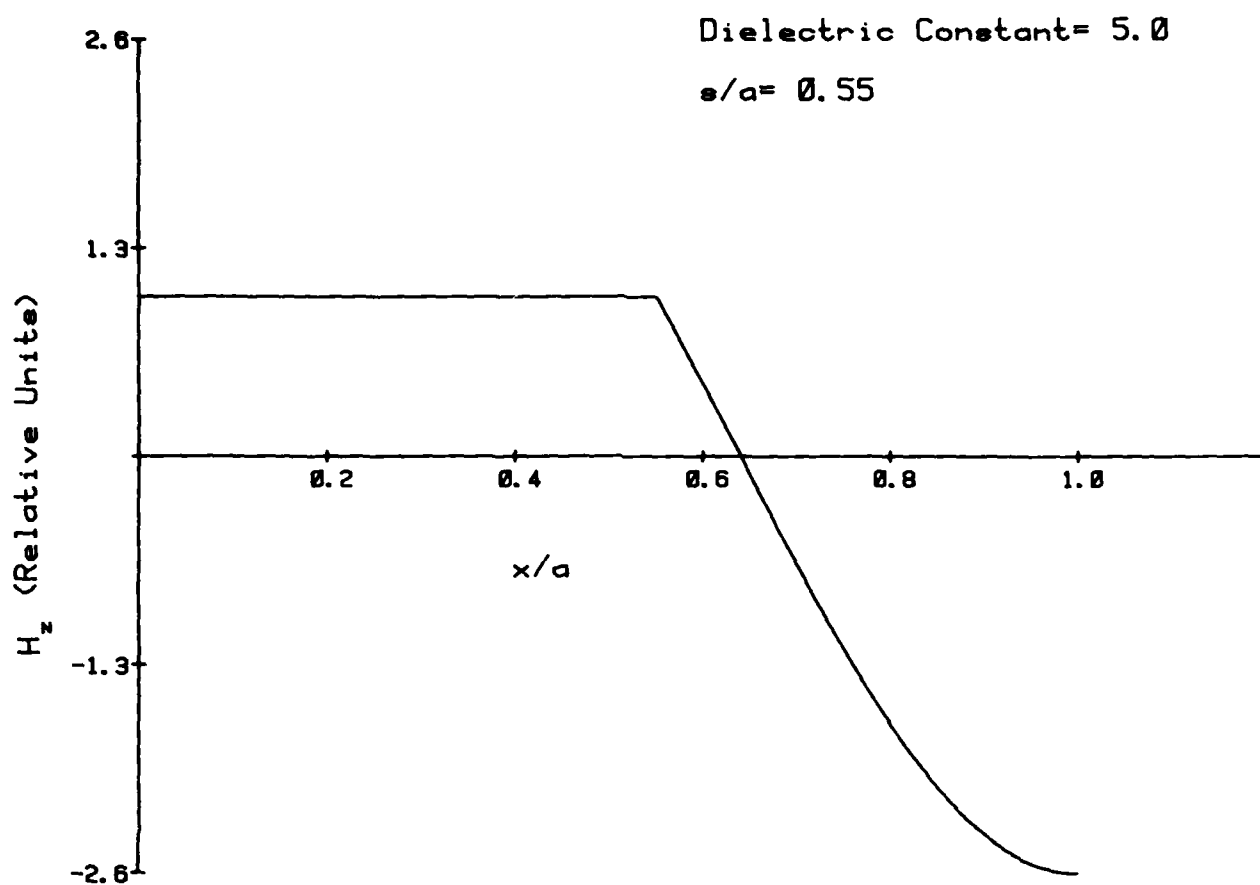


Figure 5.

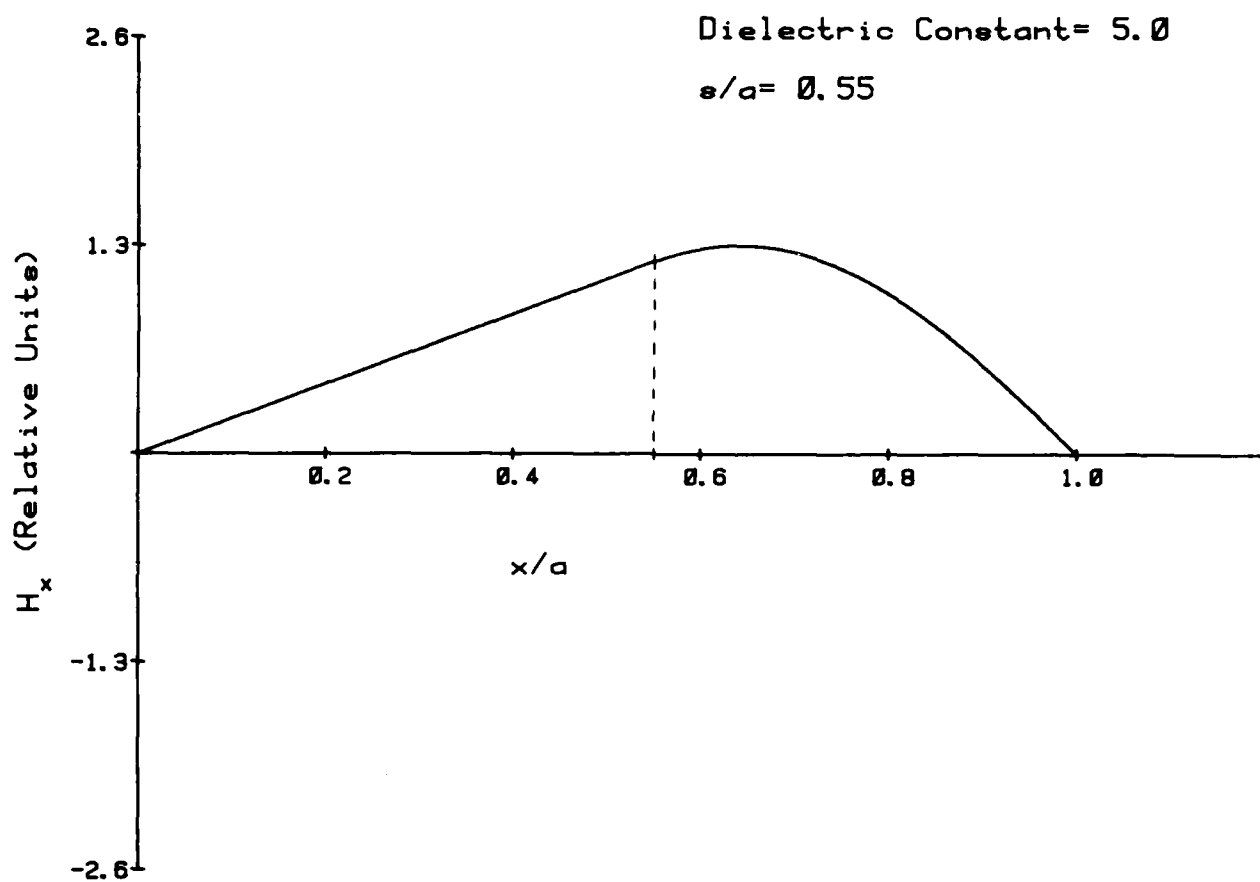


Figure 6.

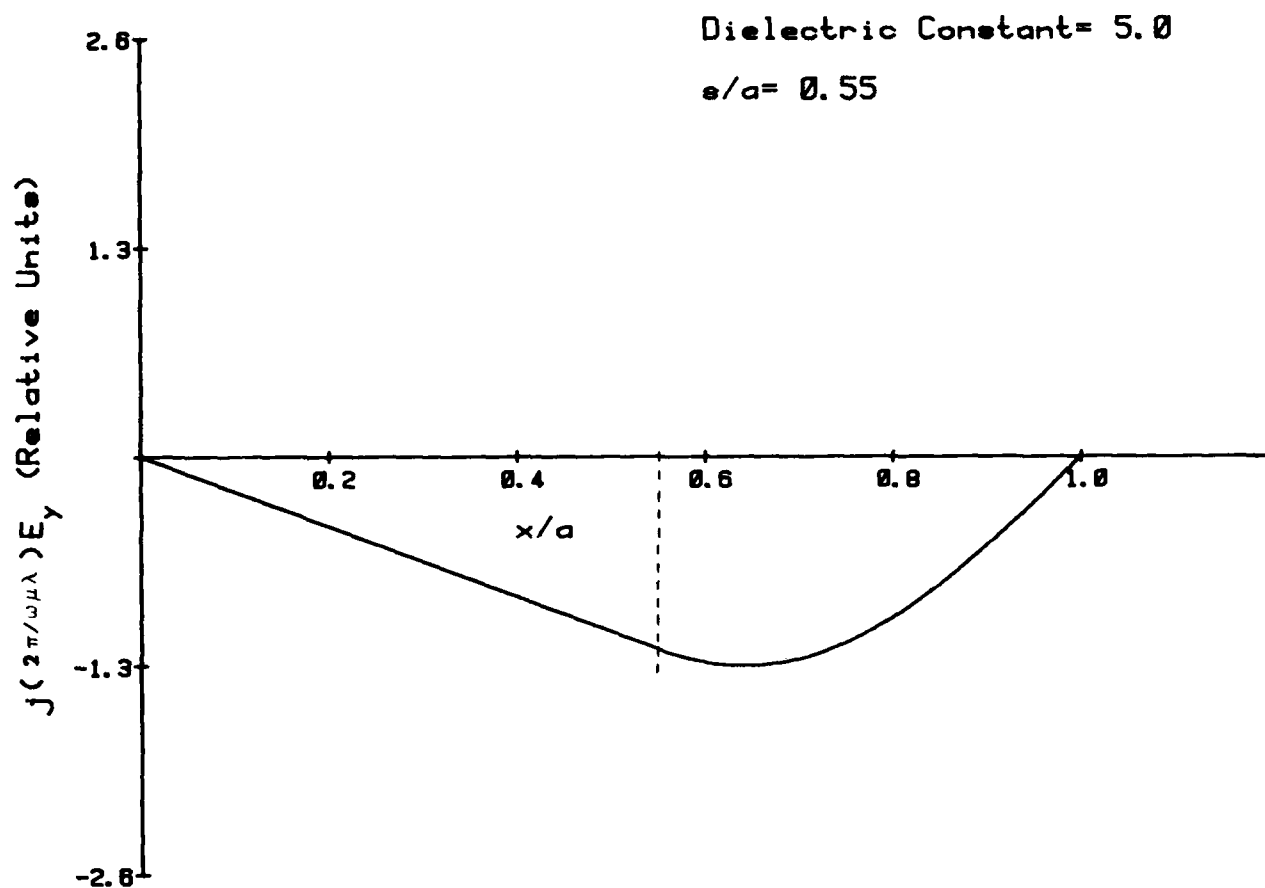
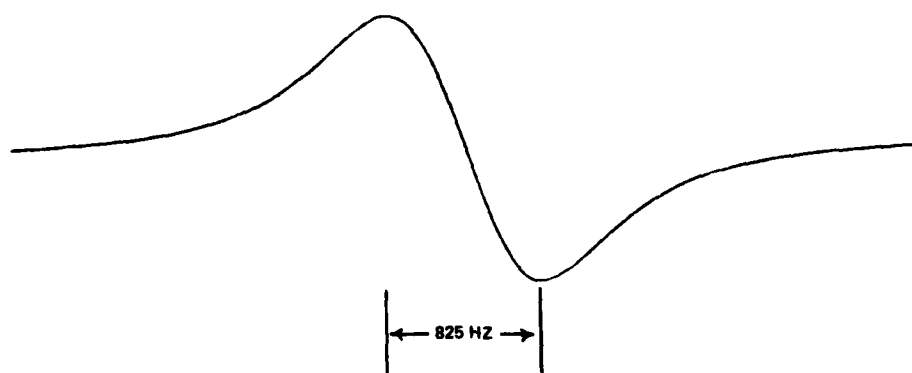


Figure 7.



$$\sigma_y (\tau = 1 \text{ SEC}) = 4.7 \times 10^{-12}$$

Figure 8. Rb Dispersion Signal using Dielectrically Loaded TE_{101} Cavity

IMPROVED VIBRATION PERFORMANCE IN PASSIVE ATOMIC FREQUENCY STANDARDS BY SERVO-LOOP CONTROL

T.M. Kwon and T. Hahn

Litton Guidance and Control Systems
Woodland Hills, California 91364

Abstract

Due to its small size, rapid warm-up, and low power consumption, the rubidium vapor-cell frequency standard unit appears to be the best candidate for most current and near-term tactical frequency standard applications. Its performance under vibration environments, however, remains to be improved. Sources of vibration-induced frequency instabilities in a rubidium frequency standard are reviewed. Demonstrated in this paper is the improved vibration performance of a rubidium frequency standard with the VCXO frequency being tightly servo-controlled to the atomic resonance reference. Technical challenges associated with such a mechanization are briefly discussed. The measured phase spectral density of a rubidium frequency standard with servo-loop bandwidth of ~ 100 Hz is shown.

Introduction

The need for an airborne tactical frequency standard is growing rapidly in both quantity and quality. Quartz crystal oscillators have been used almost exclusively for these applications. Today, several tactical military applications require frequency performance beyond the level that the state-of-the-art crystal oscillators can offer. In addition, the frequency standards for tactical applications demand small size, rapid warm-up, and low power consumption. Among the available technologies, the rubidium vapor-cell frequency standard appears to be the best fit for most current and near-term tactical frequency standard applications.

The output of a rubidium frequency standard is that of a voltage controlled crystal oscillator (VCXO), the frequency of which is servo-controlled to the clock transition frequency of rubidium atoms. Bandwidth of the servo-loop in a typical rubidium frequency standard is a few Hz. The vibration-induced sidebands in the VCXO, therefore, appear directly in the output of a rubidium frequency standard, when the unit is subjected to vibrational environments.

The vibration inputs under the tactical environments, may be a few Hz to a few thousand Hz with accelerations as high as $\pm 10g$ peak, a situation in which no rubidium frequency standard available to date can maintain its stable frequency output. Overall frequency performance of a rubidium frequency standard under vibration can be considered to be worse than that of a free running VCXO, if one takes into account the erroneous output frequency resulting from the vibration-induced perturbation of the control servo-loop. In order to utilize the rubidium frequency standard technology advantageously for the tactical applications, it is imperative to improve its vibration performance beyond what is available to date.

Vibration performance of a rubidium frequency standard is improved by a servo-loop bandwidth that is substantially higher than that of a conventional rubidium frequency standard. The idea behind this is to lock the inherently vibration sensitive VCXO output frequency tightly to the inherently vibration-insensitive clock transition frequency of rubidium atoms. This idea is not new having been the subject of discussions within the time and frequency community, but has not been realized, to our knowledge, in a rubidium frequency standard.

Sources of Vibration Sensitivities

The output frequency instabilities induced by vibration in a rubidium frequency standard have been the subject of a few published papers.¹ Observations of such instabilities have been

made typically in the time domain (rather than in the frequency domain). The vibration frequencies at which significant frequency instabilities appear to occur may be categorized as follows:

1. $f_v \approx f_m$, where f_v and f_m are the vibration frequency and the frequency of modulation of the phase-sensitive detector in the servo-loop, respectively.
2. $f_v \approx 2 f_m$.
3. Low vibration frequencies.

The vibration simulation tests indicate that, in the region of $f_v \approx f_m$, the frequency fluctuations are caused by the spurious modulations in the light intensity present at the photodetector. The interference signal is generated by the mechanical translations of physics components with respect to each other, and is seen as a perturbation to the servo-loop. When $f_v \approx 2 f_m$, the significant effect comes from the vibration-induced sidebands of the VCXO. The sideband intermodulates the modulated signal of rubidium resonance resulting in a spurious frequency component at f_m . This component is seen as a perturbation by the servo-loop in an analogous manner to the case of $f_v \approx f_m$. The subject of servo-loop perturbation is beyond the scope of this paper, and will be published in a separate paper.

The frequency instabilities at low vibration frequency arise from vibration-induced sidebands of the VCXO. Although such sidebands exist at all vibration frequencies, observation of the frequency instabilities is limited to the bandwidth of the time domain measurement system.

At the lower end of the low vibration frequency spectrum, R. Filler has shown that there exist discrete vibration frequencies at which the power at the resonance interrogation frequency 6.8 GHz, becomes identically zero.² Results are elaborated in Figures 1 and 2. The power of 6.8 GHz, which is generated by frequency multiplication of 10 MHz VCXO output, is plotted as a function of vibration frequency at constant acceleration of $\pm 10g$ peak. Acceleration sensitivity of a VCXO, $\gamma = 2 \times 10^{-9}/g$, is assumed in Figure 1. Complete loss of the 6.8 GHz power is apparent up to $f_v \approx 70$ Hz. At such discrete vibration frequencies, the rubidium frequency standard no longer functions as a frequency standard. Figure 2 is the case for $\gamma = 2 \times 10^{-10}/g$ corresponding to the state-of-the-art VCXO, where no loss of the 6.8 GHz power occurs beyond $f_v \approx 7$ Hz. The choice of the appropriate acceleration sensitivity of VCXO appears to be the key to the tactical rubidium frequency standards.

At the higher end of the low vibration frequency spectrum, say, $f_v \gtrsim 7$ Hz in the example of $\gamma = 2 \times 10^{-10}/g$, the vibration-induced sidebands appear directly in the output if they are not controlled by the servo-loop.

The vibration-induced frequency instabilities in this region are to be improved by the increased servo-loop bandwidth as discussed below.

Frequency Stability Under Vibration

Transfer of frequency stability from an atomic frequency reference to the VCXO has been discussed by J. Vanier, et. al.³ In a passive atomic frequency standard, the phase spectral density, $S_{\phi}(f)$, of its output frequency is shown to be

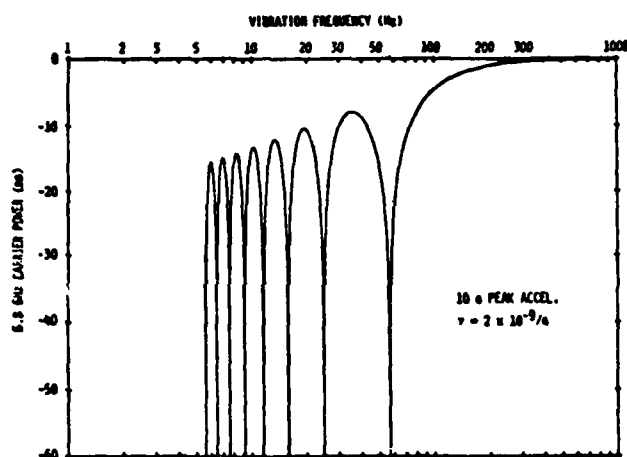


Figure 1. Microwave Power vs. Vibration Frequency for $\gamma = 2 \times 10^{-9}/g$

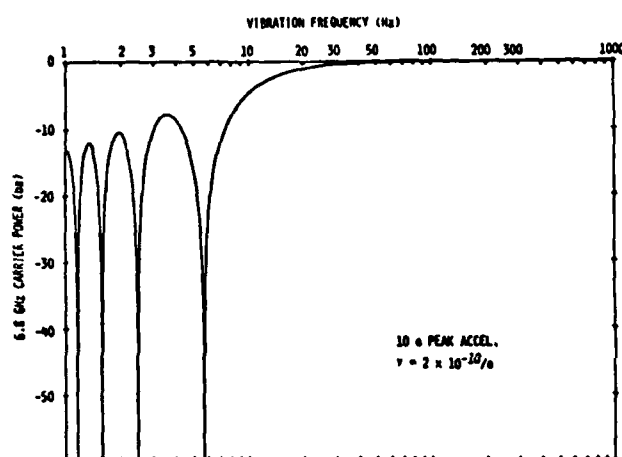


Figure 2. Microwave Power vs. Vibration Frequency for $\gamma = 2 \times 10^{-10}/g$

$$S_{\phi}(f) = \left[\frac{(f/f_n)^2}{1 + (f/f_n)^2} \right] S_{\phi q}(f) + \left[\frac{1}{1 + (f/f_n)^2} \right] S_{\phi r}(f) \quad (1)$$

where $S_{\phi q}(f)$ and $S_{\phi r}(f)$ are the phase spectral density of the free running VCXO and of the atomic reference, respectively, as a function of frequency, f . f_n is bandwidth of the servo-loop, in which an integrator is used as a filter.

We consider the case where only the phase spectral density of the VCXO is affected by vibration:

$$S_{\phi q}(f) = S_{\phi q}^o(f) + S_{\phi q}^v(f) \delta(f - f_v) \quad (2)$$

where f_v is the vibration frequency, and the superscripts, o and v, denote the VCXO phase spectral density under static and vibration conditions, respectively.

In a typical VCXO, the vibration-induced sidebands are described by frequency modulation with small modulation index, in which case most of the power is in the carrier and only a small amount is in the first upper and lower sideband. The ratio of the amount of power carried by the first sidebands to the power remaining in the carrier, ξ , is shown to be²:

$$\xi(f_v) \approx \left(\frac{\gamma A_p f_o}{2 f_v} \right)^2 \quad (3)$$

where, f_o is the carrier frequency of VCXO, and A_p is the peak acceleration in units of g. The vibration-induced sideband represents deterministic noise. It shows up, however, as a discrete frequency component in the plot of $S_{\phi}(f)$ versus f at $f = f_v$. The magnitude of phase spectral density of such a frequency component is equivalent to that given in Equation (3) from the fact that the peak phase deviation of the VCXO frequency is much less than one radian⁴, i.e.,

$$S_{\phi q}^v(f) \approx \xi(f_v) \quad (4)$$

Substituting Equations (2) and (4) into Equation (1), we arrive at an expression for phase spectral density of a servo controlled VCXO in a passive frequency standard subjected to sinusoidal vibrations

Allan variance is calculated through the relation³:

$$\sigma^2(\tau) = 2 \int_0^{\infty} \left(\frac{f^2}{f_o^2} \right) S_{\phi}(f) \frac{\sin^4(\pi f \tau)}{(\pi f \tau)^2} \left[\frac{1}{(1 + f/f_c)^2} \right] df \quad (5)$$

where, τ is the averaging time and f_c is the cutoff frequency of the one-pole filter assumed to be present at the input of the Allan variance measurement system.

For graphic demonstration of the effect of servo-loop bandwidth on the frequency stabilities, the following values are assumed:

$$S_{\phi r}(f) = 2.0 \times 10^{-18} \frac{1}{f^2} + 1.8 \times 10^{-11} \frac{1}{f^3} \quad (6a)$$

$$S_{\phi q}^o(f) = 1.0 \times 10^{-9} \frac{1}{f^3} + 1 \times 10^{-13} \quad (6b)$$

$$\gamma = 3 \times 10^{-10}/g \quad (6c)$$

$$f_o = 10 \text{ MHz} \quad (6d)$$

Shown in Figure 3 is the phase spectral density calculated from Equation (1). The two lower curves in Figure 3 are the phase spectral density under static conditions for bandwidth of 1 Hz and of 100 Hz. For the values assumed in Equations (6a) and (6b), the increased servo-loop has resulted in degraded short-term noise at a static environment. The upper two curves in Figure 3 represent the point-to-point connections of the varying peak values of vibration-induced sideband for bandwidth of 1 Hz and of 100 Hz as the vibration frequency is varied. Constant acceleration of $\pm 5g$ peak is assumed throughout the vibration frequency spectrum. Improved servo-loop bandwidth is explicit. Figure 4 is the data taken from Litton engineering model tactical rubidium frequency standard (TRFS). The vibrating VCXO is simulated in the measurement. Good agreement between Figures 3 and 4 justifies the analytical predictions of Equation (1).

Allan variances calculated numerically from Equation (5) are shown in Figures 5 and 6 for a servo-loop bandwidth of 1 Hz and of 100 Hz, respectively. Sinusoidal vibration of 5 Hz, $\pm 5g$ acceleration is assumed for both.

Implementation of Increased Servo-Loop Bandwidth

Implementation of increased servo-loop bandwidth has its own set of technical challenges. In the conventional passive frequency standards, the excellent short-term stability of VCXO is directly transferred to the output of the frequency standard. With increased servo-loop bandwidth, the short-term stability of its output frequency becomes that of the atomic reference. As was demonstrated in Figure 3, there exists a potential of degraded short-term stability

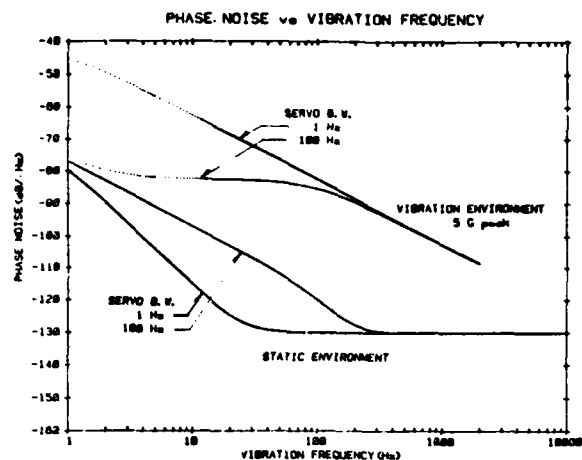


Figure 3. Phase Spectral Density vs. Frequency as a Function of Servo-Loop Bandwidth

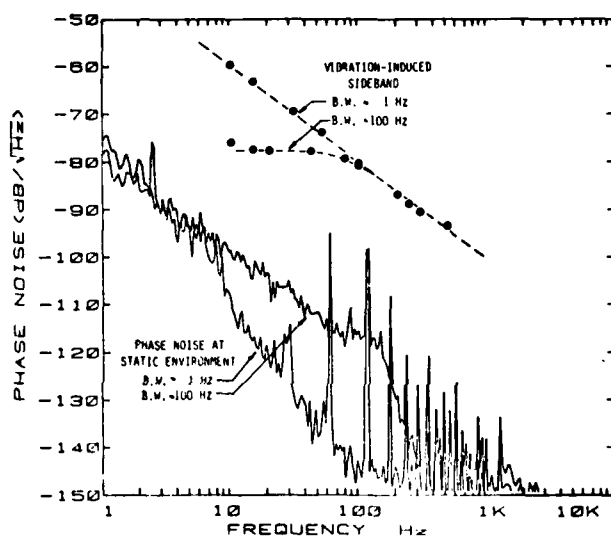


Figure 4. Measured Phase Spectral Density vs. Frequency as a Function of Servo-Loop Bandwidth

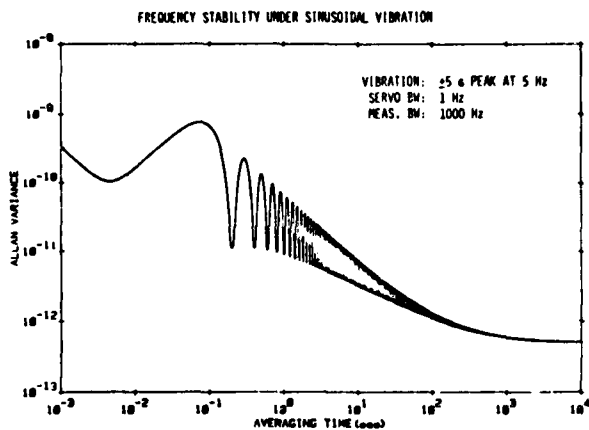


Figure 5. Allan Variance for Servo-Loop Bandwidth of 1 Hz

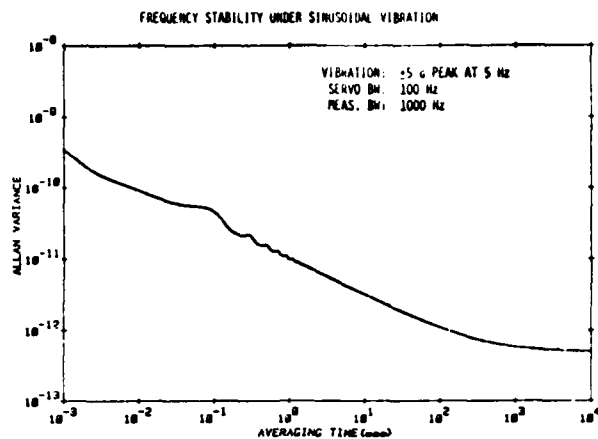


Figure 6. Allan Variance for Servo-Loop Bandwidth of 100 Hz

with an increased servo-loop bandwidth, if the atomic reference stability is not better or comparable to that of the VCXO.

In order to mechanize the increased servo-loop bandwidth, the modulation frequency must be higher than the desired bandwidth⁵. It has been determined empirically in this laboratory and by others analytically^{5,6} that the increased modulation frequency may be employed with no serious degradation in stabilities of the atomic resonance reference. Both sinusoidal and squarewave modulation may be implemented with modulation frequency lower than or higher than the linewidth of the atomic resonance.

For a given slope of the atomic resonance frequency discriminator, the servo-loop bandwidth can be further increased by increasing electronic gain of the servo-loop. The increased gain makes the control loop susceptible to spurious noise.

Summary and Acknowledgement

We have shown that frequency performance of a rubidium frequency standard under tactical vibration environments is improved by increasing the bandwidth of the VCXO servo-control loop. Discussed briefly are the technical challenges associated with the increased bandwidth. Bandwidth ~ 100 MHz is implemented, for the first time to our knowledge, in a rubidium frequency standard unit.

Authors are indebted to H. Williams and W. Debley for their helpful discussions. This work is supported in part by the U.S. Air Force Rome Air Development Center Electronic Systems Division.

References

1. Norman Houlding, Proc. 13th Ann. PTTI Meeting (1981).
2. R. Filler, Proc. 35th Ann. Freq. Control Symposium (Philadelphia, 1981).
3. J. Vanier and M. Têtu, IEEE Trans. Instr. Meas. **28** 188 (1979).
4. J. Gibbs, "Computer-Aided Phase Noise Measurement", HP Publication.
5. B. Ginzburg and V. Mogilyoukin, Opt. Spectrosc. (USSR), **50**(1) 44 (1981).
6. C. Audoin, J. Viennet, N. Cyr, and J. Vanier, Proc. 14th Ann. PTTI Meeting (1982).

AD P002447

ULTRA-STABLE LASER CLOCK, SECOND GENERATION

Lieutenant Roger Facklam
AFWAL/AAAN-1
Wright-Patterson AFB, Ohio

Summary

The Air Force has requirements for both high accuracy timing sources and precision inertial navigation systems for advanced fighter aircraft. The author has devised a method of using a modified multi-frequency ring laser gyro (RLG) simultaneously as a gyro and a clock. This device uses a multi-frequency RLG with auxiliary detectors to sense two beat frequencies (each at about 582 MHz) and necessary electronics to combine them and produce a 5MHz clock signal. The advantage of the laser clock will be size, costs, and redundancy. The aircraft that already has an RLG navigator will then only require the addition of detection and output electronics to implement the clock function. This additional equipment will add very little to the acquisition and maintenance costs for the RLG navigator. A three axis navigator could have three clocks, thereby providing redundancy.

This paper is a continuation of the work presented last year at the Frequency Control Symposium. The ring laser used is a Raytheon RB-25 which is a non-planar He-Ne laser whose optical path length is 25.69 cm. The best data from last year was taken on laser #68 which was an operational prototype RLG. Two main error sources were evaluated; (1) the quantum limit imposed by the Heisenburg uncertainty principle, and (2) the temperature induced drift in the frequency shift caused by the Faraday rotator. The current work does not attempt to reduce the quantum limit, which could be done by increasing the laser power and/or the cavity Q. Improvements in the quantum limit would require building a new laser and funding was not available for a new laser. The current effort addresses improvement in the Faraday rotator frequency drift as a function of temperature. The new electronics has one detector to measure the clockwise beat frequency and a second to measure the counterclockwise beat frequency. The two beat frequencies are summed and then divided down in frequency to near 5MHz.

The RLG clock is a transfer frequency standard which must be set to a reference before use. Setting could be accomplished by using the Global Positioning System, the Joint Tactical Information Distribution System, radar techniques, or physically connecting a time standard to the RLG clock.

Key Words: Helium-Neon, Laser, Ring Laser, Clock, Frequency standard, Transfer Frequency standard.

I. Background

The objective of this paper is to demonstrate the feasibility of using a four-frequency ring laser to simultaneously provide gyro and clock outputs. The test of the first generation clock took place in March and September, 1982. The second generation tests were conducted in February, 1983. Some application of precision timing sources are bi-static radar, secure/anti-jam communication, and aircraft identification avionics. Precision navigation is needed to support fighter aircraft missions.

II. Theory

The two desirable characteristics of a frequency standard are the narrowness of the signal in frequency space and the stability of the center frequency as a function of time. The ultimate narrowness of the laser clock frequency is determined by the quantum limited laser linewidth. The first generation laser clock had a long term frequency drift caused by the temperature induced drifting of the Faraday rotator frequency shift. The second generation laser clock had a long term drift frequency caused by dispersion effects. The theory section will cover clock characterization, laser cavity physics, gyro operation, first generation clock/gyro operation, second generation clock/gyro operation, and optical dispersion.

Clock Characterization

Clock accuracy is characterized by the standard deviation of the frequency fluctuations and may be written as

$$\sigma_y(\tau) = \frac{1}{\nu_c} \sqrt{\frac{m}{\tau} \sum_{k=1}^m \frac{(\Delta f_{k+1} - \Delta f_k)^2}{2m}} \quad (1a)$$

where

$$\Delta f_k = \frac{\Delta \phi_k}{2\pi\tau} \quad (1b)$$

and ν_c is the nominal frequency, m is the number of measurements, τ is the integration time, $\Delta \phi_k$ is the integrated phase error, Δf_k is the k^{th} measurement of the frequency, and $\sigma_y(\tau)$ is the Allan variance.¹

Laser Cavity Physics

The ring laser clock/gyro produces four laser frequencies. The cavity, shown in Figure 1, is non-planar and allows only circularly polarized light to exist.

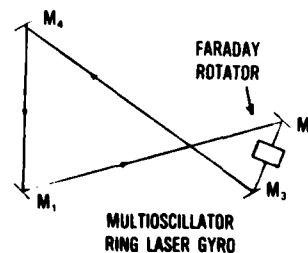


Figure 1 Ring Laser Gyro/Clock

The cavity also produces a frequency difference between the right and left hand circularly polarized (RHCP & LHCP) modes.

The average beat frequency between RHCP and LHCP is given by

$$\nu_{\text{beat}} = \frac{c}{2nl} \quad (2)$$

where c is the speed of light in a vacuum, n is the index of refraction, l is the physical pathlength and nl is the average optical pathlength.²

The configuration shown in Figure 1 contains a Faraday rotator, which is a non-reciprocal optical device that uses the Faraday effect to shift the frequency of the CW mode from the frequency of the CCW mode.

The Faraday induced frequency shift is given by

$$\nu_F = \frac{c}{n_l} \frac{\Delta\rho}{2\pi} \quad (3)$$

where $\Delta\rho$ is the phase shift between CW and CCW modes.³ This is shown as the separation between 1 and 2, or 3 and 4 in Figure 2.

Gyro Operation

The ring laser operates as a gyro due to the relativity effect that causes the frequency of the CW and CCW modes in a ring laser to differ depending on rotation rate. The output is obtained by optically beating the LHCP modes and the RHCP modes, and then differencing the two beat frequencies by means of a mixer to produce the gyro output.

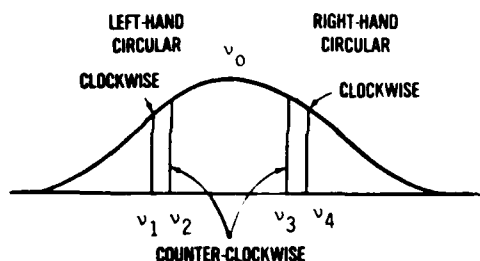


Figure 2 Four modes on gain curve with Faraday rotator (not to scale)

This can be done as shown in Figure 3. The gyro output ν_{gyro} is given by

$$\nu_{\text{gyro}} = (\nu_4 - \nu_3) - (\nu_2 - \nu_1) = \Omega S^{-1} \quad (4)$$

where Ω is the angular input rate, ν_i is the frequency of the i^{th} mode, and S is the scale factor of the gyro.³

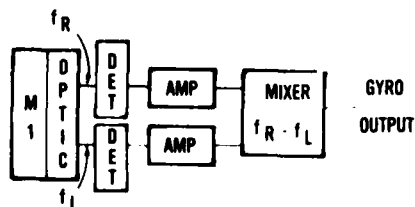


Figure 3 Possible Gyro Electronics

First Generation Clock/Gyro Operation

The first laser clock produced a clock fre-

quency given by

$$\nu_{\text{clock}} (\text{CW}) = \nu_4 - \nu_1 \quad (5a)$$

or

$$\nu_{\text{clock}} (\text{CCW}) = \nu_3 - \nu_2 \quad (5b)$$

where the $\nu_{\text{clock}} (\text{CW})$ is about 800 KHz higher than $\nu_{\text{clock}} (\text{CCW})$. The clock processing electronics are shown in Figure 4.

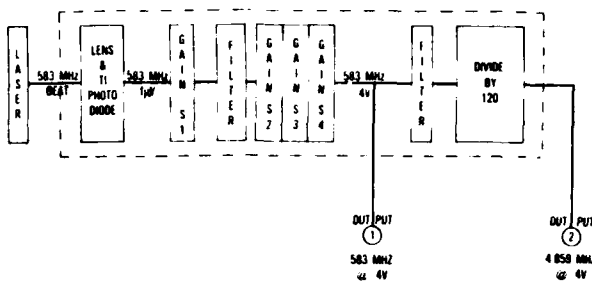


Figure 4 First Generation Clock Electronics

The two optical frequencies (having a 538 MHz beat frequency) are focused on an avalanche photodiode. The signal is amplified and filtered. The output can then be taken directly or counted down by 120.

Second Generation Clock/Gyro Operation

The clock frequency drift in the first generation laser clock was due to un-compensated temperature effects in the Faraday rotator. The Faraday frequency temperature drift $\Delta\nu_F$ is given by

$$\frac{\Delta\nu_F}{\nu_F} = (\alpha_i + \alpha_e (n-1)) \Delta T \quad (6)$$

where ν_F is the Faraday induced frequency shift,

$\Delta\nu_F$ is the Faraday frequency temperature drift, ΔT is the change in temperature, α_i is the temperature coefficient of the index of refraction and α_e is the coefficient of expansion of the Faraday rotator.³ The Faraday rotator temperature sensitivity can be compensated by use of dual detectors and implementing the second generation electronics are shown in Figure 5.

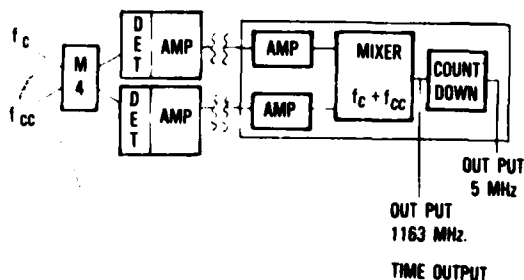


Figure 5 Second Generation Clock Electronics

The Faraday stabilized clock frequency is given by

$$\nu_{\text{clock}} = (\nu_4 - \nu_1) + (\nu_3 - \nu_2) \quad (7)$$

When the Faraday effect increases, $(\nu_4 - \nu_1)$ is increased by the same amount that $(\nu_3 - \nu_2)$ decreases, thus holding a constant sum. This should allow the clock to remain at the quantum limit for longer time periods. The quantum limited frequency uncertainty $\Delta\nu_Q(\tau)$ is given by

$$\Delta\nu_Q(\tau) = \frac{\nu_0}{Q} \sqrt{\frac{h\nu_0}{2P\tau}} \quad (8a)$$

where

$$Q = \frac{\nu_0 E}{2\pi P} \quad (8b)$$

and ν_0 is the laser frequency, h is Planck's constant, Q is a measure of the resonators ability to store energy, P is the circulating power, τ is the measuring time, and E is the stored energy.⁴ The quantum limit for the clock (QLC) is the ratio of the quantum limited frequency uncertainty to the clock frequency and is given by⁵

$$\text{QLC}(\tau) = \frac{\Delta\nu_Q(\tau)}{\nu_{\text{clock}}} \quad (9)$$

Optical Dispersion

The clock frequency in terms of the free spectral range is given by

$$\nu_{\text{clock}} = \nu_{\text{FSR}} - \nu_{d,\text{tot}} \quad (10)$$

where ν_{clock} is the clock frequency, ν_{FSR} is the free spectral range, and $\nu_{d,\text{tot}}$ is sum of the frequency pull-in for all four modes.⁶ The frequency pull-in for the i^{th} mode is given by

$$\nu_{d,i} = \frac{2|\nu_0 - \nu_i|}{\Delta\nu_D} \frac{\omega_0}{2Q} \frac{1}{1 + 4((\nu_0 - \nu_i)/\Delta\nu_D)^2} \quad (11)$$

where ν_0 is the linecenter frequency of the gain curve, ν_i is the frequency of the i^{th} mode, ω_0 is the angular frequency of linecenter, and $\Delta\nu_D$ is the normalized Doppler linewidth.⁶ The normalized Doppler linewidth is a function of discharge current and the temperature. The Q of the cavity can be a function of time if the cavity is aging. An active current controller for the discharge current might be necessary to enable the device to work over a temperature range and to compensate for aging.

III. Experiment

The experimental setup shown in Figure 6 was used to measure the stability of the laser clock frequency. The frequency measurement system used is the Hewlett Packard 5390A. The laser clock frequency was compared to an external frequency

standard by using a mixer and measuring the difference frequency using a counter. The counter measures the frequency for an integration time τ and the calculator uses this data to compute the Allan variance. The laser clock output was fed to the mixer (HP 10830A), as was the frequency synthesizer (HP 8660C) output. The reference for the frequency synthesizer and the counter was the Loran C cesium based system. The output of the mixer was fed into the counter which was controlled by the HP 9825B calculator. The calculator controls the counter and calculates the Allan variance as given in (1). The calculator outputs were produced using the HP 9871 printer/plotter. This configuration provides the highest accuracy system available.

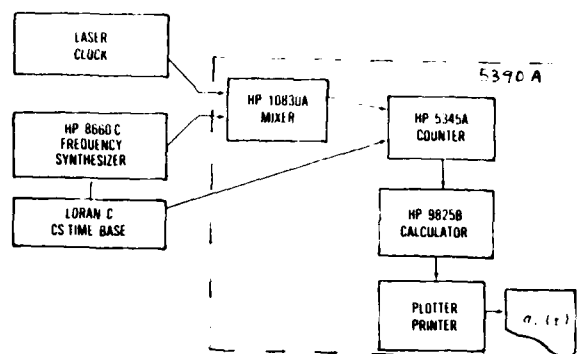


Figure 6 Experimental Setup

The test clock data for the first generation laser clock is shown in Figures 7 and 8. Figure 7 shows the data sets containing the best data points for measurement times of 1 msec, 10 msec, and 100 msec.

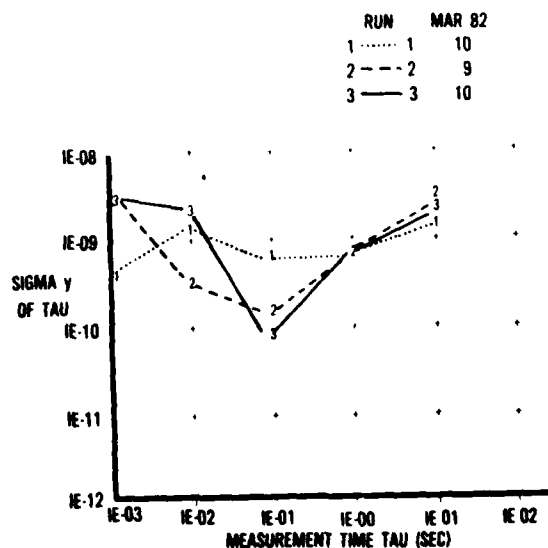


Figure 7 First Generation Clock Data

Figure 8 shows the data sets containing the best data points for measurement times of 1 sec, 10 sec, and 100 sec. The data for measurement times from 1 to 100 msec are quantum limited. The Faraday drift affects the data for measurement times greater than 0.1 sec and therefore the value of $\sigma_y(\tau)$ increased.

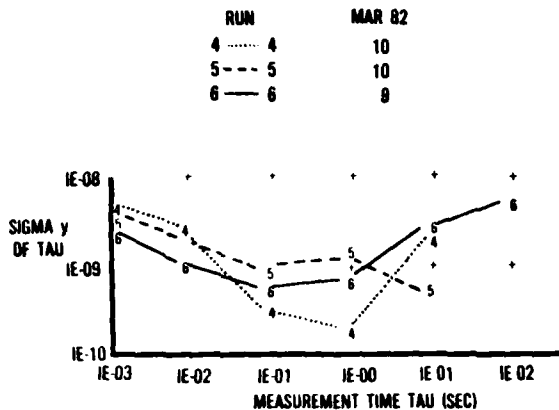


Figure 8 First Generation Clock Data

Figure 9 shows a comparison of the frequency stability of an HP 10811 quartz oscillator (Q), an HP 5065A rubidium atomic clock (R), and an HP 5062C cesium atomic clock (C). The best laser clock data falls in the region between the cesium and rubidium standards for measurement times of 1 msec to .3 sec. The Faraday drift then starts to degrade the laser clock stability.

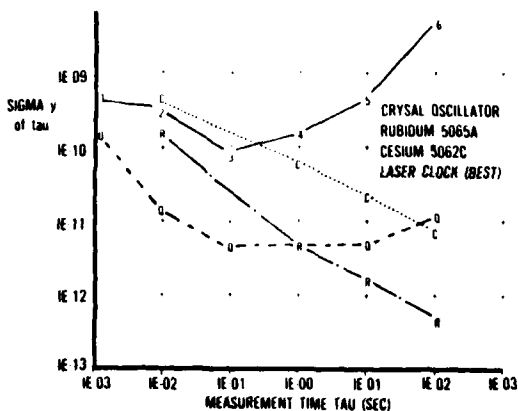


Figure 9 First Generation Stability Comparison

The long-term drift shown in Figure 10 is 30.5 mHz/sec. Figure 10 also shows that the narrowest linewidth measured was 50 mHz.

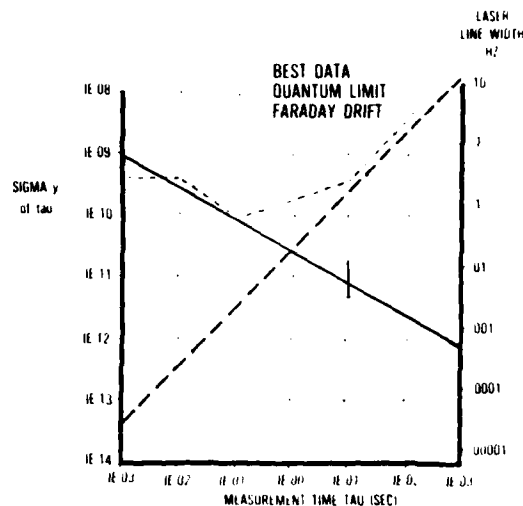


Figure 10 First Generation Error Analysis

It is estimated there was also a higher frequency fluctuation in the operating temperature as shown in Figure 11.

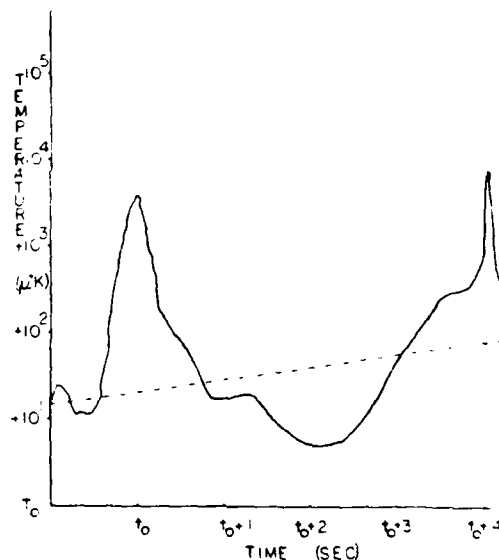


Figure 11 First Generation Estimated Temperature Drift

The short-term data was affected by this temperature fluctuation and the associated impact on the Faraday rotator but by picking the best data, this effect was minimized. The first generation longterm data always has the Faraday induced drift present; and understandably even the best data shows the drift. The single error bar in Figure 10 shows plus or minus a factor of two. The data obtained in the first test was taken while the laser was not temperature controlled. The data taken under temperature control is shown in Figure 12.

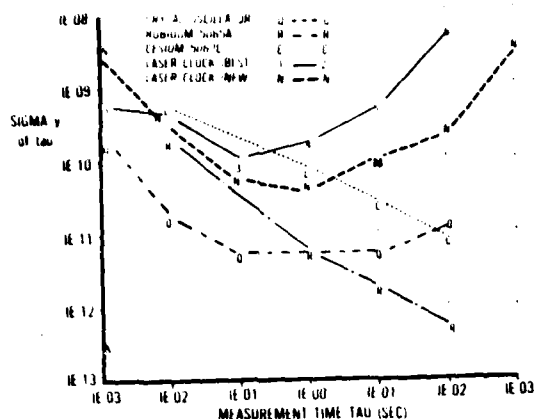


Figure 12 Test With First Generation Laser In Temperature Controlled Oven.

The second generation laser clock electronics does eliminate the error caused by the Faraday rotator. The errors due to optical dispersion effects caused by the aging of the laser and temperature are not eliminated. The next two paragraphs outline two methods of active dispersion control.

Figure 13 shows the first method of optical dispersion control. The laser power level of the right hand circularly polarized (RHCP) and left hand circularly polarized (LHCP) modes is sensed by the pathlength control circuit. The power level of either the RHCP or LHCP modes is measured by a voltmeter. The Fluke computer controls the measurement gate time of the voltmeter and reads in the voltmeter measurement via an IEEE addressable bus. The Fluke then calculates a laser discharge current correction value in digital form. This digital signal is sent to the digital to analog converter where it is converted to an analog control signal. This signal is sent to the laser discharge control circuitry which adjusts the laser discharge current depending on the analog control voltage. The control loop is implemented with an attack time which is approximately equal to the voltmeter measurement gate time plus computation time.

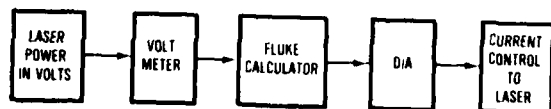


Figure 13 Method One For Active Dispersion Control

Figure 14 shows the second method of dispersion control. This method compares the laser beat frequency between adjacent modes to an external reference frequency standard and generates a control signal which is sent to the laser discharge control circuitry. In detail this is done as follows: the laser clock output is input to a mixer along with the output from a LORAN C stabilized frequency synthesizer. The output of the mixer is the beat frequency between the laser

clock and LORAN C signals. The mixer output is input to a counter which is controlled by the Fluke calculator. The calculator uses the beat frequency as an input to a program which computes the laser discharge current correction in digital form. This digital signal is sent to the D/A converter where it is converted to an analog control voltage.

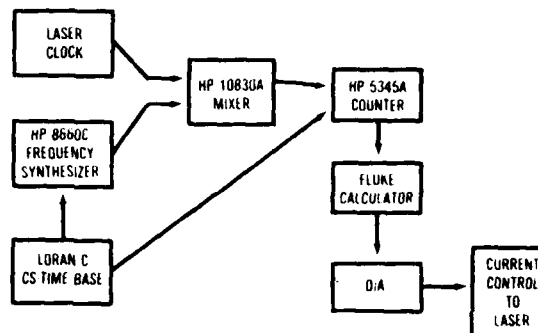


Figure 14 Method Two For Active Dispersion Control

Figure 15 shows a plot of the Allan Variance for different configurations vs. integration time in seconds. There are three sets of laser clock data. The first set of data is shown as triangles and shows the long term aging drift of 40 Hz per hour. The second data point is shown by the box and was obtained by using the first method of optical dispersion control. The third set of data is shown by diamonds and was obtained by using the second method of optical dispersion control. Figure 15 also shows the Quantum limit for the configuration tested.

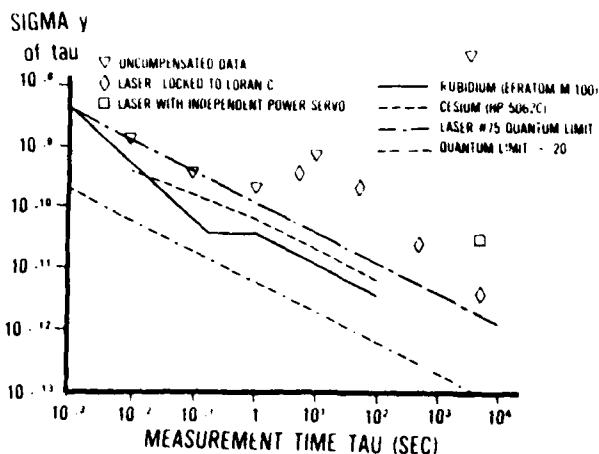


Figure 15 Second Generation Laser Clock Data

IV. Conclusions

In conclusion, good short-term stability has been obtained, but long-term stability needs to be improved. The ring laser clock is a transfer frequency standard. A transfer frequency standard used with GPS, JTIDS, and radar techniques will meet high performance fighter application

requirements so that an absolute frequency standard is not required. If implemented in conjunction with existing PLG systems, the additional cost of the clock function would be one to three thousand dollars depending upon the level of integration. This could also help reduce life cycle cost compared to using a separate clock and gyro. Careful packaging could also yield a reduced volume when compared to separate devices.

V. Recommendations

The two methods of dispersion control need to be improved and tested. Active dispersion control would allow quantum limited performance. The current quantum limited performance of the laser clock is shown in Figure 15. The other area of improvement is the quantum limit. This would require building a new laser. If the quantum limit was reduced by a factor of 20, then the clock performance would be increased by a factor of 20. This is shown in Figure 15 by the dot-dash-dash line.

VI. Acknowledgments

The experimental part of this effort could not have been completed without the help of the Raytheon Company. The experiment was conducted at their plant in Sudbury, Massachusetts. I wish to thank Jim Matthews, Al Zampello, Bill Nelson, Iri Smith, Terry Dorschner, Ed Perdue, Bob Gauthier, Dennis Berg, and the Raytheon staff.

At RADC/ESE, Hanscom AFB, I wish to thank Nick Yannoni and Ferdinand Euler for background material on clocks and timing.

At Wright-Patterson AFB (AFWAL/AAAN) I wish to thank Ron Ringo, Reference Systems Branch Chief, for supporting my effort; Bob Witters, Reference System Technology Group Leader, for his continuing support; and Jerry Covert for many discussions on clock applications and listening to laser physics discussions.

The earlier work described was performed in partial fulfillment for a Master of Science degree in Engineering Physics, AF Institute of Technology. The thesis is available through DTIC.

VII. Bibliography

1. Barnes, et al "Characterization of Frequency Stability," "IEEE Transactions on Instrumentation and Measurements," IM-20 (2): 105-120 (May 71).
2. Smith, I. W. and T. A. Dorschner. "Electromagnetic Wave Ring Resonator," U.S. Patent #4, 110,045, 29 Aug 78.
3. Smith, I. W. and T. A. Dorschner, "Biasing the Raytheon Four-Frequency Ring Laser Gyroscope," SPIE, 157 : 21-29 (1978).
4. Dorschner, T. A., et al, "Laser Gyro the Quantum Limit," "IEEE Journal of Quantum Electronics," 16 (12) : 1376 - 1379 (Dec 80).
5. Facklam, R. L. "Ultra-Stable Laser Clock" 36th Annual Frequency Control Symposium : 361 - 369 (Jun 82).
6. Facklam, R. L. "Ultra-Stable Laser Clock" MS thesis, Air Force Institute of Technology, WPAFB, Ohio, January 1983 (AFIT/GEP/PH/82d-28) available through DTIC.

MICROPROCESSOR FOR THE NR SERIES OF HYDROGEN
MASER FREQUENCY STANDARDS

Edwin E. Mengel
Johns Hopkins University/Applied Physics Laboratory
Johns Hopkins Road
Laurel, Md. 20707

Donald W. Stover
Bendix Field Engineering Corporation
9250 Route 108
Columbia, Md. 21045

Summary

The NR series of hydrogen maser frequency standards was designed by the Applied Physics Laboratory with funding and direction supplied by NASA Goddard Space Flight Center for use in the crustal dynamics investigations. In order to develop a high level of confidence that hydrogen masers were performing throughout extended weeks of field experiments, a microprocessor monitor and control system has been incorporated into the NASA-NR model hydrogen masers. This system provides control of the resonant frequency of the cavity, synchronization of the clock output, and a data link for archiving the vital performance characteristics of the maser that can be added to the data base for VLBI data processing. The monitoring and remote command features provide diagnostic information for preventative maintenance and control to set any of the maser operating parameters. The monitoring information can be observed at the local site and via phone/modems at remote locations as well. The flexibility that the microprocessor provides allows the maser frequency standard to become a complete, self-contained time and frequency instrument.

Introduction

An overall picture of the NR Maser is shown in Figure 1. The microprocessor is mounted at the top of the maser with the keyboard, displays and alarm being mounted on the front panel. A closeup picture of the front panel is shown in Figure 2. Figure 3 shows the assembled microprocessor with the large wire wrap boards tilted back as they would be for maintenance access to the system. The smaller display module is mounted on the front panel. Figure 4 is a picture of the computer board that shows several sets of switches are also included on the board. These switches are used to isolate the various separate operating sections from the computer's data bus to aid in troubleshooting processor failures.

System Capabilities

Several distinct advantages have been realized by including the microprocessor in the design of the maser:

1. The autotuning function is flexible allowing the various autotuning parameters to be customized to match the reference oscillator whether it is another maser, a cesium, an external crystal, or the internal crystal oscillator.

2. Once the aging rate of the resonant cavity has been determined, the autotuner register can be incremented or decremented on a timed basis to compensate for this drift rate.

3. The two one pulse per seconds signals can be directly controlled to 200 nsec.

4. The clock can be slewed to subnanosecond resolutions by changing the synthesizer for a specified period of time and then restoring it to its original value because frequency changes of the synthesizer have continuous phase in the output signal.

5. The remote communications capability allows verification of proper performance by knowledgeable people at a central facility without the necessity of having a maser expert at the field installation.

6. The remote communications capability allows a data base of operating masers to be developed so that long term trends can be observed using graphics to display the data.

7. The processor contains the software to input a desired fractional frequency offset and calculate the new synthesizer register settings to realize this offset. This calculation requires at least 12 decimal digits of precision which is not readily available in many calculators.

System Description

Figure 5 is a block diagram of the system showing the processor, 16K of program memory, (13K is currently used) 2K of read/write memory (RAM), and an I/O section that interfaces to the keyboard, display, alarm buzzer and three RS-232 channels. This part of the system is housed on one wire wrap module and performs the computer related functions. The display and keyboard circuitry are housed on a small wire wrap module that mounts on the front panel.

The rest of the system is on a separate wire wrap board and is used as the interface between the computer and the maser. It has the autotuner, clock, digital interface section and the analog section consisting of a 64 channel A/D converter and two D/A converters.

The microprocessor used is the Intersil 6100, a CMOS processor that uses the PDP-8E/F instruction set. This processor was selected because the only other low power competitor at the time of selection (1977) was the RCA 1802 which has a more awkward hardware interface and did not have the extensive software support that was available for the PDP-8E/F.

It was decided in the early stages of the design process to include the microprocessor and this allowed us to include various sensors, extra switch contacts and other controls in the original design instead of retrofitting these features.

The power consumption of the entire processor system is seven watts with the displays off and 10 watts with the displays on. The processor can detect the loss of a-c power and turn off the displays to conserve battery capacity.

Autotuner

The autotuner's function is to measure the frequency of the maser with a high flux of hydrogen atoms entering the cavity and then with a low flux of hydrogen atoms entering the cavity. If there is a difference between the two frequencies, then the 18 bit cavity control register is changed. This results in the cavity thermal control servo changing the size of the aluminum cavity and bringing the cavity to its tuned condition. Long term drifts of the cavity control elements (resistors, thermistors, capacitors...) make the tuning capability a requirement. The autotuner circuitry outputs an 18 bit word to the cavity control circuitry and a Hi/Low signal to the flux control circuitry in the maser. The software allows the operator to control the various parameters of the autotuning process such as the time durations of the measurement sequence and the gain of the cavity control servo. The software also provides a means to automatically change the cavity register on a timed basis to remove predictable drift rates.

The period measurement circuitry that is included in the autotuner and the software that controls the autotuning parameters can be used to allow the maser to evaluate the stability of other oscillators. At the Laboratory we are programming an Apple computer to control the measurement process and make the stability calculations.

Clock

The clock consists of two counters that produce one pulse per second signals from 5 MHz. Each counter's output can be shifted in time directly by the software in steps of 200 nsec. The two counters can be synchronized and a hardware alignment will allow the two signals to agree to 1 nsec. Software control of the synthesizer allows the clock pulse to be slewed to subnanosecond resolution. Both clocks are slewed together during this process. The software displays UTC from seconds to year on the front panel displays and will output the time of occurrence of a signal on the time transfer input to a resolution of 200 nsec. The UTC time is printed in all status requests as a self-documenting feature of the status printout.

Digital Interface

This section interfaces to the synthesizer and the various digital indicators from the maser. The digital indicators include the flux control settings for the shutter and state selector and separate contacts on the various control switches in the maser. The software interrogates these items and displays them on the front panel or in the status message. The microprocessor can read the settings of the A, B, and C counters in the synthesizer and can send new values to change the counters under operator

command. Software is included to allow an operator to input a desired fractional frequency shift so the microprocessor can calculate the new A, B, and C values to effect this change and can send these new values to the synthesizer. Figure 8 is a list of the calculations performed during this operation. The capability to calculate and change the synthesizer was also used in a software routine that allows the operator to slew the one pulse per second signal to subnanosecond resolution.

Synthesizer Calculations

$$k_o = \frac{1}{C_o} - \frac{1}{B_o C_o} - \frac{1}{4A_o B_o C_o} \quad k = \text{synthesizer division ratio}$$

$$f = \text{fractional frequency change} \quad ABC = \text{synthesizer settings}$$

$$k = k_o + f$$

$$C = \text{Int} \frac{1}{k}$$

$$B = \text{Int} \frac{1}{1 - Ck} + 1 \quad 750 \leq B \leq 9999$$

$$A = \text{Int} \left[\frac{1}{4(B(1 - Ck) - 1)} + .5 \right] \quad 170 \leq A \leq 3750$$

Precision required - 12 decimal digits - 40 bits

FIGURE 8
SYNTHESIZER CALCULATIONS

Analog Circuitry

This section consists of a 64 channel 12-bit A/D converter and two 12-bit D/A converters. All of the significant maser signals are input to the A/D converter circuitry to allow the operator to monitor them on the display, individually, or on a terminal sequentially. These signals include system voltages, power supply voltages, currents, temperatures and the IF signal level. The software checks each reading against preset limits and activates an alarm or gives an indication on the printout if a limit is exceeded. All readings are converted to engineering units. The temperature calculations are customized to the differences of the eight individual thermistors used and the IF curve is customized for each maser. One external signal is input to the system so that an external thermistor can be monitored for a local environment temperature reading.

At times it is desirable to display a signal on a strip chart recorder. Two D/A converters are provided and any of the analog signals that are input to the A/D circuitry can be sent out to one or the other of the D/A converters. Since the data is in digital form, it can be scaled so that small variations can be amplified and displayed full scale on the strip chart recorder. It is also possible to output two digital measurements to the D/A converters; the 18-bit autotuner register and the 36-bit autotuner measurements. The 12 bits that are sent to the converter can be selected by the operator so the variations can be scaled to be observable.

Diagnostic Capabilities

The microprocessor system is designed to have a bus structure with all elements of the system communicating over the same data lines and using the same control signals. In the event of a failure that pulls one of these lines to ground, the location of the failed element could be difficult to isolate. Switches and software have been included with this system to implement signature analysis as an aid to troubleshooting these types of failures. Switches are provided so that each section of the microprocessor can be removed from the bus to isolate the failed chip. All of the feedback paths to the microprocessor can be eliminated so that the basic operations and external address register can be verified. Special routines are provided in PROM to verify the proper operation of all of the instructions that the microprocessor can use, the data in the ROM, the operation of the RAM and the paths to output data on the bus.

In addition to the signature analysis, additional routines are provided to verify the proper operation of the various elements from a system perspective. These routines are used for bench checkout of the microprocessor system and include checkout of the displays, keyboard, terminal out on channel 1, terminal in on channel 1, RAM Hi, RAM Low, digital to analog #1, digital to analog #2, digital input data, general purpose output circuitry, clock A and B, time transfer test and autotuner test. This program can be used with or without a terminal, since the sequencing commands are input through the keyboard. However, a terminal provides the operator with more information than is available on the displays.

Long Term Drift Correction

Mature masers typically experience a long term drift that requires them to be retuned periodically to bring them back on frequency and reduce their susceptibilities to other environmental influences. Once a maser has matured, it can be characterized by a linear drift rate. The software in the maser allows a known rate to be inserted and the autotuner cavity can be changed to compensate for the known drift rate.

Figure 6 is a plot showing the change of the 18 bit cavity register for NR-1 for a period of 120 days when this maser was tuning continuously. The straight slope is obvious up to day 252 when the temperature control of the room broke down. After the room was brought back under control, it can be seen that the register recovered to its former value and that rate remained constant. This demonstrates the aging rate typical of this maser design.

Figure 7 demonstrates the improvement in the long term Allan Variance that could be realized by removing the aging rate. The raw data is between masers NR-3 and NR-6 when both masers were in environmental chambers and NR-6 had a rate removed between periods of autotuning. A least squares fit was made to the data and the second curve shows the Allan Variance remaining with the first order term removed. For this measurement it appears that the long term drift of $4 \pm 2 \times 10^{-15}$ could be achieved by using the rate compensation.

Communication Capabilities

Three RS-232 channels and the front panel keyboard and display can be used to input commands and display the status and autotuning results. A local terminal can be connected to the RS-232 channels to allow a full printed record of the performance of the maser to be maintained. Although more information may be provided to a terminal, complete control is available through the keyboard on the front panel as well as with a terminal. A computer can be connected to one of the channels that would allow the record to be stored electronically. A phone modem can be connected to one of the RS-232 channels to allow a person at a remote site with a modem and terminal to printout the status and send commands to the maser. This ability to be able to remotely printout all of the critical signals and compare to earlier printouts allows a knowledgeable person to be able to diagnose an improperly operating maser without having to be at the maser location.

If a computer can communicate with the maser, either through a direct connection or remotely via phone modems, then the status information can be archived and displayed in a manner which allows long term trends can be observed. Figure 9 shows the power supply history for NR-4 during the year 1981 using a special graphics program. There are eight power supplies shown (the + and -18 volt supplies overlap) and at a glance it can be seen that there were no problems experienced or developing in these items. Figure 10 is a plot of various system currents for NR-4. The top curve shows the variations of the vacuum pump current and on day 152 the pump failed and was replaced.

Conclusion

The autotuner software and hardware provides a flexible autotuner control that allows the autotuner process to be customized to the type of reference oscillator used. In addition, the control of the frequency measurement intervals and the printout of the results can allow an external computer to measure the stability of another oscillator.

The use of a microprocessor answered several needs that the Laboratory saw when taking care of the previous generations of masers in the field. In order to diagnose a problem, it was necessary to send an experienced person to the field site trained to operate and repair masers. By providing a means of remotely monitoring and controlling the masers, the experts can quickly deduce the problem and initiate the correct maintenance actions rather than having to visit the site or ship the maser back for diagnosis.

The ability to collect data at a remote site also allows the data to be archived and displayed in a manner which allows long term trends to be observed.

The ability to observe performance over long periods of time have provided a continuing basis for product performance and reliability improvements under typical field operations.

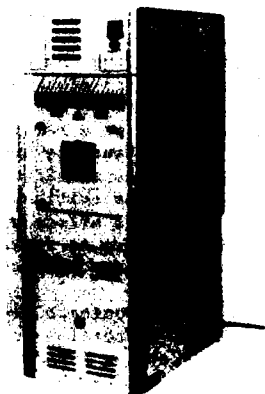


FIGURE 1
NR-MASER

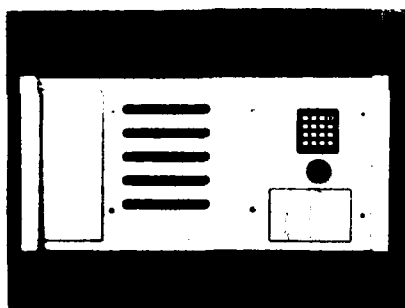


FIGURE 2
MICROPROCESSOR FRONT PANEL

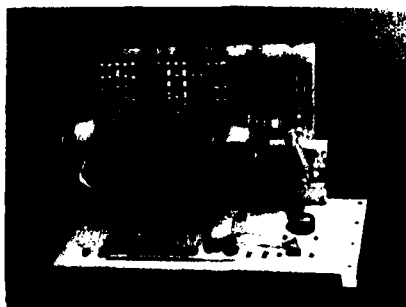


FIGURE 3
MICROPROCESSOR-INTERNAL VIEW

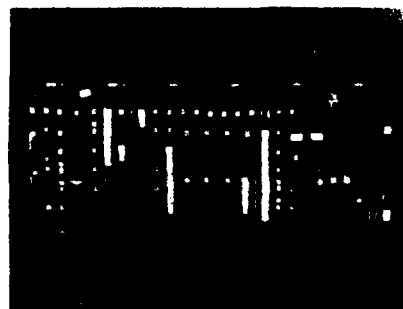


FIGURE 4
CPU BOARD FOR MICROPROCESSOR SYSTEM

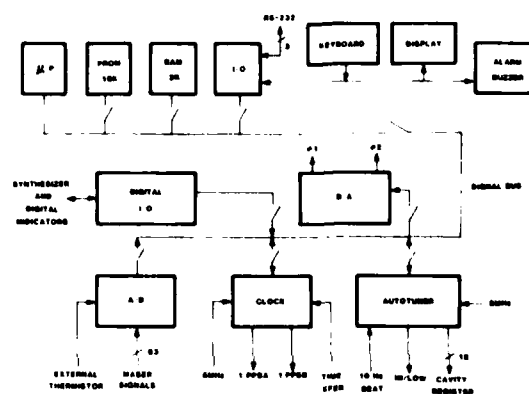


FIGURE 5

NR-1 LONG TERM DRIFT CORRECTIONS TO CAVITY IN AUTOTUNE MODE

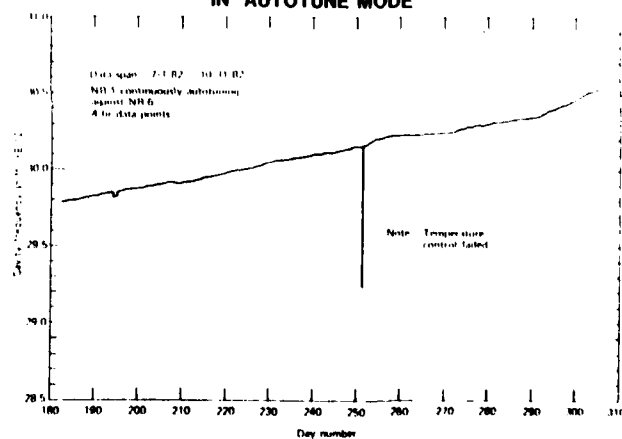


FIGURE 6
NR-1 LONG TERM DRIFT CORRECTIONS
TO CAVITY IN AUTOTUNE MODE

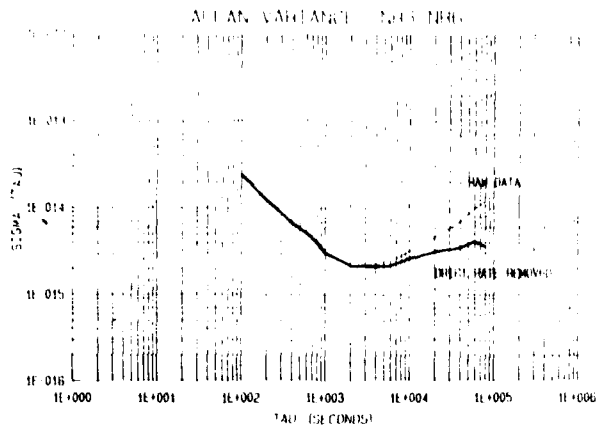


FIGURE 7
ALLAN VARIANCE WITH AND WITHOUT
DRIFT REMOVED

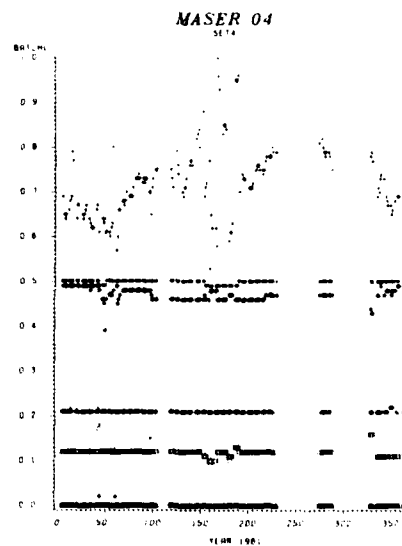


FIGURE 10
HISTORY OF SYSTEM CURRENTS FOR NR-4

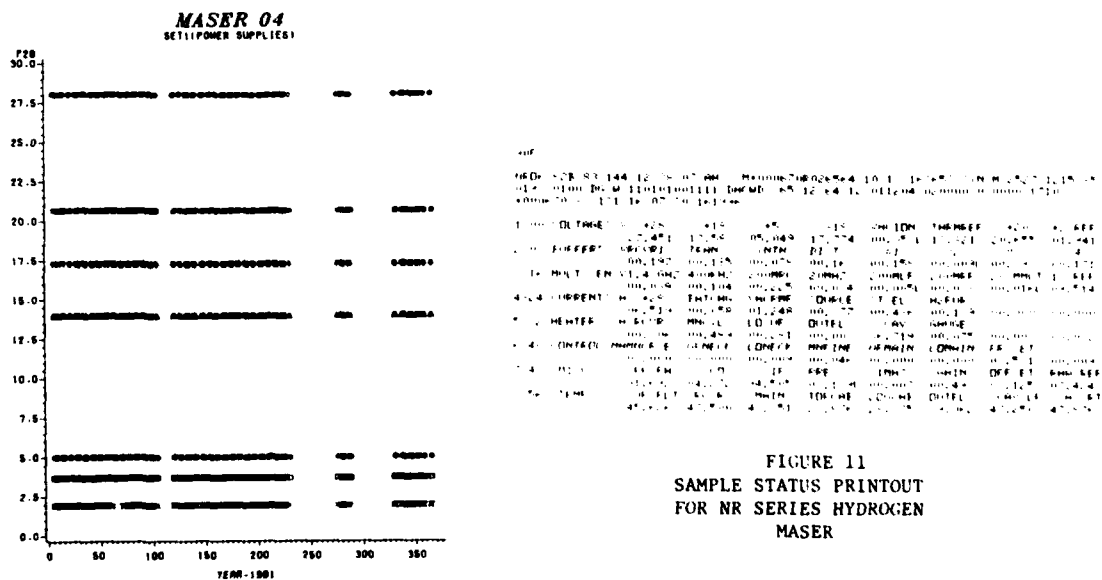


FIGURE 11
SAMPLE STATUS PRINTOUT
FOR NR SERIES HYDROGEN
MASER

FIGURE 9
POWER SUPPLY HISTORY OF NR-4
DURING 1981

MERCURY-199 TRAPPED ION FREQUENCY STANDARD: RECENT THEORETICAL PROGRESS AND EXPERIMENTAL RESULTS

L. S. Cutler, R. P. Giffard, and M. D. McGuire
Physical Sciences Laboratory
Hewlett-Packard Laboratories
3500 Deer Creek Road
Palo Alto, CA 94304

SUMMARY

The 40.5 GHz hyperfine resonance of ^{199}Hg ions stored in an rf quadrupole trap has been shown to have a very small fractional linewidth, suggesting its use as a high-precision frequency standard. The most significant offset in such a standard would be the second-order Doppler shift resulting from the motion of the stored ions. We have recently analyzed the situation in which the secular motion is cooled to a temperature of about 300K by the presence of a light background gas at low pressure, resulting in an ion cloud whose density is almost completely determined by the balance of pseudopotential and space-charge forces. Under these circumstances we have found that the second-order Doppler shift can be calculated accurately from the trapping parameters, the temperature, and the total number and mass of the stored ions.

We have carried out a large number of experiments using Helium as the cooling gas, and have obtained an rf pulse-length limited linewidth of 0.85 Hz. The lineshape observed implies a natural width smaller than 0.1 Hz. Closed-loop regulation of the number of stored ions has allowed us to demonstrate internal consistency of the observed second-order Doppler shift at the level of a few parts in 10^{13} .

INTRODUCTION

The precision of atomic frequency standards is fundamentally limited by the length of time during which individual atoms interact with the resonant radio-frequency field. Since the development of ion trapping by electromagnetic fields, and its application to microwave spectroscopy [1,2], it has been recognized that the use of this technique might permit relatively long collision-free interaction times. If the environment provided by the trap were sufficiently benign, it would be possible to make atomic frequency standards of considerably increased accuracy and performance [3-7].

Following a suggestion of F. G. Major, a number of experiments have now been carried out on rf-trapped $^{199}\text{Hg}^+$, optically pumped by resonance radiation from a ^{202}Hg lamp [8-14]. The results of these experiments confirm the potential of this system, and underline the necessity to control the second-order Doppler shift resulting from ion motion.

In the rf or "Paul" trap the average force applied to an ion by quadrupole electric fields is equivalent to a three-dimensional, approximately harmonic, potential well whose depth is determined by the strength of the fields and the size of the apparatus. If no attempt to cool the ions is made, it has been found experimentally [10,15-17] that the average kinetic energy of the ions is about 10% of the well depth, which typically exceeds 10 eV. For $^{199}\text{Hg}^+$ the

fractional second-order Doppler shift amounts to 5×10^{-12} per eV of kinetic energy, and the ion motion therefore leads to a substantial frequency shift which is difficult to model accurately.

We have recently analyzed the situation in which the secular motion of a number of interacting ions moving in the pseudopotential well is assumed to have been cooled to a temperature of about 300K. The results of this analysis, which will be presented in detail below, show that the average second-order Doppler shift for a given ion mass depends on the number of ions trapped, the amplitude and frequency of the trapping fields, and, to a relatively small extent, on the motional temperature. Thus, for a given number of trapped ions, the second-order Doppler shift is constant and can be calculated from the design and operating conditions of the experiment.

In practice we have been able to cool ^{199}Hg ions adequately using a low pressure of Helium gas in the trap. The frequency shift due to collisions with the cooling gas is small. Accurate closed-loop regulation of the number of ions trapped has enabled us to demonstrate full consistency with the cooled ion cloud theory. These results indicate that by using cooled rf-trapped ^{199}Hg ions one could construct a compact frequency standard whose accuracy would constitute an improvement of approximately an order of magnitude over existing portable Cesium standards. The requirements are somewhat different from those recently discussed by Wineland [6] who shows that to approach a fractional accuracy of 10^{-15} , laser cooling of ions in a Penning trap would be essential.

THEORETICAL DESCRIPTION OF COOLED ION CLOUD

In an rf quadrupole trap a single ion feels an average force generated by the superposition of oscillating and static electric fields. In an ideal trap the instantaneous electric potential $\phi(t)$ is given in cylindrical coordinates by

$$\phi(t) = (U - V \cos \Omega t)(r^2 - 2z^2)/\zeta^2, \quad (1)$$

where U and V are the dc and rf voltages applied, and ζ is a dimensional parameter characterizing the trap electrodes. It has been shown [1,2] that the effect of the oscillating part of the field is to create an approximately harmonic restoring force on an ion as long as its displacement during the rf cycle is small compared with its distance from the center of the trap. This is equivalent to a cylindrically symmetric pseudopotential $\psi(r,z)$ given by

$$\psi(r,z) = (q^2 V^2 / m \Omega^2 \zeta^4)(r^2 + 4z^2), \quad (2)$$

where q is the charge and m is the mass of the ion.

The superposition of the field due to the dc voltage U and the pseudopotential creates an effective potential $\phi(r, z)$. The general form of this effective potential is

$$\phi(r, z) = Ar^2 + Bz^2, \quad (3)$$

where the coefficients A and B depend on U, V, ω , and the charge to mass ratio of the ion. Although any shape of effective potential can be used, it is convenient theoretically and experimentally to choose values of the parameters, U, V and ω yielding the spherical potential corresponding to $A = B$. Equation (3) may then be written in the form

$$\phi_s = (m\omega^2/2)(r^2 + z^2) \quad (4)$$

where

$$\omega = 2qV/m\omega_c^2 = (4qV/m\omega_c^2)^{1/2}.$$

In the effective potential (4), a single ion would undergo three-dimensional simple harmonic motion at a characteristic micromotion frequency ω . Superposed on this motion is a driven oscillation at frequency Ω in the direction normal to the equipotential surfaces of equation (2). The mean square velocity $\langle v^2 \rangle$ of this micromotion is given by

$$\langle v^2 \rangle = (\omega^2/2)(r^2 + z^2) \quad (5)$$

Both the macromotion and the micromotion of a single ion in the potential well are therefore completely described by the macromotion frequency parameter ω whose value is implied in equation (4).

If the well contains a large number of ions, the effective potential is further modified by the ionic space-charge. The total effective potential ϕ_{tot} seen by an ion is then given by

$$\phi_{tot} = \phi(r, z) + q\phi_q(r, z) \quad (6)$$

where the electric space-charge potential $\phi_q(r, z)$ is related to the ionic charge density ρ by Poisson's equation. The simplest conditions under which to solve equation (6) for the ionic charge density correspond to the assumption that the ion temperature is low enough for the micromotion energy to be negligible. The space charge density will then adjust itself so that the gradient of the total potential is zero everywhere within the volume occupied by the ions [1,18]. For the spherical potential of equation (4) the result is a spherical ion cloud with a uniform number density n given by

$$n = 3\epsilon_0 m\omega^2/q^2. \quad (7)$$

The total number N of ions in the cloud and the cloud radius r_c are thus related by $4\pi r_c^3 n/3 = N$. In a typical ion trap ω might be of the order of $2\pi \times 50$ kHz resulting in a cold ion cloud density of $3.4 \times 10^4 \text{ mm}^{-3}$ for mercury ions. A cloud of 10^6 ions would thus have a radius of 1.9 mm. It appears that in all practical cases quantum-mechanical effects will not significantly affect the distribution of ions.

The assumptions made above allow one to calculate the second-order Doppler shift averaged over all positions in the cloud. Since the ion velocities are small compared with c , the fractional shift due to micromotion is given by

$$\Delta f/f = -\langle v^2 \rangle / 2c^2 \quad (8)$$

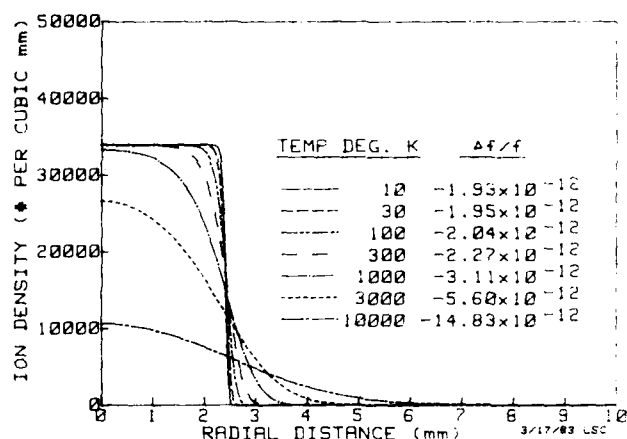


Fig. 1. Calculated number-density distributions for 2×10^6 $^{199}\text{Hg}^+$ ions in a spherical effective potential with $\omega = 2\pi \times 50$ kHz. It is assumed that the ion cloud can be described by a temperature T . $\Delta f/f$ is the total second-order Doppler shift.

where the spatial average denoted by the bar is carried out for $(r^2 + z^2) < r_c^2$. Integration of equation (5) thus gives the micromotion second-order Doppler shift for a spherical cold cloud of N ions

$$\Delta f/f = -(3/10c^2)(\omega N q^2 / 4\pi\epsilon_0 m)^{2/3}. \quad (9)$$

Normalizing to typical values of ω and N we obtain for ^{199}Hg ions

$$\Delta f/f = 1.2 \times 10^{-12} (\omega/2\pi \times 50 \text{ kHz})^{2/3} (N/10^6)^{2/3}. \quad (10)$$

Using a method recently suggested by Knight [19] we have extended the calculation to allow a non-zero macromotion temperature T in a spherical pseudopotential. A self-consistent calculation is used to find the ion number density function $n(r, z)$ which satisfies a Boltzmann distribution. The results we have obtained from this numerical calculation are shown in Figure 1 for a cloud of 2×10^6 mercury ions in a spherical pseudopotential with $\omega = 2\pi \times 50$ kHz. At low temperatures the density approaches a rectangular distribution with n given by equation (7), and at high temperatures the density approaches the gaussian distribution which has been observed experimentally in the absence of ion cooling [15-17].

The complete density distributions can be used to calculate the second-order Doppler shift. The modified number-density weights the integral involved in calculating the average micromotion energy, and thermal kinetic energy equal to $3k_B T/2$ is added. Data in Figure 1 shows how the calculated shift varies with temperature for a typical ^{199}Hg ion cloud. The rate of change of the shift with temperature at 300K is $1.2 \times 10^{-15}/\text{K}$, and the total shift at 300K is about 18° greater than that given by equation (9).

EXPERIMENTAL ARRANGEMENT

We have previously described the apparatus and reported the results of exploratory measurements [14]. For the experiments to be described below a continuous source of neutral atomic mercury-199, and means for monitoring and controlling the pressure of the helium cooling gas have been added. Magnetic field gradients, which previously led to significant relaxation,

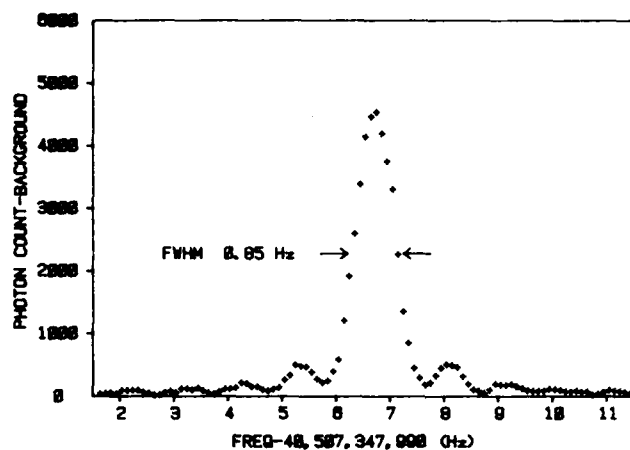


Fig. 2. Hyperfine resonance of $^{199}\text{Hg}^+$ ions measured with a microwave pulse length of 1.0s. The line is the average of 50 sweeps, and the trap contained about 2×10^6 ions. The frequency has been corrected for the 0.1 gauss magnetic field.

have been considerably reduced, and the average magnetic field is now regulated. The measurement cycle has been extended from 2.0 to 2.75s to permit a 1.0 sec interaction time with the microwave signal. Light shift is eliminated by turning off the pumping light during the interaction time and randomizing the microwave phase between pulses.

An important development is a means for monitoring the total stored ion charge: such measurements have been described previously [2, 19-22], although those which empty the trap are clearly not useful in this application. In our system a coherent oscillation of the ion cloud is excited by a short voltage burst at the micromotion frequency ω applied to the cap electrode of the trap. The free decay of the ion cloud oscillation is then observed by means of image currents induced in the cap circuit. In this way, a measurement of the trap population with a typical precision of about 2% can be made during each 2.75s cycle of the frequency standard without disturbing its operation. Closed-loop control of the trap population is achieved by varying the current in the electron gun pulses used to generate new ions.

The output of the closed-loop ion frequency standard is monitored by continuous phase comparison with a local Cesium standard which is itself compared using TV line 10 with an ensemble of 5 such standards located at the Santa Clara division of Hewlett-Packard.

RESULTS

Figure 2 shows an example of the lineshape obtained using a 1 second rf pulse. The shape is close to the expected Rabi line, implying that relaxation effects are small. The derivative of the response with respect to frequency, and the average photon count of 2.2×10^4 , imply a statistical fractional frequency noise of 0.42×10^{-12} for a single measurement cycle of 2.75s. All of the results discussed in this paper have been obtained in the presence of a background of Helium gas. As expected, cooling the ions has been found to increase the ion storage time and the stored ion density while stabilizing the second-order Doppler shift. With a Helium pressure of 1.5×10^{-5} torr the ion storage time was measured to be about 2.5×10^4 s under typical operating conditions.

Figure 3 shows the effect of varying the cooling gas pressure on the frequency of the line-center with the trap population maintained approximately constant. As the pressure rises to 10^{-5} torr the frequency increases rapidly due to the cooling of the micromotion and the reduction of the micromotion shift as the ion cloud contracts. At higher pressures the data approaches a linear dependence with a slope of 5.7×10^{-9} per torr resulting from the shift of the mercury hyperfine frequency due to collisions with the cooling gas. The effect observed is of the same order as that for $^{137}\text{Ba}^+$ ions in He [23].

Further evidence that the secular motion of the ion cloud can be effectively cooled was provided by observations of the first-order Doppler sidebands which result from modulation of the microwave phase by ion motion. The sideband spectrum was found to consist of two broad peaks symmetrically shifted from the narrow central line. The measured sideband frequency shift was in agreement with the approximate value $(3k_B T/m)^{1/2}/4r_c$, where $T=300\text{K}$, predicted by a model in which fully thermalized ions make elastic collisions with a spherical barrier without disturbances in phase, and collisions with the cooling gas are relatively infrequent.

In order to investigate the validity of the cold cloud model result given in equation (9), we measured the closed-loop frequency as a function of the trap population. The experiments were performed at five different micromotion frequencies on approximately spherical clouds. The data shown in Figure 4 demonstrates that the shift varies with the two-thirds power of the trap population as expected. The frequency has been corrected for magnetic field and background gas effects. Extrapolation of the data to zero population enables us to estimate the frequency at zero micromotion shift. The zero population intercepts of linear regressions for the 5 different micromotion frequencies studied agreed to within $\pm 1 \times 10^{-13}$, although the slopes do not vary exactly with $\omega^{2/3}$ as would be expected from equation (9).

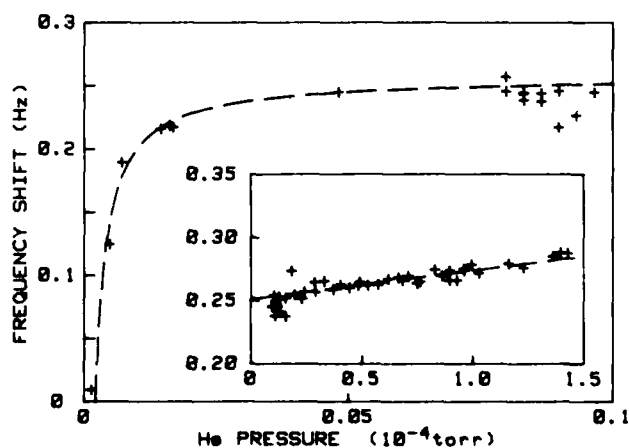


Fig. 3. Observed variation of the line-center frequency f as a function of the pressure p of the He cooling gas. The inset shows the points for pressures exceeding 10^{-5} torr on a reduced scale. The dashed line represents a fit to the data given by $f = f_0 + 233p - 5.5 \times 10^{-8}/p$ based on a simple model of the variation of cloud temperature with pressure and a linear collision shift. The zero of the vertical axis is arbitrary.

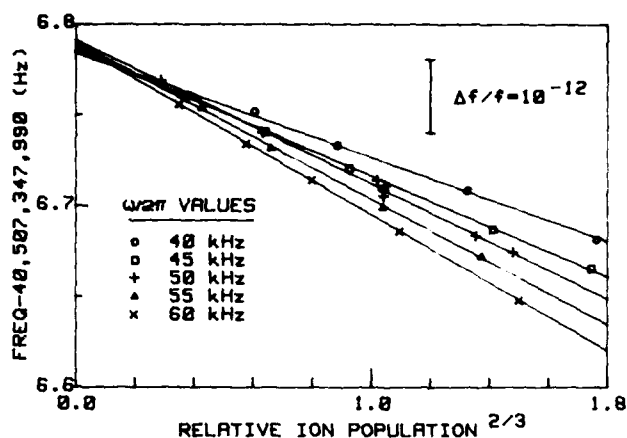


Fig. 4. Observed variation of the line-center frequency with relative ion population to the two-thirds power for various values of the micromotion frequency ω . The data is for approximately spherical clouds, and the frequency has been corrected for magnetic field and He pressure.

CONCLUSIONS

The results presented above demonstrate the effectiveness of a light background gas in cooling a cloud of $^{199}\text{Hg}^+$ ions. The cold cloud model of equation (9) is found to give a good account of the variation of the second-order Doppler shift with the total stored charge. The Stark effect has an identical variation with trapping parameters, and is comparatively small [24]. By controlling the experimental conditions we estimate that the Doppler shift can be regulated to better than 1%. Using the finite temperature model it should be possible to calculate the total shift for a given set of operating conditions with an uncertainty of less than 10%. Although the shift caused by the cooling gas cannot be calculated from first principles, it is typically small ($\leq 1 \times 10^{-13}$), and the gas pressure can be regulated to a satisfactory degree. We have not been able to detect any line asymmetry or rf power shift at the level of 10^{-13} .

The performance which could be obtained in practice with a $^{199}\text{Hg}^+$ frequency standard would depend on the chosen tradeoff between systematic errors and signal-to-noise ratio.

Table 1 shows the systematic and random effects that might characterize the performance of a portable $^{199}\text{Hg}^+$ standard and estimates of the accuracy which might be achieved. The accuracy and precision are considerably better than those of currently manufactured Cesium frequency standards.

ACKNOWLEDGMENTS

We thank J. A. Marshall of HP-Santa Clara for his cooperation in establishing the local time scale. D. A. Weigel constructed the time comparison system and much of the circuitry of the $^{199}\text{Hg}^+$ frequency standard.

TABLE 1

Effect	Magnitude	Precision	Accuracy
Micromotion (Doppler Shift)	-2×10^{-12}	$\pm 1.3 \times 10^{-14}$	$\pm 2 \times 10^{-13}$
Thermal Shift (Doppler Shift)	-2×10^{-13}	$\leq 10^{-14}$	$\pm 2 \times 10^{-14}$
Statistical Noise	$10^{-12} - 1/2$	$10^{-12} - 1/2$	-
Asymmetry	None Observed	$\leq 10^{-14}$	$\leq 10^{-13}$
rf Power	None Observed	$\leq 10^{-14}$	$\leq 10^{-13}$
Stark Shift	1×10^{-15}	-	-
Stability	$\{10^{-24}, -1, +5 \times 10^{-28}\}^{1/2}$		
Accuracy	2.5×10^{-13}		

Table 1 shows estimates of the fractional frequency errors which might be obtained in a portable $^{199}\text{Hg}^+$ frequency standard. The table shows the magnitude of each effect, the precision with which the effect might be controlled over time, and the absolute accuracy which might be obtained in manufacture. The totals have been obtained assuming no correlation between the various effects.

REFERENCES

1. H. G. Dehmelt, Advan. At. Mol. Phys. 3 53 (1967).
2. H. G. Dehmelt, Advan. At. Mol. Phys. 5 109 (1969).
3. M. D. McGuire, Proceedings of the 31st Annual Symposium on Frequency Control (1977), Electronic Industries Assoc, p.612.
4. D. J. Wineland, Proceedings of the 11th Annual PTTI Appl. & Planning Meeting (1979), NASA Conf. Publ. 2129, p.81.
5. D. J. Wineland, W. M. Itano, J. C. Bergquist and F. L. Walls, Proceedings of the 35th Annual Symposium on Frequency Control (1981), Electronic Industries Assoc, p.602.
6. D. J. Wineland, Proceedings of the 13th Annual PTTI Appl. & Planning Meeting (1981), NASA Conf. Publ. 2220, p.563.
7. L. Maleki, Proceedings of the 13th Annual PTTI Appl. & Planning Meeting (1981), NASA Conf. Publ. 2220, p.593.
8. F. G. Major and G. Werth, Phys. Rev. Lett. 30, 1155 (1973).
9. F. G. Major and G. Werth, Appl. Phys. 15, 201 (1978).

10. M. D. McGuire, R. Petsch and G. Werth, Phys. Rev. A 17, 1999 (1978).
11. M. Jardino and M. Desaintfussien, IEEE Trans. Inst. Meas. IM-29, 163 (1980).
12. M. Jardino, M. Desaintfussien, R. Barillet, J. Viennet, P. Petit and C. Audoin, Appl. Phys. 24, 107 (1981).
13. M. Jardino, M. Desaintfussien, and F. Plumelle, J. de Physique 12, Colloque C8, 327 (1981).
14. L. S. Cutler, R. P. Giffard and M. D. McGuire, Proc. 13th Annual PTTI Appl. and Planning Meeting (1981). NASA Conf. Publ. 2220, p.563.
15. R. Ifflander and G. Werth, Metrologia, 13, 167 (1977).
16. R. D. Knight and M. H. Prior, J. Appl. Phys. 50, 3044 (1970).
17. H. Schaaf, V. Schmeling, and G. Werth, Appl. Phys. 25, 249 (1981).
18. D. J. Wineland in *Precision Measurements and Fundamental Constants II*, Eds. B.N. Taylor and W.D. Phillips, NBS Special Publication 617 (in press).
19. R. D. Knight, PhD Thesis, U. of California, Berkeley (1979). Lawrence Berkeley Lab Publication LBL 9082.
20. H. D. Schuessler and C. H. Holder, J. Appl. Phys. 50, 5110 (1979).
21. M. N. Gaboriaud, M. Desaintfussien, and F. G. Major, Int. J. Mass Spectr. and Ion Phys. 41, 109 (1981).
22. F. L. Walls, D. J. Wineland, and R. E. Drullinger, Proc. of the 32nd Annual Symp. Frequency Control (1978), Electronic Industries Assoc, p.453.
23. J. Vetter, M. Stuke and E. W. Weber, Z. Physik, A273, 129 (1975).
24. W. M. Itano, L. L. Lewis and D. J. Wineland, Phys. Rev. A 25, 1233 (1982).

LASER COOLED ${}^9\text{Be}^+$ ACCURATE CLOCK

J. J. Bollinger, Wayne M. Itano, and D. J. Wineland
Time and Frequency Division
National Bureau of Standards
Boulder, Colorado 80303

Summary

The use of laser cooled stored ions in an atomic frequency standard has the potential of very high accuracy because Doppler effects are greatly suppressed. A clock based on the ground-state hyperfine transition in ${}^{201}\text{Hg}^+$ has potential accuracy and stability exceeding 1 part in 10^{15} . However, laser cooled ${}^9\text{Be}^+$ ions are experimentally easier to obtain. Therefore a ${}^9\text{Be}^+$ based frequency standard is investigated in order to study the generic problems of laser cooled stored ion frequency standards. Approximately 300 ${}^9\text{Be}^+$ ions are stored in a Penning trap and laser cooled. The 303 MHz ground state $(M_J, M_I) = (-3/2, 1/2) \rightarrow (-1/2, 1/2)$ nuclear spin flip hyperfine transition is observed at a magnetic field ($\sim 0.82\text{T}$) where the transition frequency is independent of magnetic field to first order. The time domain Ramsey method of interrogation is used and yields a linewidth of 25 mHz. The stability of an oscillator locked to this transition has been measured for $400 \leq \tau \leq 3200\text{s}$ to be $\sigma_y(\tau) \approx 2 \times 10^{-11} \tau^{-1/2}$. By measuring the velocity distribution of the ions, the second-order Doppler shift is determined to be on the order of 5×10^{-14} . The magnetic field instability contributes a 3×10^{-14} uncertainty in the present experiment. All other systematic uncertainties are estimated to be no greater than 3×10^{-14} .

Introduction

Because stored ion techniques provide long confinement times with minimal perturbations, they provide the basis for improved time and frequency standards¹. Several groups have sought to develop a microwave frequency standard based on ${}^{199}\text{Hg}^+$ ions in an rf trap²⁻⁵. The choice of the ${}^{199}\text{Hg}^+$ ion for a microwave stored ion frequency standard is based on its 40.5 GHz ground-state hyperfine separation, which is the largest of any ion which might easily be used in a frequency standard (hence high Q for a given interrogation time), and its relatively large mass, (hence small second-order Doppler shift at a given temperature). In addition, a ${}^{202}\text{Hg}$ lamp

source can be used to optically pump the ${}^{199}\text{Hg}^+$ ground state. Experiments on Hg ions to date have been unsuccessful in cooling the ions to the ambient room temperature. This results in a relatively large second-order Doppler or time-dilation shift which may be difficult to evaluate to better than about 10^{-13} . NBS has proposed⁶ an ion frequency standard based on a ground-state hyperfine transition in ${}^{201}\text{Hg}^+$ ions stored in a Penning trap and cooled to below 1 K by laser radiation pressure. The laser cooling of the Hg ions is made difficult by the requirement for a 194 nm tunable, narrowband source and by depopulation optical pumping of the ground state. Laser cooling of ${}^9\text{Be}^+$ ions is experimentally much easier. Therefore it is useful to investigate a ${}^9\text{Be}^+$ based frequency standard in order to study the generic problems of stored ion frequency standards, even though the potential performance is not as high as for Hg ions, due to the lower transition frequency.

Experimental Method

Clouds of ${}^9\text{Be}^+$ ions are confined by the static magnetic and electric fields of a Penning trap and stored for hours (see Fig. 1). The ions are created inside the trap by electron bombardment of Be atoms evaporated from an oven located outside the trap electrodes. During a run, the oven is turned off, and the residual pressure is estimated to be less than 10^{-7} Pa. The trap is made of gold mesh endcaps and a molybdenum mesh ring electrode. The center of the trap is at one focus of an ellipsoidal mirror; the second focus is outside the vacuum system. A lens is used to collimate the fluorescence light into a photomultiplier tube. The ions are laser cooled, compressed, and pumped into the $(M_I, M_J) = (-3/2, -1/2) \rightarrow 2p^2P_{1/2}(-3/2, -3/2)$ ($\lambda = 313\text{nm}$) transition.⁷ The 313 nm light is obtained by generating the second harmonic of the output of a single mode

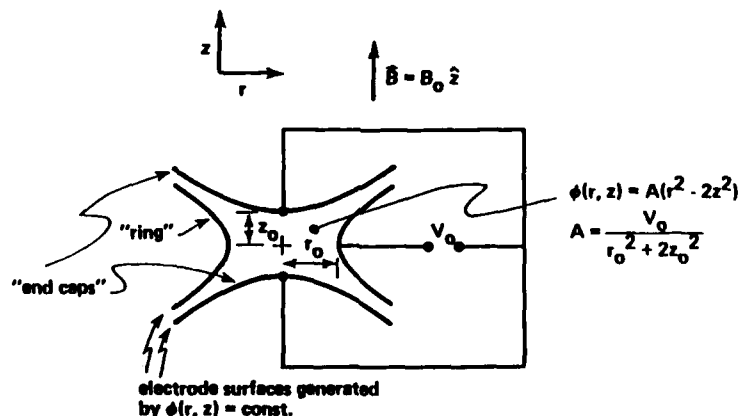


Figure 1. Electrode configuration for a Penning trap. In this experiment $B_0 = 0.8194\text{T}$, $1.64z_0 = r_0 = 0.417\text{cm}$, and $V_0 = 1\text{V}$.

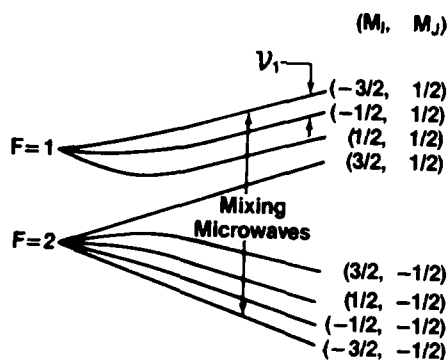


Figure 2. Hyperfine structure of the ${}^9\text{Be}^+ 2s^2S_{1/2}$ ground state as a function of magnetic field. ν_1 is the clock transition.

cw dye laser in a 90° phase-matched crystal of rubidium dihydrogen phosphate (RDP). The power is typically 20 μW . The size, density, and temperature of the ion clouds are determined by using a second probe laser.⁸ Typical clouds consist of at most several hundred ions with cloud densities of $1\text{--}2 \times 10^7$ ions/ cm^3 and cloud diameters ranging from 100 to 300 μm . Ion temperatures of around 100 mK are obtained with the laser on continuously.

At a magnetic field of about 0.8194 T, the $(-3/2, 1/2) \rightarrow (-3/2, -1/2)$ ground state hyperfine transition ν_1 (see Fig. 2), depends only quadratically on the magnetic field deviation ΔB .

$$\frac{\Delta \nu_1}{\nu_1} = -0.017 \left(\frac{\Delta B}{B} \right)^2$$

Linewidths which, to a high degree, are independent of the magnetic field homogeneity and stability are obtained by using a first order magnetic field independent transition like ν_1 as the clock transition.⁹ Microwave radiation tuned to the electron spin flip resonance ($\sim 23\,914.01$ MHz) transfers half of the ion population from the optically pumped $(-3/2, -1/2)$ state to the $(-3/2, +1/2)$ state. Some of the $(-3/2, +1/2)$ state population is transferred to the $(-1/2, +1/2)$ state by application of rf near the 303 MHz clock transition frequency. This results in a decrease in the optically pumped $(-3/2, -1/2)$ state population because of the microwave mixing and a decrease in the fluorescence light detected by the photomultiplier tube.

The time domain Ramsey method is used to probe the ν_1 clock transition. First a 0.5 s rf pulse is applied to the trapped ions. This is followed by a 19 s free precession period and then the second 0.5 s rf pulse, coherent with the first one. The laser and mixing microwaves are on for a period of 3 to 5 s during which the ${}^9\text{Be}$ ions are prepared in the $(-3/2, -1/2)$ and $(-3/2, +1/2)$ states. The laser and mixing microwaves are then turned off during the 20 s rf Ramsey interrogation period in order to avoid light and ac Zeeman shifts. After the Ramsey interrogation period, the laser and mixing microwaves are turned back on, and the signal is obtained from the photomultiplier tube count rate during the first 0.3 to 0.5 s of the laser and mixing microwaves on period. The theoretical line shape for these conditions is shown in Fig. 3. Figure 4 gives the signal obtained by averaging 10 sweeps across a 100 MHz width centered near the clock transition

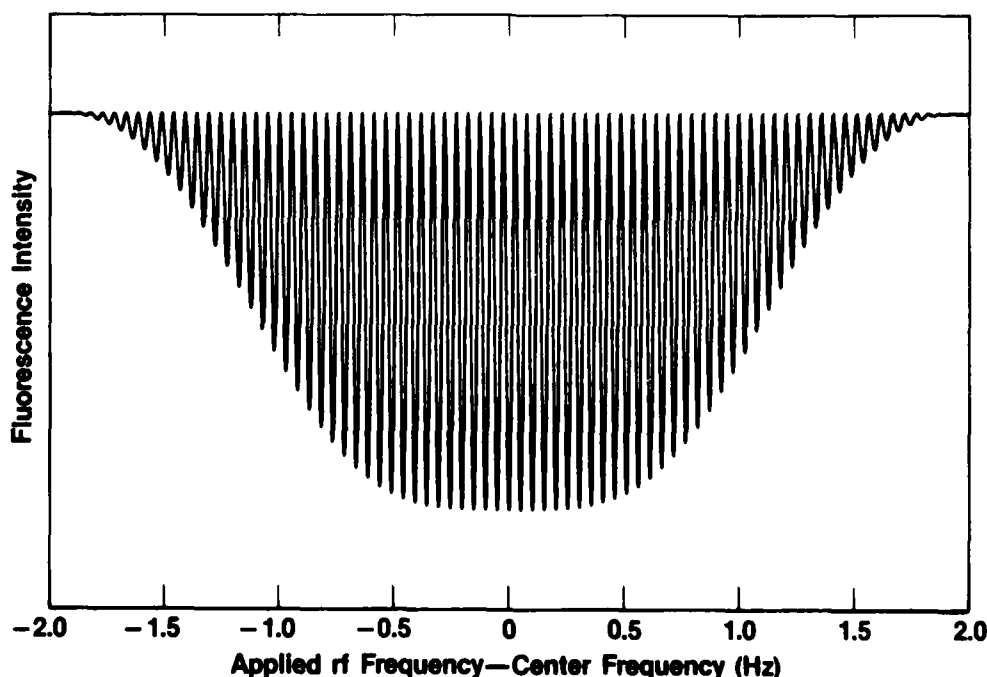


Figure 3. Theoretical resonance curve for the clock transition.

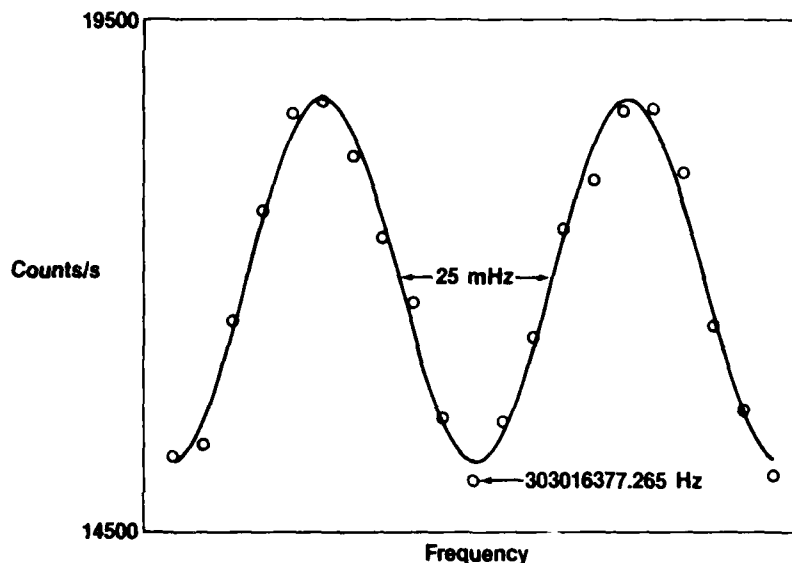


Figure 4. Ramsey signal obtained on the clock transition. The sweep width is 100 mHz and the frequency interval between points is 5 mHz. The dots are experimental; the curve is a least squares fit.

frequency (i.e., the central portion of Fig. 3). The 25 mHz linewidth gives a Q of 1.2×10^{10} on the 303 MHz clock transition.

Figure 5 is a block diagram of the system used to lock an rf oscillator to the clock transition. The 303 MHz is obtained by sum-mixing the 294 MHz frequency doubled output of a frequency synthesizer (SYNTH 1) with the output of another frequency synthesizer (SYNTH 2) near 9 MHz. A 5 MHz voltage controlled crystal oscillator (VCXO) is used as a reference for SYNTH 2. A passive hydrogen maser ($\sigma_y(\tau) = 1.5 \times 10^{-12} \tau^{-1/2}$, frequency drift $< 3 \times 10^{-16}/\text{day}$)¹⁰ is used as a reference for SYNTH 1. SYNTH 2, the counter, the laser light shutter, and the microwave and rf switches are interfaced to a computer which controls the data acquisition sequence. The computer steps SYNTH 2 by ± 13 MHz about a frequency near the 303 MHz clock transition frequency minus 294 MHz. If C_1 , C_2 , and C_3 are three successive signals, the computer calculates an error signal equal to $C_1 + C_3 - 2C_2$.^{4,5} This error signal is independent of a linear drift in the total count rate. The computer simulates an analog integrator by incrementing a register by a number proportional to the error signal after each measurement. The contents of this register are output to a digital-to-analog converter (D/A), which steers the VCXO so as to keep the mean interrogation frequency (the average of the high and low frequency half-cycles) as close as possible to the clock transition frequency. The frequency of the VCXO is compared with the frequency of the passive hydrogen maser.

Results

The Allan variance of the VCXO was measured out to 3.2×10^3 s. The VCXO is used as a reference for SYNTH 2 which provides only 9 MHz of the 303 MHz rf. By dividing the Allan variance of the VCXO by the ratio of the clock transition and SYNTH 2 frequencies (equal to 33.6), the Allan variance, $\sigma_y(\tau)$, of an oscillator stabilized by the entire 303 MHz clock transition is obtained, since frequency

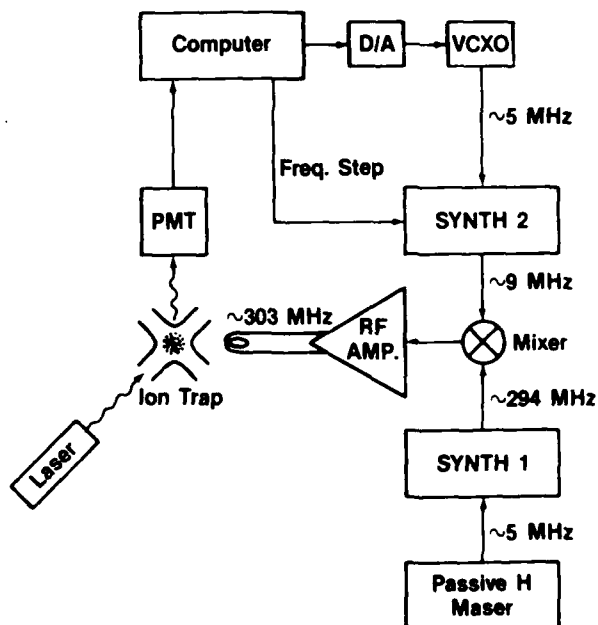


Figure 5. Block diagram of the frequency servo.

fluctuations in the frequency doubled 294 MHz output of SYNTH 1 are negligible. Figure 6 shows $\sigma_y(\tau)$ for six runs of this experiment. The attack time is between 300 and 400 s. For times greater than the attack time, $\sigma_y(\tau)$ is falling off as $2 \times 10^{-11} \tau^{-2}$. A measurement of the ν_1 clock transition frequency is obtained from each run. A weighted average frequency, $\bar{\nu}_1$, is obtained for fifteen 6400 s runs by weighting each run according to its stability $\sigma_y(\tau)$ at $\tau=1600$ s.

$$\bar{\nu}_1 = \frac{\sum (f_i / \sigma_i^2)}{\sum (1 / \sigma_i^2)}$$

$$= 303\,016\,377.265\,077 \text{ Hz}$$

The standard deviation of this average is

$$\sigma = \left[\frac{\sum (f_i - \bar{f})^2 / \sigma_i^2}{(N-1) \sum (1 / \sigma_i^2)} \right]^{1/2}$$

$$= 44 \text{ } \mu\text{Hz} \text{ } (1.5 \times 10^{-13}).$$

As noted above, the frequency of the stabilized VCXO was measured relative to the passive hydrogen maser. The uncertainty of the measured frequency of the passive hydrogen maser relative to the international definition of the second, based on the Cs hyperfine separation, is 64 μHz (2.1×10^{-13}).

Various experimental parameters were varied to test for possible systematic errors. The mixing microwave power was varied by over 20 dB. In the initial stages of this experiment, a 1×10^{-12} shift in the measured clock transition frequency was observed when the mixing microwave power was lowered by 20 dB. It was determined that this shift was

related to the small rf leakage during the repumping (laser, mixing microwaves on) part of the data cycle. By switching the SYNTH 2 frequency by ± 1 kHz during the repumping part of the data cycle, this shift was made to disappear. Other parameters which were varied in order to test for systematic frequency shifts were the laser power and the number of ions, each of which was varied by about a factor of two, and various dead times which were inserted between parts of the data cycle. In order to test for frequency offsets due, for example, to an asymmetric component of the resonance line, the frequency steps were changed to ± 0.47 Hz, so that the sides of the ninth sidelobes, rather than of the central lobe, were sampled (See fig. 3). No systematic shifts were observed at the level permitted by the signal-to-noise ratio (a few parts in 10^{10}).

Table 1 lists the estimated systematic errors for this experiment. The 3 parts in 10^7 peak-to-peak fluctuations in the magnetic field contribute a 3×10^{-14} uncertainty in the clock transition frequency. State of the art superconducting magnets have stabilities 1,000 times better than the magnet used in this experiment, which could lower this uncertainty by a factor of 10^3 . Nonlinearities and frequency drift in the VCXO could cause servo offsets at the 3×10^{-14} level. By improving the servo system (for instance, linearizing the VCXO response) this uncertainty could be substantially reduced. The temperature of the ions with the laser on was measured to be approximately 100 mK which corresponds to a second-order Doppler shift of 1.5×10^{-15} . With the laser off, some heating of the ions was observed. During the 20 s rf Ramsey period while the laser was off, the ion temperature increased to about 3 K. This results in a second-order Doppler shift of about 5 parts in 10^{14} . We believe that the heating observed in this experiment can be

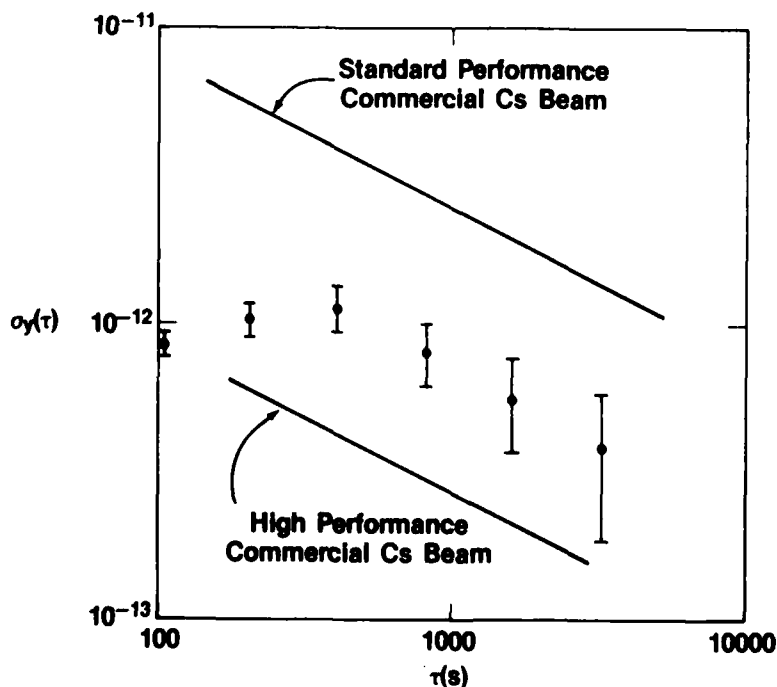


Figure 6. $\sigma_y(\tau)$ in this experiment compared with standard and high performance commercial Cs beam tubes.

TABLE 1. Estimated systematic errors

Systematic effect	Size of effect	Uncertainty
Magnetic field freq. shift	$\nu_1 = 303 \text{ MHz}$	3×10^{-14}
Servo offsets	3×10^{-14}	3×10^{-14}
2nd order Doppler	5×10^{-14}	5×10^{-14}
Pulling due to spurious signals	$< 10^{-15}$	
Stark shifts	$< 10^{-15}$	
1st order Doppler	$< 10^{-15}$	
Collisions with background gas	$< 10^{-15}$	
Background slopes	$< 10^{-15}$	
Coherence between cycles	$< 10^{-15}$	

understood and controlled in future experiments, and that second-order Doppler shifts at the 1×10^{-15} level will be obtained. A potential systematic effect for stored ion frequency standards is a light shift due to an incomplete reinitialization of the ion population during the repumping part of the data cycle¹¹. Such a shift may be present even though the light is off during the rf interrogation period, if coherence survives the repumping period. In this experiment, the computer random function generator was used to randomize the phase of the rf before each rf Ramsey period. This eliminates any systematic light shift. All other systematic effects listed in Table 1 are less than 1×10^{-15} . Unfortunately, the signal-to-noise ratio in this experiment did not permit reaching the level of the anticipated systematics of Table 1 in a reasonable amount of time.

Future Work

As stated in the introduction, the purpose of this experiment was to study the generic problems of stored ion frequency standards with cold ions. Ultimately, the same experiment could be performed on the 26 GHz field independent ground state hyperfine transition in $^{201}\text{Hg}^+$.⁶ Based on previous work with Mg^+ ions⁷, 10 MHz linewidths should be obtainable. This would give a line Q of 2.6×10^{12} . In future experiments, we also expect to load clouds with as many as 10^5 ions.¹² From the results obtained in this experiment on small clouds and a 303 MHz clock transition frequency, we anticipate that a stability of $2 \times 10^{-15} \tau$ should be obtainable with a frequency standard based on large clouds of $^{201}\text{Hg}^+$ ions.⁶ The uncertainty of the anticipated systematic effects should be comparable to those discussed in the previous section; thus an accuracy on the order of 1×10^{-15} is expected. The development of the cw, tunable, narrowband source at 194 nm required for cooling the Hg^+ ions has now been accomplished.¹³ Experiments on trapping and cooling Hg^+ ions are being initiated.

Acknowledgments

We wish to thank the Air Force Office of Scientific Research and the Office of Naval Research for their support of this work. We would also like to thank S. R. Stein for his aid in using the time measurement system and for many helpful discussions, C. Manney for his work in designing the rf probe, and J. S. Wells and R. Blatt for carefully reading the manuscript.

References

1. H. G. Dehmelt, *Advan. Atomic and Mol. Physics* **3**, 53 (1967) and **5**, 109 (1969); D. J. Wineland, W. M. Itano, and R. S. Van Dyck, Jr., *Advan. Atomic and Mol. Physics* **19**, to be published.
2. F. G. Major and G. Werth, *Phys. Rev. Lett.* **30**, 1155 (1973).
3. M. D. McGuire, R. Petsch, and G. Werth, *Phys. Rev. A* **17**, 1999 (1978).
4. M. Jardino, M. Desaintfuscien, R. Barillet, J. Viennet, P. Petit, and C. Audoin, *Proc. 34th Ann. Symp. on Freq. Control*, 1980 (Electronic Industries Assoc., 2001 Eye St., NW, Washington D.C. 20006) p. 353 and *Appl. Phys.* **24**, 167 (1981).
5. L. S. Cutler, R. P. Giffard, and M. D. McGuire, *Proc. 13th Annual Precise Time and Time Interval Applications and Planning Meeting*, Washington, D.C. 1981 (NASA Conf. Publ. 2220, NASA Scientific and Tech. Info. Branch, 1982) p. 563 and paper in these proceedings.
6. D. J. Wineland, W. M. Itano, J. C. Bergquist and F. L. Walls, *Proc. 35th Ann. Symp. on Freq. Control*, 1981 (Electronic Industries Assoc., 2001 Eye St., NW, Washington D.C. 20006) p. 602.
7. W. M. Itano and D. J. Wineland, *Phys. Rev. A* **24**, 1364 (1981).
8. J. J. Bollinger and D. J. Wineland, *Bull. Am. Phys. Soc.* **28**, 782 (1983).
9. P. Kusch and H. Taub, *Phys. Rev.* **75**, 1477 (1949).
10. Private communication with F. L. Walls, Time and Freq. Div., NBS.
11. E. I. Alekseyev, Ye. N. Bazarov, and G. I. Telegin, *Radio Eng. Electron. Phys. (USSR)* **20**, 73 (1975).
12. Only small clouds of $^9\text{Be}^+$ ions were used in this experiment because the supply of ^9Be in the oven was almost exhausted.
13. H. Hemmati, J. C. Bergquist, and W. M. Itano, *Opt. Lett.* **8**, 73 (1983).

PERFORMANCE OF COMPACT HYDROGEN MASERS

Albert Kirk

Jet Propulsion Laboratory (JPL)*
California Institute of Technology
4800 Oak Grove Dr., Pasadena, CA 91109

Abstract

Compact hydrogen masers have been built and are under continuing development at several organizations. JPL maintains a special facility which is dedicated to testing, evaluating, improving and maintaining state of the art frequency standards. This paper presents the results of a series of tests that were performed at this facility on two different types of compact hydrogen masers. One of these masers is a prototype unit developed by Hughes Research Laboratories, Malibu, CA for the United States Naval Research Laboratory. This maser uses Q-enhancement to achieve cavity oscillation and is equipped with a cavity tuning servo system. The other is a prototype passive hydrogen maser also equipped with a cavity tuning servo system and was developed by the National Bureau of Standards, Boulder, CO. The data presented includes output frequency versus certain environmental parameters, Allan variance and drift. The work on the Hughes Research maser was sponsored by the United States Naval Research Laboratory (NRL). The NBS passive maser development was sponsored by the National Aeronautics and Space Administration (NASA) and NRL.

Introduction

Two types of tests were performed on the Hughes compact maser¹ designated throughout this paper as "CHY-2".

- Environmental effects on output frequency.
- Statistical frequency stability and drift.

For the frequency stability and drift tests we included a compact passive hydrogen maser prototype², which was built by the National Bureau of Standards (NBS) and designated in this report as "NBS-3".

Four active hydrogen masers³ and two commercial cesium beam frequency standards were used as reference frequency standards. DSN-2 and DSN-3 are the designations for the JPL built (deep space network) "DSN" type masers. These are placed inside environmentally isolated chambers and serve as the primary reference standards at the test facility. The NR-5 maser is one of the newest models from the Goddard Spaceflight Center (GSFC) and was built by the Applied Physics Laboratory (APL) of the Johns Hopkins University. The SAO-14 maser is of the latest series built by the Smithsonian Astrophysical Laboratory (SAO).

The test environment was continuously monitored in order to assess its effect on the validity and accuracy of the tests. Room temperature was maintained at $23.5^\circ \pm 0.2^\circ\text{C}$. Great care was taken to assure that the masers were not locked to each other. The CHY-2 maser was susceptible to this problem.

Environmental Effects on the Output Frequency of CHY-2

A. TEMPERATURE

The maser was placed inside a test chamber and the chamber temperature was held at $+23^\circ\text{C}$ until the maser output frequency was stable. The chamber temperature was then lowered to 21°C and about 8 hours later it was increased to 29°C . The resulting shifts in output frequency were measured and we estimate the following temperature coefficients:

For a small step $\Delta T = 2^\circ\text{C}$

$$\frac{\Delta f}{f} = 3 \times 10^{-14} \pm 2 \times 10^{-14}/^\circ\text{C}$$

and for a large step $\Delta T = 8^\circ\text{C}$

$$\frac{\Delta f}{f} = 1.2 \times 10^{-14} \pm 0.5 \times 10^{-14}/^\circ\text{C}$$

It should be noted that this was a "quick look" temperature test and was not repeated, furthermore the influence of humidity changes on the output frequency was not determined but subsequent data suggests that it may be significant.

B. BAROMETRIC PRESSURE

The chamber temperature was held constant at $+25^\circ\text{C} \pm 0.3^\circ\text{C}$ while the barometric pressure was stepped $\pm 12"$ H₂O about 7 times. This test took about 30 hours and the resulting frequency changes let us estimate a barometric pressure coefficient by averaging all but the first and last half steps.

$$\frac{\Delta f}{f} = -1 \times 10^{-14} \pm 0.5 \times 10^{-14}/\text{inch Hg}$$

C. MAGNETIC FIELD

A helmholtz coil was placed around the maser so as to provide a uniform magnetic field along the vertical axis. Three separate tests were conducted. The applied magnetic field was stepped 7 times between + and - 100 mG yielding 200 mG per step data. The test was repeated two more times using ± 300 mG and ± 500 mG steps to yield data for 600 and 1000 mG steps. The maser was operated at the hydrogen flux level specified by the manufacturer, the ambient temperature was at 23°C and the Zeeman frequency was initially at 450 Hz. The following coefficients were determined.

$$\frac{\Delta f}{f} = 8.9 \times 10^{-13} \pm 4 \times 10^{-14}/\text{gauss (200 mG steps)}$$

$$\frac{\Delta f}{f} = 8.8 \times 10^{-13} \pm 3 \times 10^{-14}/\text{gauss (600 mG steps)}$$

$$\frac{\Delta f}{f} = 7.7 \times 10^{-13} \pm 2 \times 10^{-14}/\text{gauss (1000 mG steps)}$$

*The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

AD P 002451

Frequency Stability

Allan variance (CHY-2, NBS-3) frequency stability was determined by measuring the Allan variance⁴ several times against different frequency standards. The reference frequency standard was offset in frequency to yield a desired beat frequency. The zero crossing times of this beat was recorded on magnetic tape for offline computer processing⁵. For most tests, the 5 MHz standard output frequency of the masers, under test, was multiplied to 100 MHz and the measurements were made at this frequency. The plots shown in this report are typical of all the data that was collected during this test.

Figure 1 shows the results of CHY-2 versus the JPL DSN-3 maser. This data was taken during the period between 12-22-82 and 1-7-83. As is shown later in this report the CHY-2 had two distinct changes in output frequency during the above period. These are supposed to be due to the malfunction in the CHY-2 cavity synthesizer (a problem that had occurred earlier during initial verification tests) which may be susceptible to changes in humidity.

Figure 2 uses the same set of raw data but covers a shorter time interval which excluded the periods during which the large frequency change took place. Notice the difference at $\tau > 10^4$ s.

Figure 3 is a typical Allan variance plot of the NBS-3 maser versus the JPL DSN-3 maser.

Figure 4 covers the period which excludes the effects of distinct frequency changes that occurred on NBS-3 (see Fig. 11).

Figure 5 is a plot of the JPL reference masers DSN-3 vs. DSN-2. The noise is sufficiently low to not appreciably affect the results obtained for the compact test masers, except perhaps at tau values $> 10^4$ seconds.

Long Term Drift

A primary objective was to determine the long term behavior of the output frequency vs time. The allotted test period was 45 days. Preliminary measurements indicated that a low drift rate was to be expected for both compact masers.

During the last few years it was determined that the drift of DSN-2 and DSN-3 is due to cavity aging.³ Repeatability of cavity tuning was shown to be $\pm 2 \times 10^{-14}$ for these two masers. A manual tuning technique allows completion of tuning both of these masers in less than 6 hours. We therefore proceeded in the following manner:

- Tune all reference masers and measure the frequency offset that exists between all masers at the beginning of the test. (Day 0)
- Continuously auto tune the NR-5 maser during the entire test period. (This removes nearly all of its drift.)
- Repeat the tuning procedure for DSN-3 and DSN-2 at the end of the test and again measure the frequency offset that exists between all masers.*

*We make the basic assumption that the atomic line frequency of DSN-2 and DSN-3 did not change.

After the 43 day test period we determined the following frequency changes between maser pairs:

$$\begin{aligned} \text{DSN2} - \text{DSN3} &= 0 \pm 2 \times 10^{-14} \\ \text{DSN2} - \text{NR5} &= +2.0 \times 10^{-14} \pm 5 \times 10^{-15} \\ \text{DSN2} - \text{CHY2} &= -6.0 \times 10^{-14} \pm 1.5 \times 10^{-14} \\ \text{DSN2} - \text{NBS3} &= 0 \pm 2 \times 10^{-14} \\ \text{DSN3} - \text{NR5} &= +2.5 \times 10^{-14} \pm 5 \times 10^{-15} \\ \text{DSN3} - \text{CHY2} &= -5.5 \times 10^{-14} \pm 1.5 \times 10^{-14} \\ \text{DSN3} - \text{NBS3} &= +1.5 \times 10^{-14} \pm 2 \times 10^{-14} \end{aligned}$$

From these measurements we estimate the following frequency shifts for each maser over the 43 day test period:

$$\begin{aligned} \text{CHY-2} &= -6 \times 10^{-14} \pm 3.5 \times 10^{-14} \\ \text{NBS-3} &= +1 \times 10^{-14} \pm 4.0 \times 10^{-14} \\ \text{NR-5} &= +2.5 \times 10^{-14} \pm 2.5 \times 10^{-14} \end{aligned}$$

In addition to the frequency offset measurements against DSN2 and DSN3 we also recorded continuous phase data between the test masers and NR-5 which was continuously autotuning with a maser as the reference. The two commercial cesium units were too noisy for the required resolution during this test.

Figure 6 is the phase plot of the CHY-2 against the NR-5 for a 38 day period. The effect of two distinct frequency shifts can be clearly seen.

Figure 7 is the fractional frequency plot derived from the phase data of Figure 6. Note the increase from day 5 to 8, the turnaround followed by a decrease of about 1×10^{-13} and settling at a new frequency by day 11. This corresponds very closely with a significant change in humidity (approximately 25%). See Figure 16 which shows a plot of humidity during the 38 day test period. A similar shift in frequency, corresponding to humidity occurred near day 25. It is not clear why the frequency did not return to its nominal value on these two occasions or why it changed by such a relatively large amount. Measurements against the other reference masers verified these shifts on CHY-2. We observed this type of frequency shift before, on earlier tests, and found problems with the cavity control synthesizer and a cable on the CHY2 maser.

We estimate the total frequency shift of CHY-2 against NR-5 using Figure 7 to be -8.5×10^{-14} . This is an excellent comparison agreement with the measurements also made against the reference masers DSN-2 and DSN-3.

Figures 8, 9, 10 show the 10 to 26 day portion of phase, frequency and Allan variance between CHY-2 and NR-5. The peak to peak noise of approximately 2 nS masks any long term drift that may be present. The residuals suggest environmental effects and appear to correlate with humidity. Of course CHY-2 could be drifting at about the same rate and in the same direction as the NR-5. As measured directly against the offset masers the upper bound of drift for the NR-5 is estimated to be 1×10^{-15} /day when the auto-tuner is being utilized (in NR-5).

The main question that needs to be addressed in regards to CHY-2 is whether the behavior shown in

Figure 7 is the intrinsic behavior of frequency versus time for this type of maser design or rather the result of a curable malfunction which may be induced by environmental changes.

If we remove the two distinct frequency changes of 1×10^{-13} shown in Figure 7 we can estimate a drift for CHY-2:

$$\frac{\Delta f}{f} = \frac{-6 \times 10^{-14} + 1.0 \times 10^{-13}}{43 \text{ days}}$$

$$= 1 \times 10^{-15} / \text{day} \pm 8 \times 10^{-16}$$

Figures 11, 12 show the phase and frequency data of the NBS-3 maser against NR-5. Several distinct frequency changes occurred during the test period. The net frequency change over the entire period appears to be nearly zero and confirmed by direct frequency measurements against DSN-2 and DSN-3. The environmental data shows that the frequency perturbations correlate well with humidity. Note the nearly seven day cycles in the residuals.

Again we can speculate as to whether this behavior is inherent with this design or whether an abnormal malfunction, perhaps environmentally induced, is at play.

Figures 13, 14, 15 cover the 10 to 26 day period during which the maser frequency perturbations were of shorter durations. Note again the seven day periodicity often seen on frequency standards. The Allan variance shown on Figure 15 extends that of Figure 4 without the limitation of the DSN-3 drift.

The overall frequency change over the test period yields an estimate of the drift rate for the NBS-3 maser.

$$\frac{\Delta f}{f} = +2 \times 10^{-16} / \text{day} \pm 9 \times 10^{-16}$$

Correlation Functions

In order to gain more information about the frequency-humidity relationship auto and cross-correlation was performed on the humidity and test maser output frequency data.

Figure 16 is a plot of dew point/frost point (humidity) during the test period.

Figure 17 and 18 are plots of the cross correlation coefficients of humidity and frequency for the CHY2-NR5 and NBS3-NR5 respectively. The CHY2-NR5 frequency pair yields a coefficient of 0.83 with humidity with virtually no lag whereas the NBS3-NR5 frequency pair yields a coefficient of 0.89 with humidity at a lag of about 1 day.

Significant Events

Both test masers failed and dictated the end of test on the 43rd day. CHY-2 vacion pump failure caused out of lock conditions and the NBS-3 cavity servo started to go out of range.

Day 26

Brief loss of lock on CHY-2 due to vacion glitch. Maser was restored and phase data corrected. The impact was minor.

Day 38 to Day 44

Vacion pump failures on CHY-2. Maser restored and end of test frequency measurements made. Frequency repeatability was very good after each pump failure and subsequent restoration.

Prior to environmental testing on CHY-2 several synthesizer malfunctions occurred. There was also a problem with a coaxial cable between the source oscillator and the source. Additionally some effort was required to provide sufficient shielding for the CHY-2 receiver and physics unit to ensure that no lock up to a reference maser can occur. Evidence that this could happen was discovered when a 1.42405751 GHz signal from another maser was radiated in the test facility.

Summary

Long term frequency stability measurements proved to be as difficult to make as anticipated. It is clear that much longer time periods are required to reduce the relatively large measurement uncertainties associated with this test. The unexpected frequency shifts that occurred make it nearly impossible to derive an accurate measurement of drift over a 43 day period. It appears that the cavity frequency control servo mechanism employed on the CHY-2 and NBS-3 type masers are capable of considerably reducing the frequency pulling effect of cavity aging.

It is also apparent that, for a maser with intrinsically good long term stability (low drift), environmental parameters such as those measured for CHY-2 assume a dominant role and need to be dealt with. This involves the environment itself as well as the maser's sensitivity to it.

A drift of 5×10^{-16} per day for example accumulates a time error of about 20 nS in a 30 day period. It is easy to see that for precise time keeping purposes severe requirements are imposed on the environment, given the environmental coefficients measured on CHY-2.

The data suggests the following conclusions:

Over the short term ($T < 7$ days) environmental effects and short term noise dominate and are characterized by the Allan variance measurement as white frequency noise ($\tau^{-1/2}$). Random walk frequency type noise was not evident over the given measurement period. Neither the 38 day phase data (absence of second order terms) nor the 38 day Allan variance data (absence of τ' slope) indicated significant long term drift.

Correlation with humidity is high.

Acknowledgements

The author would like to acknowledge the contribution of Roland Taylor who performed many tests and solved several difficult problems. Additionally

Phil Clements, Charles Greenhall, Paul Kuhnle and Richard Sydnor who contributed in the data analysis, plots and writing of this report.

References

1. H.T.M. Wang, "An Oscillating Compact Hydrogen Maser," Proc. 34th Annual Symposium on Frequency Control, pp. 364-369, 1980.
2. F.L. Walls, D.A. Howe, "Timekeeping Potential Using Passive Hydrogen Masers," J. de Physique, Vol. 42, C8-151 to 158, 1981.
3. A. Kirk, P. Kuhnle and R.L. Sydnor, "Evaluation of Modern Hydrogen Masers," Proc. 14th Annual Precise Time and Time Interval Applications and Planning Meeting, pp. 359-392, 1982.
4. J.A. Barnes, A.R. Chi, L.S. Cutler, D.J. Healey, D.B. Leeson, T.E. McGunigal, J.A. Mullen, W.L. Smith, R.L. Sydnor, R.F.C. Vessot and G.M.R. Winkler, "Characterization of Frequency Stability," IEEE Trans. Instrum. Meas., Vol. IM-20, No. 2, pp. 105-120, 1971.
5. C.A. Greenhall, "The Fundamental Structure Function of Oscillator Noise Models," Proc. 14th Annual Precise Time and Time Interval Applications and Planning Meeting, pp. 281-294, 1982.

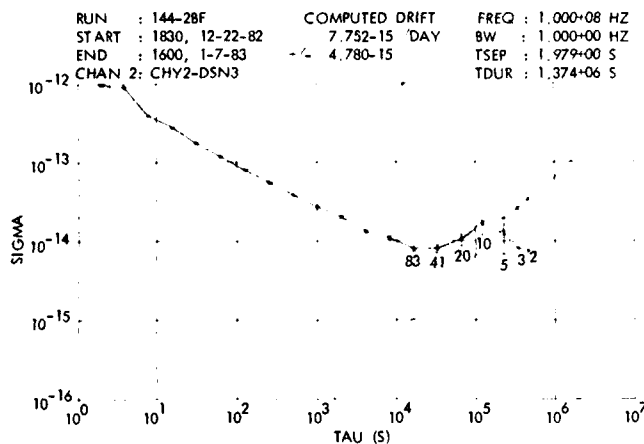


Figure 1. Allan Variance CHY2 vs. DSN3

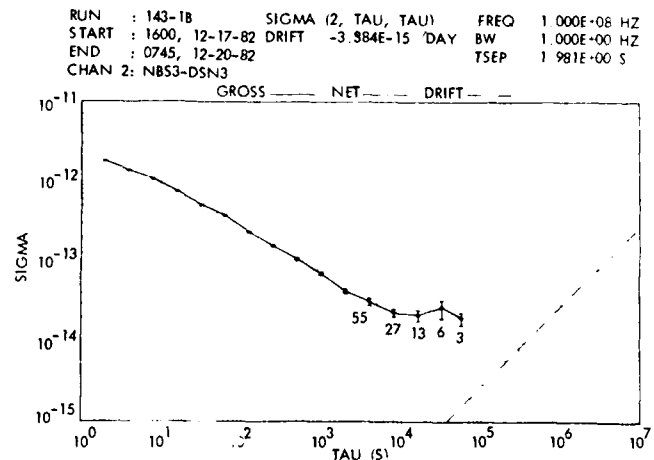


Figure 3. Allan Variance NBS3 vs. DSN3

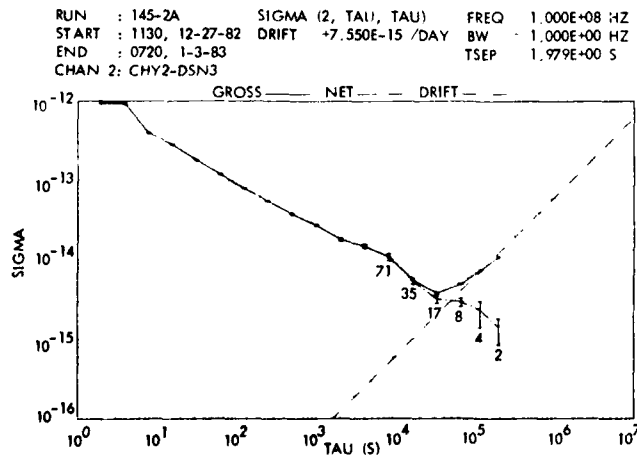


Figure 2. Allan Variance CHY2 vs. DSN3

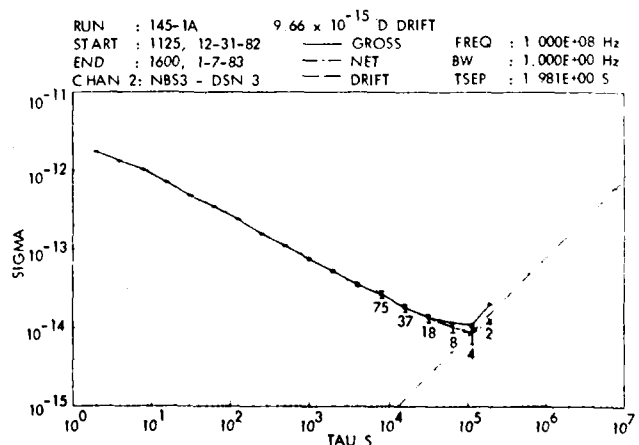


Figure 4. Allan Variance NBS3 vs. DSN3

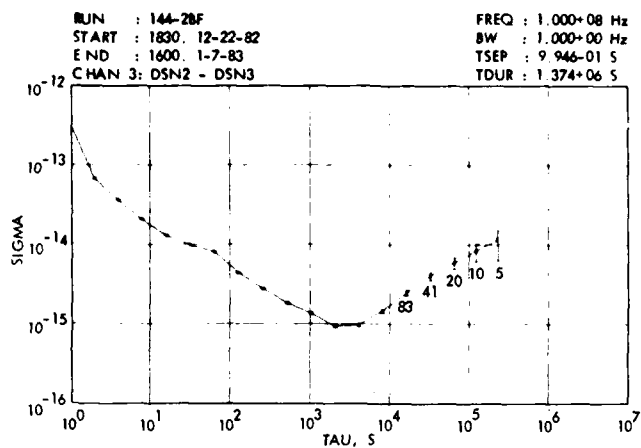


Figure 5. Allan Variance DSN2 vs. DSN3

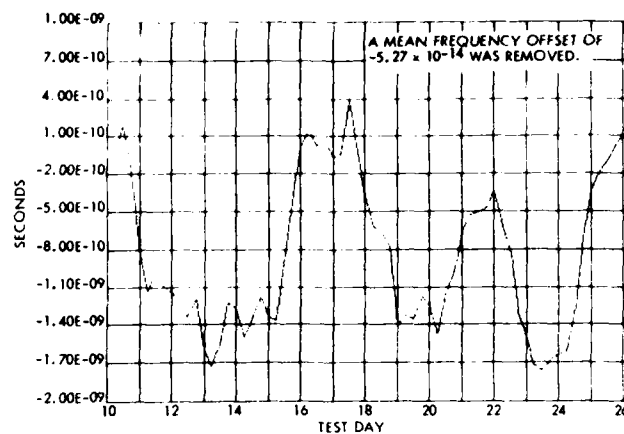


Figure 8. Time Residuals CHY2 vs. NR5

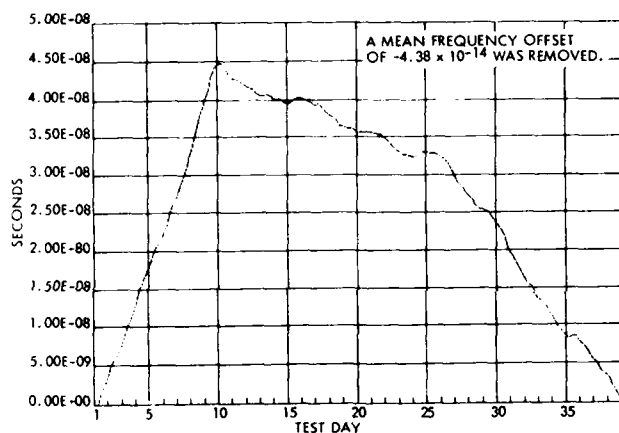


Figure 6. Time Residuals CHY2 vs. NR5

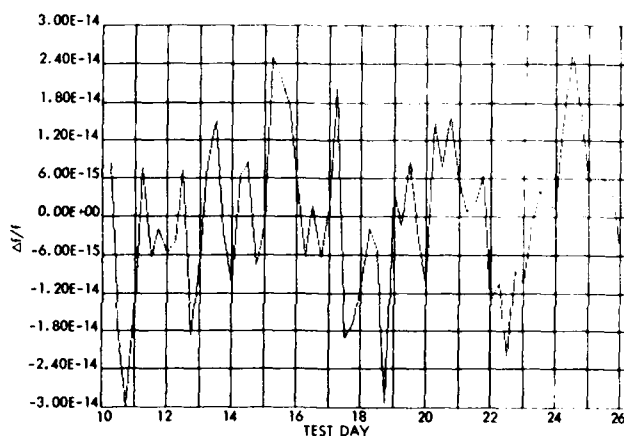


Figure 9. Frequency Residuals CHY2 vs. NR5

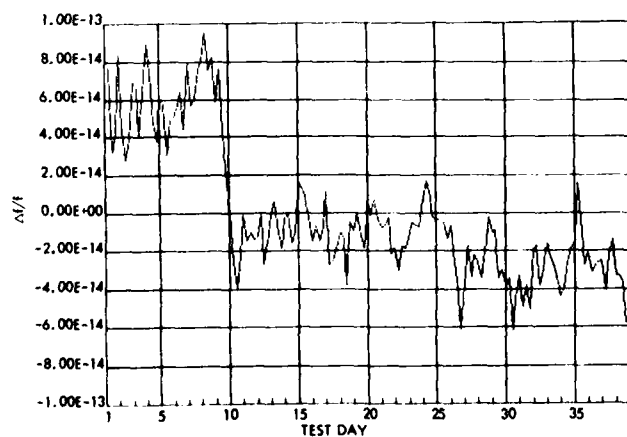


Figure 7. Frequency Residuals CHY2 vs. NR5

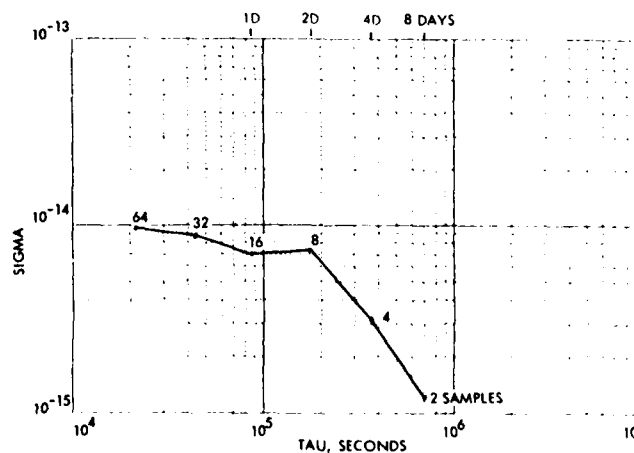


Figure 10. Allan Variance CHY2 vs. NR5
Test Day 10 thru 26

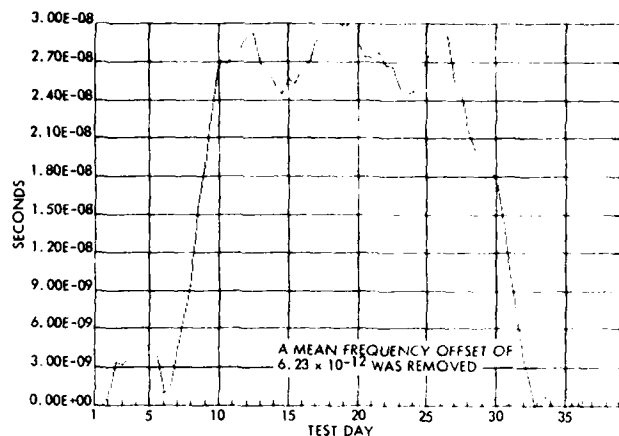


Figure 11. Time Residuals NBS3 vs. NR5

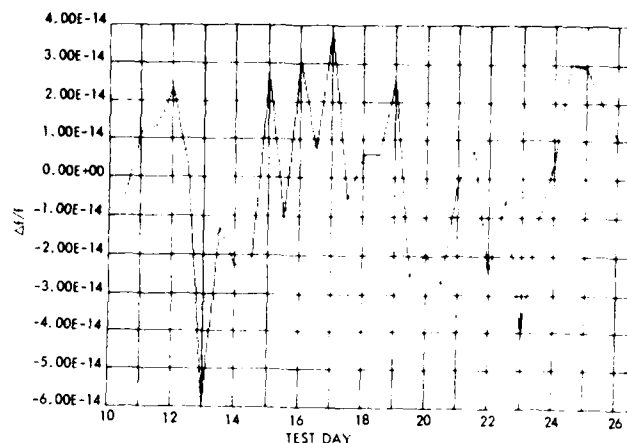


Figure 14. Frequency Residuals NBS3 vs. NR5

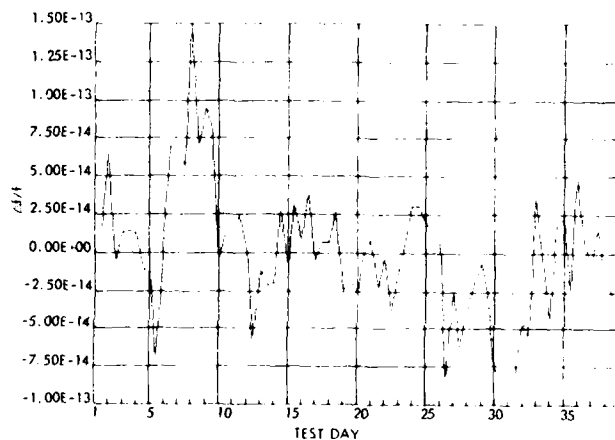


Figure 12. Frequency Residuals NBS3 vs. NR5

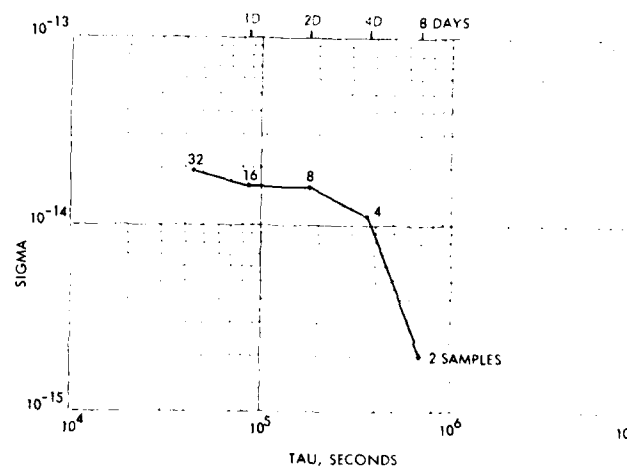


Figure 15. Allan Variance NBS3 vs. NR5
Test Day 10 thru 26

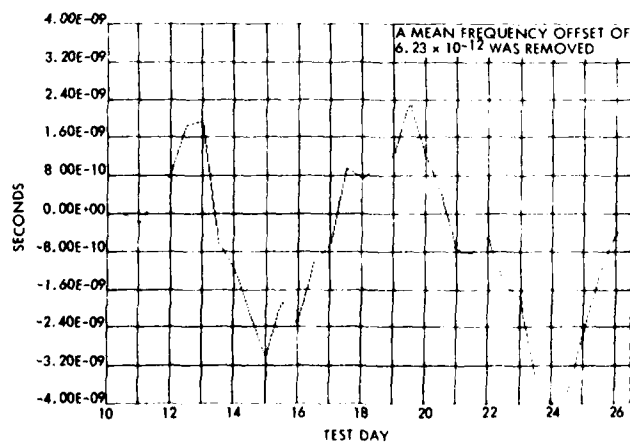


Figure 13. Time Residuals NBS3 vs. NR5

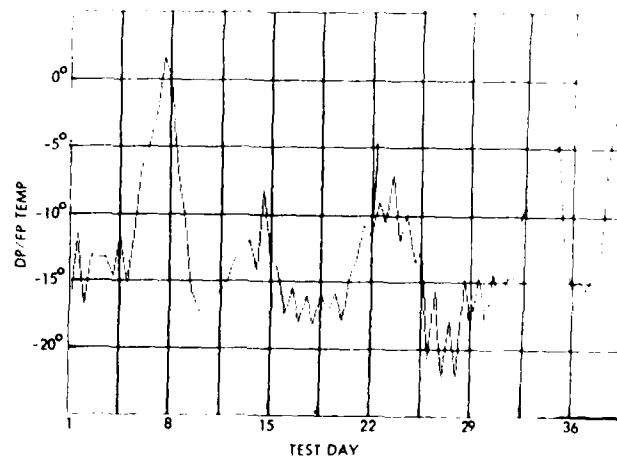


Figure 16. Ambient Humidity

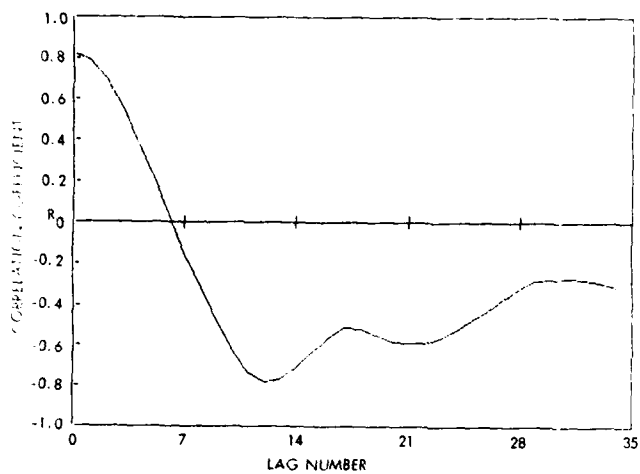


Figure 17. Cross Correlation CHY2 NR5 vs. Humidity

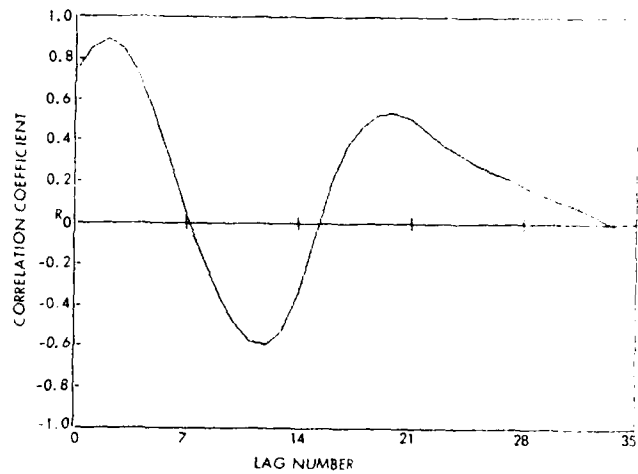


Figure 18. Cross Correlation NBS3-NR5 vs. Humidity



COLD HYDROGEN MASER RESEARCH AT SAO AND RELATED DEVELOPMENTS

R.F.C. Vessot, E.M. Mattison, and E. Imbier

Smithsonian Astrophysical Observatory and
Harvard College Observatory, Cambridge, Mass. 02138

ABSTRACT

The cryogenically-cooled maser provides a means for measuring frequency shifts and relaxation properties of storage wall coating materials that can be frozen in place from substances normally in the gaseous phase. We report initial measurements for FEP-120 Teflon between 77K and 48K and compare them with the 372K to 77K data of de Saintfiscien.¹ Some of the design features of the maser are described. The cryomaser's hydrogen dissociator is located entirely within the vacuum enclosure and operates at 77K. The in-vacuum dissociator and other developments that have been adapted and tested in a room temperature maser, and that are applicable for both space and terrestrial use, are also discussed.

INTRODUCTION

A program for studying the operation of atomic hydrogen masers at low temperatures has been under way at the Smithsonian Astrophysical Observatory since 1977. The aim of this program is to advance both the science and the technology of hydrogen masers. The work has been supported by the Office of Naval Research, the Smithsonian Institution, the National Aeronautics and Space Administration (NASA), the Jet Propulsion Laboratory, and Marshall Space Flight Center.

The principal reason for the outstanding stability of the hydrogen maser is the fact that hydrogen atoms can be confined in a storage vessel for extended periods (on the order of a second) with very little effect on the phase of their oscillating magnetic dipole moments. This moment results from the magnetic interaction between the proton and electron spins in the hydrogen ground state; the energy separation produced by this interaction provides the hydrogen maser signal, the familiar 21 cm line of radio astronomy.

Atoms stored in a hydrogen maser make many collisions with the wall surface of the storage volume. In conventional masers this surface is usually coated with Teflon, a long chain fluorocarbon. The interaction during collisions has not been extensively studied, and we are still using the same materials and techniques of 20 years ago. At present the theoretically available stability of the maser has been achieved for averaging periods between one second and one hour; in this regime the stability is limited by thermal noise, thermal movement of the atoms in the storage volume, and the rate of wall collision relaxation. This last process is what we want to investigate at low temperatures where electronic noise, kT , is reduced, and where atoms move more slowly and collide with much less energy than at room temperature.

By injecting substances that are gaseous at room temperature into the storage volume of a cryogenically cooled maser and allowing them to freeze on its walls, we can repeatedly coat the walls with different materials in a well-controlled manner. A measurement of the average advance or retardation of the phase of

the hydrogen atom's oscillating magnetic dipole (per collision) is obtained by measuring the frequency shift of the maser signal and knowing the kinetics of the atoms in the bulb. The average phase decorrelation is determined from the oscillation line Q , defined as

$$Q = \frac{\omega}{2\gamma_{2T}}$$

where $\omega = 2\pi f$ is the maser frequency, and γ_{2T} is the relaxation rate of the phase from all causes, including exit of the atoms from the bulb. The line Q is analogous to the quality factor of any conventional oscillator and appears in the following expression for the fundamental limitation of frequency stability:

$$\sigma_{fy}(\tau) = \frac{1}{Q} \left[\frac{kT}{2P_b \tau} \right]^{\frac{1}{2}}$$

This is the one-sigma expectation of the fractional frequency variation measured in the time interval τ . Here kT is the thermal noise power per unit bandwidth (k is Boltzmann's constant, T is absolute temperature) and P_b is the power generated by the oscillator, in this case the state-selected hydrogen atoms that have been focussed into the storage volume.

Additive white phase noise also contributes to the frequency instability; the variance of the fractional frequency variations from this source is given by

$$\sigma_{ay}(\tau) = \frac{1}{2\pi f \tau} \left[\frac{FkTB}{P_o} \right]^{\frac{1}{2}}$$

where F is the receiver noise figure, B is the overall system bandwidth in Hz, and P_o is the power delivered to the receiver.

There are several reasons to expect increased hydrogen maser stability at low temperatures.

1. kT is smaller. This also affects the signal-to-noise ratio of equipment that receives the maser signal.
2. P_b can be made larger. The limit on P_b is from hydrogen-hydrogen spin exchange collisions in the storage volume. The cross section for such collision diminishes rapidly with temperature making it possible to have a much greater density for a given level of spin exchange quenching.
3. The lower speed of the atoms reduces the wall collision rate and, all other things being equal, correspondingly increases the storage time.
4. Many properties of materials are much better behaved at low temperatures and excellent magnetic field control can be accomplished using superconducting shields.

OBJECTIVES OF THE COLD MASER RESEARCH PROGRAM

The principal objective of our present efforts is to determine the behavior of the various types of walls as a function of temperature and to develop an understanding of the fundamental processes governing the collision interactions.

AD P 002452

The possible advantages of operating a hydrogen maser at temperatures well below room temperature have been speculated upon for many years, but until recently the prospects for successfully operating such a device at temperatures significantly below 77K seemed marginal. Kleppner and his co-workers² demonstrated that atomic hydrogen could be contained at 4K in a vessel whose walls were coated with frozen molecular hydrogen. This gave considerable impetus to investigate the behavior of the hyperfine resonance at temperatures below 77K, which was the temperature limit explored by M. de Saintfiscien and others¹ in their research into wall relaxation and wall coatings.

In 1977 work began at SAO on an initial version of a liquid helium-cooled cryostat that contained a TE₁₁₁-mode split storage volume resonator with fluorinated ethylene propylene copolymer (FEP) Teflon coated walls. At first, oscillation could only be sustained down to temperatures of about 55K. Why it would not oscillate below this temperature was open to many questions -- were the walls of the storage volume contaminated because of the rather crude vacuum system we employed, or was there a more fundamental limit imposed by the nature of the atomic hydrogen in its collision with the FEP surfaces? To answer these questions we decided to try a different approach with the cold maser. This was to introduce carbon tetrafluoride (CF₄) gas through the hydrogen source structure and freeze a coating of this molecule on the inside surfaces of the cavity storage volume. This technique provided a fluorine-bonded-to-carbon surface comparable to the long chain fluoride molecules of Teflon. The scheme worked, and greatly to our delight, we were able to maintain oscillation down to temperatures of 26K.

This work was reported³ at the 33rd Annual Symposium on Frequency Control and we continued theoretical studies to determine what advantages would accrue if we could operate at or below 26K.⁴ From theory we know that the stability would be well below 1 part in 10¹⁶ at 1000 seconds provided no new instability problem arose. However, we know that the large temperature-dependence of the wall shift is likely to cause difficult requirements for temperature stabilization in the cold maser.

MECHANICAL CONFIGURATION OF THE COLD MASER

The design of the maser cryostat required considerable effort and imagination to anticipate as many uses as possible to which it could be put in a cold maser testing program. First, it had to be structurally reliable. Second, it had to have a reasonably large volume in order to house a variety of possible maser cavity and magnetic shield configurations. We decided that if we had uncertainties in these dimensions it would be better to have a larger, rather than smaller, volume. As an upper limit, we decided that it should be able to house the full size TE₀₁₁-mode cavity and inner magnetic shield design such as was used in the lightweight space-probe maser.⁵

The question of the range of operating temperatures was also an issue. Our previous success in obtaining oscillation at temperatures as low as 26K suggested that we should cover the range from 77K to somewhere near 4.2K with provision for stabilization at any point between these temperatures. This was done by mounting the maser cavity on a tubular support that contains a constant-loss, metered flow of liquid helium that cools the cavity and offsets heat leakage into the system. By metering a slightly greater flow of helium than required to offset the heat leak, an electrical resistance heater on the cavity support can

be used in a servo-controlled system to hold the cavity temperature at the desired level.

For temperatures below the liquid helium boiling point, we included provisions to change our method of cooling and can connect the cavity directly to the helium bath. By pumping on the helium gas, we should be able to drive the temperature well below 4K.

Additionally, we wanted to have a reasonably large capacity for helium so that disturbances from filling could be avoided for up to five days of steady operation. Since the projected cavity and shielding structures that would be cooled tended to be quite massive and the apparatus itself would be rather large, we decided to use a liquid nitrogen guard region to surround the liquid helium container and the sample region. This reduces the cost of cryogenics used for initial cooling and allows more flexibility of operation. It also eliminates the need for multi-layer reflective insulation (superinsulation) and the ensuing outgassing problems that could lead to surface contamination in the atomic hydrogen storage volume of the maser.

The final design of the SAO cold hydrogen maser cryostat is shown in Figure 1. The fact that the liquid helium and nitrogen tanks are above the region to be cooled is dictated by the need for a gravity feed of cryogen to the cavity region.

The annular 30 liter liquid nitrogen tank is thermally connected to an equipment mounting plate of copper beneath it and to a heavy copper shroud surrounding the working volume, which is roughly 22 inches (56 cm) in diameter x 22 inches (56 cm) deep. Inside it is the 30 liter liquid helium tank, which has a flat base plate for mounting equipment to operate near 4K. The cavity resonator is suspended from a tubular sample holder that is thermally isolated from the liquid helium reservoir by being attached to a point near the top of the dewar, allowing it to operate at temperatures substantially above 4K. Liquid helium flows to the sample holder through a needle valve that controls its flow rate. The magnetic shield assembly is suspended by thermally conducting rods from the nitrogen tank and is kept near 77K to prevent the excessive loss of magnetic shielding owing to the reduction of magnetic permeability at very low temperatures.

All the components are mounted on the 26 inch (66 cm) diameter top plate of the cryostat which is vacuum sealed by an O-ring to the vacuum enclosure. Access is gained to the assembly by lifting the plate out of the enclosure using a crane as shown in Figure 2.

To obtain reasonably fast pumping, as well as good enough cleanliness to permit outgassing the system at the at the 10⁻⁸ torr level, we use a turbo-molecular pumping system with 450 liter/second throughput. This is intended to pump the system when it is at or above 77K. When liquid helium is used, the 4 inch diameter gate valves are closed and the system cryopumps, so that all gases but helium condense on the helium tank.

The cavity resonator used in the first tests was a 15 cm diameter x 18 cm long TE₁₁₁-mode r.f. cavity⁶ (instead of the usual 28 cm diameter x 28 cm diameter long TE₁₁₁-mode cavity). This cavity is equipped with a Teflon septum to separate the two regions of r.f. magnetic field that are 180 degrees out of phase. The remaining interior surfaces of the cavity are coated with FEP-120 (fluorinated ethylene propylene copolymer). The entry collimator is split by the sep-

tum, each side leading to one-half the cavity.

The gas handling system for the wall coating gases utilizes a ballast volume, metering valves, and pressure- and flow-measuring equipment. A mechanical scavenging pump allows control of the gas ballast tank pressure, as well as purging and cleaning of the system. The gas is admitted through a valve and led to two jets located on the downstream side of the hexapole magnet, each aimed into one of the input collimators of the TE₁₁₁ cavity's storage volume.

THE COLD HYDROGEN DISSOCIATOR

Hydrogen masers require a source of hydrogen atoms in the form of a directed beam. In conventional masers operating at room temperature the source of atoms is nearly in thermal equilibrium with the storage region. In the cold masers, however, a conventional atomic hydrogen source would be much hotter than the storage volume. This can cause at least two types of problems. First, there is radiative heat transfer from the source to the cold cavity and second, there is uncertainty in the method of thermal equilibration of the hot, fast-moving hydrogen atoms after they enter the cold cavity and the possibility of their dislodging the wall coating atoms frozen to the surface. A beam source cooled at least to liquid nitrogen temperature would substantially help to diminish these problems.

We have designed a cryogenically-cooled hydrogen maser using an r.f. plasma dissociator operating at liquid nitrogen temperatures (77K) in conjunction with a state selector magnet whose dimensions are suitable for slow atoms. The focusing characteristics for a hexapole state selector magnet with maximum fields at the pole tips, H_m , provide a maximum acceptable angle θ_m for atoms at the most probable velocity in the beam given by

$$\theta_m = \left[\frac{2\mu_0 H_m}{3kT} \right]^{1/2}$$

where μ_0 is the Bohr magneton.

Since the acceptance solid angle is proportional to θ_m^2 , the magnet is approximately 4 times more effective at 77K than at 300K. Furthermore -- although this is not a very important consideration -- the magnet's length is reduced by about one-half.

By thermally isolating the r.f. circuitry from the dissociator glassware, only dielectric losses in the glass and the r.f. energy coupled to the plasma will result in the boil-off of liquid nitrogen. We estimate that this is about one watt and thus anticipate a loss rate of approximately .022 litres per hour.

The design of the dissociator is particularly simple and was later used in the four experimental- and advanced-development models of passive hydrogen masers, developed and tested by SAO for the U.S. Naval Research Laboratory.⁷ As with the other systems, these operated from an r.f. oscillator and power amplifier at 80-100 MHz, and are capable of delivering about 5 watts to the dissociator. The load impedance presented by the dissociator is highly variable because the hydrogen presents very different electrical characteristics to the generator when it is ionized in a plasma state than when it is a neutral gas.

The glass of the dissociator bulb is thermally connected to the liquid nitrogen reservoir and is enclosed in the vacuum system. A hexapole state selector magnet and beam stopping disc with dimensions

suitable for the cryogenically cooled maser are attached to the cold structure.

Figure 3 shows the device as it is built into the cryogenic maser. The liquid nitrogen cooled attachment ring is bolted to the cooled copper shroud as shown in Figure 1.

The very low power dissipated by the glass and the plasma within it has led us to try this type of design in non-cryogenically cooled masers. Here the heat from the glass is conducted to the vacuum envelope, and the excitation coil, made of heavy copper, is also thermally connected to the vacuum envelope.

This design, where the dissociator is wholly within the vacuum envelope, eliminates the need for elastomer seals and is particularly attractive for use in spaceborne masers where the thermal control of the dissociator is best done by conductive means rather than by a recirculated air flow system as used in the 1976 Redshift space maser.⁵

RESULTS OF WALL COATING MEASUREMENTS AS OF MAY 1983

The two measurable wall properties are:

1. The wall collision frequency shift. This is expressed in terms of the average phase shift per collision.
2. The wall relaxation rate. This is expressed in terms of the probability of loss of phase per collision.

Our present efforts have been to measure the wall frequency shift. The temperature dependence of the phase shift per collision gives some important clues about the nature of the interaction between impinging hydrogen atom and the wall surface. The perturbation results from polarization of the hydrogen's electron wave function owing to a dipole-dipole attractive force as the atom approaches the wall and a strong, repulsive interaction at short ranges arising from Pauli exclusion exchange forces. Two types of frequency shifting effects are encountered: 1) The polarization of the hydrogen atom leads to negative frequency shifts because of the diminished interaction of the displaced electron cloud with the proton, and 2) Pauli forces lead to positive frequency shifts resulting from compression of the electron cloud about the hydrogen atom and stronger interaction.

These two can offset and cancel each other. The temperature at which this occurs depends on the surface properties⁸⁻¹¹. At low temperatures the negative shift dominates and can be expressed as an exponential function of the temperature, $e^{E/kT}$. Here E/k can be considered an energy expressed in degrees Kelvin.

Our data are shown in Figure 4, where we plot the logarithm of the average phase shift per collision against inverse temperature. From 77K to approximately 52K the data follow the relationship:

$$\Delta\theta = 6.17 \times 10^{-6} e^{215/T}$$

As had occurred in the earlier experiments,³ during the cooling and measuring process the maser stopped oscillating at about 50K. Whether this was due to contamination or was a consequence of the nature of the FEP Teflon surfaces is still not clear. At this point we admitted gaseous CF₄ into the system, raising the pressure to 10⁻⁴ Torr for 30 seconds, and waited for the surplus to pump away. The next day we took data, then allowed the system to warm up to about 60K and took more data as the system cooled again. These data are shown as open circles in Figure 4 and were

obtained at temperatures down to 48K. These data have a different relationship with temperature than that of the bare FEP Teflon surface:

$$\Delta\theta = 11.6 \times 10^{-6} \text{ } ^\circ\text{C}/\text{T}$$

From this run we observe that the re-cooled data do not fit the first data point obtained after injecting CF_4 . It is possible that the surfaces were not heavily coated with CF_4 and that reevaporation (or resublimation) occurred when we went to 60K, leaving some of the Teflon surface exposed.

During this run we suffered a blockage in the helium flow and were unable to cool lower than 48K. However, in our earlier cryostat we were able to get oscillation at temperatures as low as 26K, but were unable to run at low enough magnetic fields to make accurate wall shift measurements.

The data reported in Figure 4 agree well with those of de Saintfuscien.¹ Figure 3 shows his data and ours plotted on the same graph. Here we see the effects of phase changes in the Teflon that occur near 300K and 200K. A summary of the properties of FEP Teflon as they would apply for a maser with a 7 inch (17.8 cm) diameter bulb is given in Table 1. We show the wallshift, its temperature coefficient, and the quantity $(df/f)/(dT/T)$, which measures the fractional temperature sensitivity. In general, at low temperatures the fractional temperature can be controlled at least as well as at higher temperatures. Nevertheless, we see that to realize frequency stability of 1×10^{-16} will require temperature control at the $5 \times 10^{-6}\text{K}$ level at 50K, not a simple matter.

CONCLUSIONS

It is clear that we are looking for a surface material that does not alter the chemical state of the impinging hydrogen atom. Beyond this, we require as low an interaction potential with hydrogen as possible so that the atom is reflected without dwelling too long in a potential well.

So far we are aware of only a small number of solid and liquid materials that have been investigated as storage surfaces. The wallshift interaction energies and their applicable temperatures are shown in Table 2 for some interesting examples. Of the Teflons, we include here only the FEP fluorocarbon. Surfaces of liquid helium⁴ and helium³ have been measured at very low temperatures by Hardy. These are extremely interesting surfaces with very low interaction energies and should lead to some interesting results when applied to hydrogen masers.^{12,13}

So far the fluorocarbon surfaces appear to work as low as 25K. If we consider what the fluorine atom looks like when bound to carbon, we note that its electronic structure is that of the noble gas neon, which has the next closed electron shell occurring after helium. It would appear that fluorine, when it looks like neon to an outsider, works well. It is logical, then, to try neon itself as the surface coating material. This is where we are headed as soon as we can get our system to function at temperatures between 8K and 14K.

The expected frequency stability of cold masers, based on thermodynamics and spin exchange behavior, is well below 1×10^{-16} for averaging times of 100 seconds.^{4,13} The potential for greatly increased stability is a strong incentive for continuing the study of wall interactions.

ACKNOWLEDGEMENT

We are indebted to Mr. Martin Andonian of Northeast Cryogenics, Inc., for his assistance in designing the cryostats.

REFERENCES

1. de Saintfuscien, M., Etude des Effets de Paroi dans le Maser a Hydrogene. PhD Thesis, L'Universite Paris XI Centre, D'Orsay. 1975.
2. Crampton, S.B., et. al., Hyperfine resonance of gaseous atomic hydrogen at 4.2K, Phys. Rev. Lett. **42**, 1039 (1979)
3. Vessot, R.F.C., E.M. Mattison, and E.L. Blomberg, Research with a cold atomic hydrogen maser, Proc. 33rd Ann. Freq. Control Symposium, USAERADCOM, Ft. Monmouth, NJ 07703, May 1979, p. 511.
4. Vessot, R.F.C., M.W. Levine, and E.M. Mattison, Comparison of theoretical and observed hydrogen maser stability limitation due to thermal noise and the prospect for improvement by low temperature operation, Proc. 9th Ann. Precise Time and Time Interval Planning Meeting, March 1978. NASA Technical Memo 78104, p. 549
5. Vessot, R.F.C., et. al., Space borne hydrogen maser design, Proc. 8th Ann. Precise Time and Time Interval Planning Meeting, Nov. 1976. NASA document X-814-77-149, p. 277.
6. Mattison, E.M., M.W. Levine, and R.F.C. Vessot, New TE₁₁₁-mode hydrogen maser, Proc. 8th Ann. Precise Time and Time Interval Planning Meeting, Nov. 1976. NASA document X-814-77-149, p. 355.
7. Mattison, E.M., E.L. Blomberg, G.U. Nystrom, and R.F.C. Vessot, Design, construction, and testing of a small passive hydrogen maser, Proc. 33rd Ann. Freq. Control Symposium, USAERADCOM, Ft. Monmouth, NJ, May 1979, p. 549.
8. Zitzewitz, Paul W., Surface Collision Frequency Shifts in the Atomic Hydrogen Maser. PhD Thesis, Harvard University, Cambridge, Mass, 1970.
9. Vessot, R.F.C. and M.W. Levine, Studies of hydrogen maser wallshift for high molecular weight polytetrafluoroethylene, Proc. 24th Ann. Freq. Control Symposium, USAERADCOM, Ft. Monmouth, NJ, May 1970, pp. 270-274.
10. Vessot, R.F.C and M.W. Levine, A method for eliminating the wallshift in the atomic hydrogen maser, Metrologia **6**, 116-117 (1970).
11. Vanier, J. and R.F.C. Vessot, Hydrogen maser wallshift, Metrologia, **6**, 52-53 (1970).
12. Hardy, W.N., and M. Morrow, Prospects for low temperature H masers using liquid helium coating walls, Journal de Physique, **42**, p. C8-171 (1981)
13. Berlinsky, A.J., and W.N. Hardy, Cryogenic Masers, Proc. 13th Ann. Precise Time and Time Interval Planning Meeting, Washington, D.C., Dec 1981. NASA Conference Publication 2220, p. 547.

Performance of Maser with 7-inch (17.8 cm) Diameter Bulb

Table 2

Interaction Energies and Temperatures of Various Surfaces with Atomic Hydrogen

Fig. 2. Cryogenic maser suspended from vacuum housing.



Fig. 1. Cryogenically cooled hydrogen maser with low temperature dissociator.

Fig. 3. Liquid nitrogen cooled hydrogen dissociator

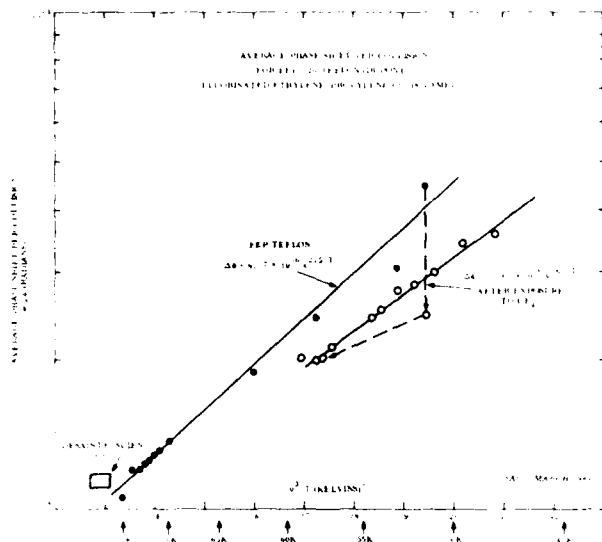


Fig. 4. Average phase shift per collision, $-\Delta\theta$ (radians), from 77K to 48K.

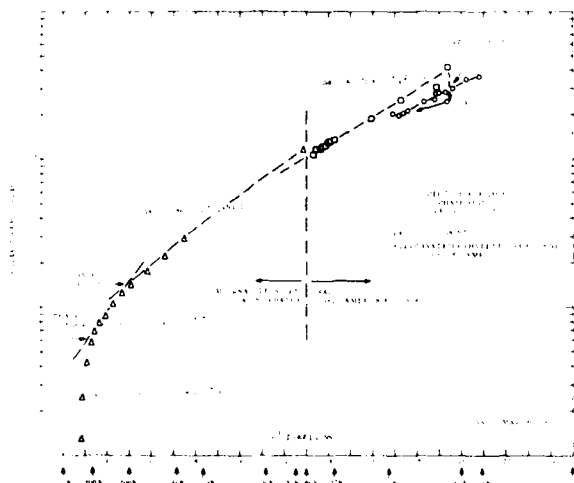


Fig. 5. Average phase shift per collision, $-\Delta\theta$ (radians), from 330K to 48K.

David W. Allan
Time and Frequency Division
National Bureau of Standards
Boulder, Colorado 80303

Abstract

The advent of satellite time and frequency comparison techniques has provided the opportunity for measuring the time and frequency difference between remote clocks with greatly improved accuracies. The paper will give a brief review of various remote clock comparison techniques; in particular the Global Positioning System (GPS) will be highlighted.

The advent of GPS provides the opportunity for the first time for cost effective, high accuracy, operational (automatic), international time and frequency comparisons. It has been demonstrated that where the ephemerides of the satellites are known to within several meters, this translates to an error of only a few nanoseconds in measuring the time difference between remote clocks on the surface of the earth as long as the remote clocks receive the GPS time signal simultaneously. The time difference between the two remote clocks is obtained simply by subtracting the two readings taken at each site resulting from the common-view measurement minus a differential delay constant, which can either be calculated or measured at the outset. Other common-mode errors either cancel or are reduced using this simultaneous viewing approach; e.g., the GPS clock error, the ionosphere, etc. Because of the high inclination to the ecliptic of the GPS satellite orbits, simultaneous common-view is possible between essentially all principal sites in the Northern hemisphere, and principal sites in the Southern hemisphere have common-view with key timing centers in the North.

Some of the other techniques which will be reviewed and discussed will be the "two-way" satellite technique, LASSO, STIFT, the meteorological satellite system including GOES, and some other relevant techniques. The current status of these techniques will be discussed along with their functionality and essential characteristics. Some projections as to what the future holds will also be discussed.

With the advent of atomic clocks we have seen a rapid improvement in the accuracy capabilities within the time and frequency community. Figure 1 is a plot of the accuracy of the primary standards at the National Bureau of Standards indicated by the circles. We are currently using NBS-6 with an accuracy of 8×10^{-14} . We anticipate with optical state selection in cesium to obtain accuracies of a part in 10^{14} . The physics is essentially completed for a standard featuring laser cooling of mercury ions stored in an electromagnetic ion trap, which hopes to yield, within this decade, an accuracy of a part in 10^{15} . The sloping line drawn to fit the circles indicates one decade improvement every seven years.

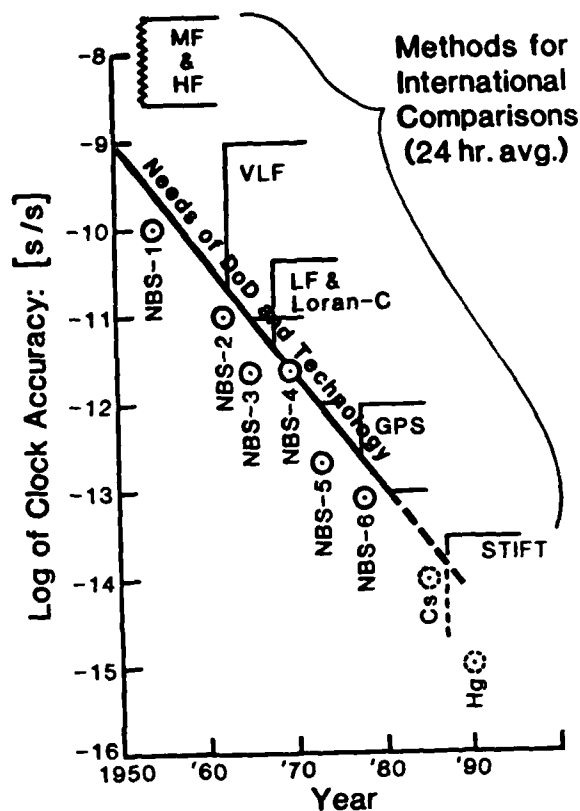


Figure 1. A plot of the accuracies of NBS primary frequency standards compared with some methods of comparing frequencies nationally and internationally.

Comparing standards between one laboratory and another at these accuracy levels is a significant metrology challenge. The comparison methods have usually been much less accurate than the primary standards and less stable than their associated time scales. It has only been of recent time with the advent of satellite techniques that one is able to compare with accuracies that are comparable to, and even better than, the current accuracies of primary standards.

Since 1969 the international method of comparing time has been the Loran-C system. Loran-C has the problem that it is not stable enough to compare state-of-the-art clocks until time averages of a couple of months are taken, and also an annual term gives an additional instability in this technique. Figure 2 shows a bar graph comparison of some satellite techniques that have been studied -- compared with Loran-C. The Shuttle, TDRSS, and LASSO experiments have no near term experimental realization and though hardware has been built, there are some uncertainties as to whether they will be conducted. [1,2,3,4,5] The first row in

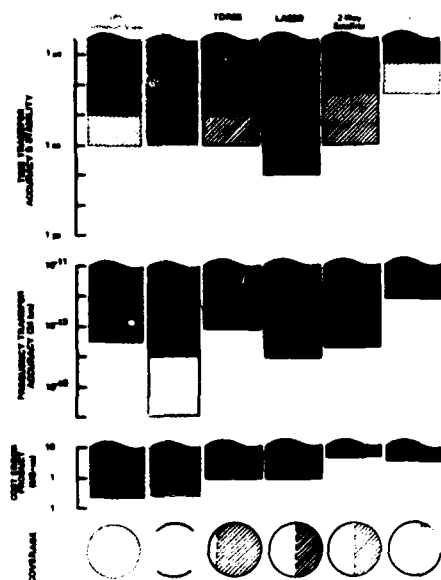


Figure 2. Time and Frequency Transfer Accuracy Comparison of some satellite techniques and Loran-C. The bottom of the bar indicates the accuracy or the stability as explained in the text.

this bar chart illustrates the time transfer accuracy and stability of these various techniques. The bottom of the dashed or shaded part of the curve is the time stability; the bottom of the solid part of the bar is the time transfer accuracy; i.e., the ability to measure the absolute time difference between two remote points. The second row is the frequency transfer accuracy averaged over a 24-hour period. There is a special case for the shuttle experiment (STIFT) which may allow frequency comparisons of a part in 10^{16} accuracy. This will be explained later. The third row is a cost-error product in megadollar-nanoseconds. The bottom of the bar is the economic-accuracy estimate. A lower number means less cost for more accuracy. The bottom row indicates the degree of global coverage.

Using two-way satellite technique [5] in a comparison between Boulder, Colorado, Ottawa, Canada, Washington, DC and Brittany (500 km from Paris, France) time stabilities of a few tenths of a nanosecond were realized and in a two-way double hop experiment between Brittany in Europe and Boulder via NRC in Ottawa a time stability of 6 nanoseconds was realized. The accuracy of this system, interestingly, is not limited by the reciprocity assumption for the nearly 80,000 km up and down signal path which is used, but by the non-reciprocity of the transmitter and receiver delays. The accuracy in the past, due

to this problem, has been of the order of 50 nanoseconds.[6] With some recent work [7] the error due to this problem has been pushed down to the order of six nanoseconds.

The shuttle experiment called STIFT (shuttle time and frequency transfer) features three things [2]; first a laser for ranging to the shuttle from the ground with an accuracy of a few centimeters. Second is featured a three frequency microwave technique which accomplishes a Doppler cancellation and a cancellation of the ionospheric delay. Third, the Shuttle will carry an active hydrogen maser clock with frequency stabilities of the order of one part in 10^{14} and better for sample times

FREQUENCY STABILITY

Long-dashed lines are 1984 projections

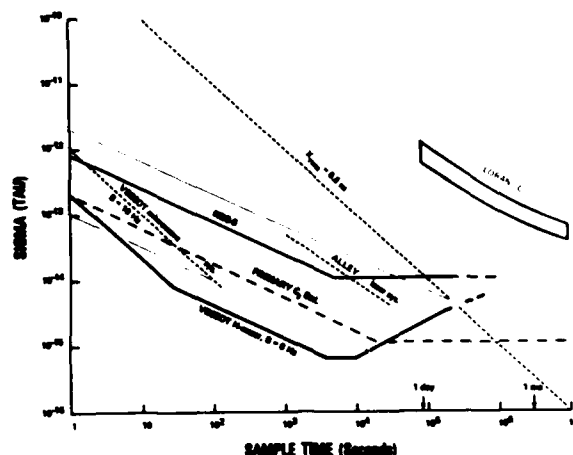


Figure 3. Frequency Stability of 3-frequency microwave (Vessot) Doppler cancellation system compared with other systems. It is denoted by the $1/\tau$ dashed line marked "VESSOT microwave system" and "B = 10 Hz".

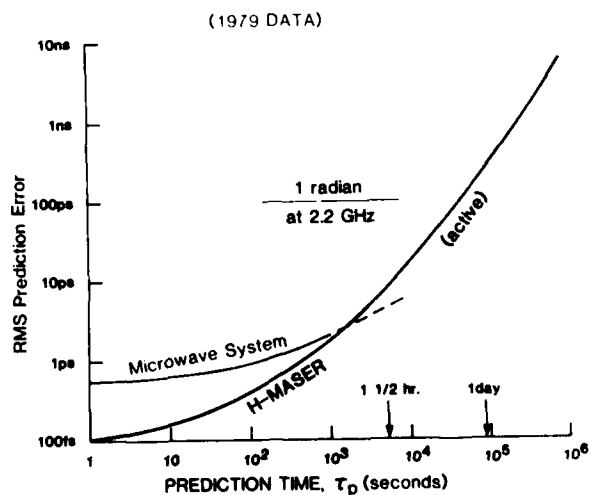


Figure 4. RMS time deviation error as a function of prediction interval for 3-frequency microwave Doppler cancellation system and for a hydrogen maser compared to 1 radian at 2.2 GHz.

convenient to the tracking and orbit times. In Figure 3 the stability of the microwave system is shown which translates to about one picosecond phase stability in the short term (during a single track) and about 10 ps over an orbit period. Figure 4 shows a rms time error analysis of the system. This figure shows that it may be possible to remove the cycle ambiguity of the 2.2 GHz carrier. In which case if this could be maintained over one day, this would translate to about one part in 10^{16} frequency transfer accuracy.

The GPS will feature an 18 satellite constellation. There are nominally four different techniques for using the GPS system as shown in Figure 5. One can also use a combination of the common-view and the viewing of several satellites and reduce the system measurement noise even more than using common-view alone. With this combination approach, a few parts in 10^{14} have been achieved for sample times of one day. Figure 6 gives some estimates, using the common-view technique, of the rms errors in comparing time between two remote sites, A and B, assuming a 25 nanosecond ephemeris error. The satellite is at an altitude of 4.2 earth radii. The common-view technique yields a large amount of common-mode error cancellation.[8] This technique has been studied between the Naval Observatory and NBS in Boulder, CO. Nine portable clock trips over the course of about one year agreed with the GPS in common-view to within an rms error of ten nanoseconds. Because of the high latitudes of most timing centers and because of the 63° inclination of the GPS satellite orbits, very good common-view is available between most sites as can be seen in Figure 6. The analysis illustrated in Figure 6 does not consider other important error sources, such as the ionosphere, troposphere and multipath distortion. Though the ionospheric model has been studied in a lot of detail over the years, the model elements used in the GPS are only about 50% accurate. Fortunately, in the common-view technique, some significant common-mode ionospheric error cancellation can be achieved -- to the order of a few nanoseconds. During nighttime viewing, the total delay for the nighttime sky at $\pm 40^\circ$ latitude is of the order of 5 nanoseconds so even for fairly large hour angles the differential delay may be stable to about a nanosecond. The tropospheric delay for high viewing angles is only a few nanoseconds at the GPS frequencies and can be estimated to the order of a nanosecond if desired. The errors due to multipath distortion can amount to a few nanoseconds, but can nicely be dealt with from a stability point of view by making the common-view measurement once per sidereal day i.e., the viewing geometry for the two sites to the satellite stay the same. Receivers are now being made by several different organizations. We have built some prototype units at the NBS. One can quite economically build a receiver with a long-term stability of about 1 ns with an antenna about the size of a man's thumb. Depending on the noise figure and configuration of the receiver's front end short-term stabilities may vary from 1 ns to nearly 30 ns. Their instabilities are typically characterized as a white noise phase modulation process hence the values are amenable to averaging, and the confidence on the mean will improve as the square root of the number of values. The NBS prototype units are configured with a very powerful and friendly software system using a Z-80 microprocessor and a built-in 0.1 ns resolution time interval counter. All that is needed from the local site clock are 5 MHz and 1 pps signals. The receiver is RS-232 compatible and coupled to a modem and can automatically be dialed by a computer -- making computer

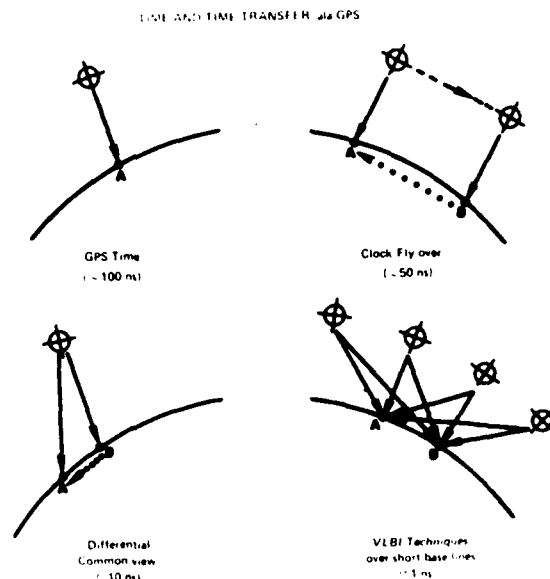


Figure 5. Four different methods of doing time transfer using GPS.

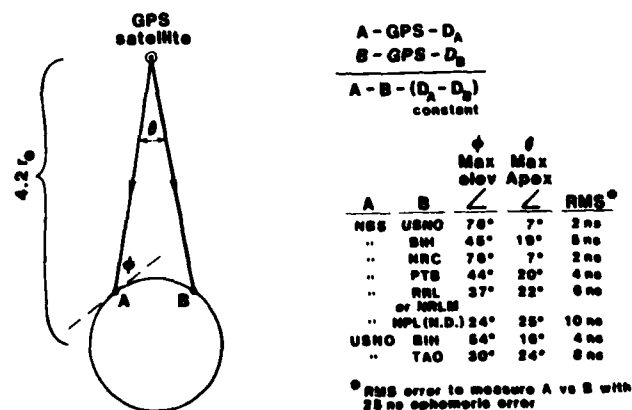


Figure 6. RMS error estimates to do time transfer between two sites A and B with an RMS GPS satellite ephemeris error of 25 ns -- showing the advantage of using GPS in common-view. The locations are: NBS (Boulder, CO); USNO (Washington, DC); BIH (Paris, France); PTB (Braunschweig, W. Germany); RRL, TAO, and NRLM are near Tokyo, Japan; and NPL (New Delhi, India).

analysis very straight forward. Internationally, the Mark III computer system is being used for comparing GPS common-view data.

The frequency stability measured between the U.S. Naval Observatory clock ensemble and the NBS clock ensemble is shown in Figure 7. The stability of the difference between these two ensembles was measured at the one or two in 10^{14} level at sample times of the order of a week and longer. An assessment was made of the GPS in common-view measurement limit and it appears to be about 1 part in 10^{13} at one day and going down on a Mod. $\sigma_y(\tau)$ plot as white noise phase modulation ($\tau^{-3/2}$). In fact we have a data point at 3.5×10^{15} at $\tau = 10$ days. This stability value was verified over several months of data. The NBS receiver stability at $\tau = 1$ day is 8 parts in 10^{15} -- making the receiver noise negligible for almost all clocks for $\tau > 1$ day.

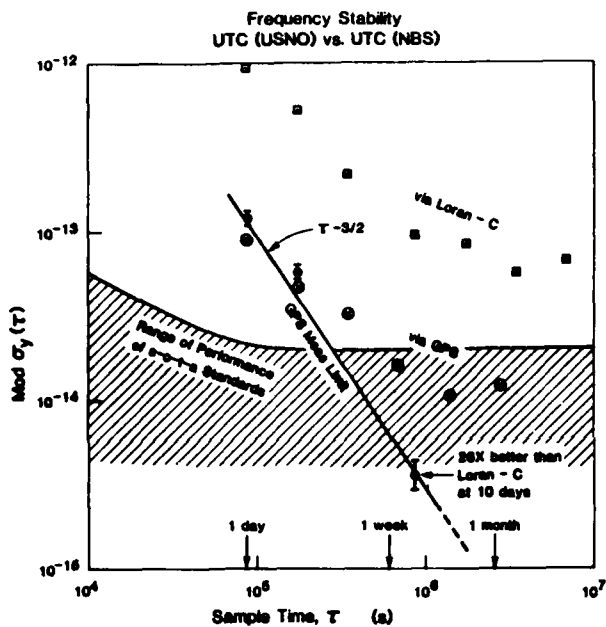


Figure 7. Frequency stability comparisons using Mod. $\sigma_y(\tau)$. The $\tau^{-3/2}$ line corresponds to white noise PM. The s-o-t-a are state-of-the-art frequency standards.

Consider now a combination of common-view comparisons employing three or more satellites. If the three satellites, denoted i, j, and k, are available to make independent common-view measurements between sites A and B, and the measurements are made within a few hours of each other so that the random time deviations between the clocks at A and B are small, then a variance analysis can be performed on the sidereal daily common-view values as follows: Let

$$\sigma_{AB}^2 = \sigma_A^2 + \sigma_{sv_i}^2 + \sigma_B^2, \quad (1)$$

where σ_{AB}^2 denotes the measurement variance from the common-view data time series, which is composed of the variance of each of the site clocks plus the variance of measurement fluctuations resulting from using GPS space vehicle in common-view. Also, assume on the basis of the satellites being independent that

$$\sigma_{sv_{ij}}^2 = \sigma_{sv_i}^2 + \sigma_{sv_j}^2, \quad (2)$$

which can be measured by taking a variance of the difference between the i and j measurements, since the measurements are taken at nominally the same time. Performing this same analysis on the difference between the i and k and the j and k measurements allows the computation,

$$\sigma_{sv_i}^2 = \frac{1}{2}[\sigma_{sv_{ij}}^2 + \sigma_{sv_{ik}}^2 - \sigma_{sv_{jk}}^2], \quad (3)$$

which can be subtracted from equation (1) yielding an estimate of $\sigma_A^2 + \sigma_B^2$ with the measurement contribution being only a few parts in 10^{14} at $\tau = 1$ day, and improving as $1/\tau$.

In the Hafele-Keating[9] experiment, performed some years ago with some clocks from the Naval Observatory, the clocks were flown in both directions around the globe and verified the Sagnac effect, since a disparity of about 200 nanoseconds was observed. One can do that same experiment now using either GPS or other satellite techniques using the photon as the portable clock. The size of the effect is proportional to a projected area on the equatorial plane -- the area being that made by the circumnavigating photons. The coefficient of proportionality is 1.6 ns/Mm^2 . For GPS signals going around the globe the Sagnac effect will be about 200 to 300 nanoseconds depending on the particular geometry. We plan to do this experiment this summer and fall, as soon as we have around-the-globe capability. Tables 1 and 2 show the characteristics of major time and frequency dissemination systems. Table 1 shows the operational type of systems (VLF, LF and high frequency stations, etc). Table 2 shows estimates of the state-of-the-art satellite techniques giving the time stability and frequency accuracy of each of these techniques. Figure 8 is the $\sigma_y(\tau)$ plot showing the fractional frequency stability of these various operational and state-of-the-art satellite techniques as a function of sample time. Clearly, satellite techniques for time and frequency metrology on a national and international basis have been and will be extremely important in keeping up with state-of-the-art time and frequency standards now and in the future.

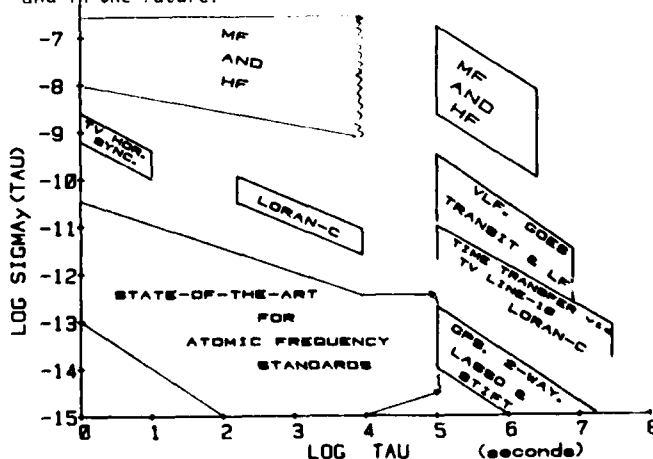


Figure 8. Frequency stability of different remote comparison techniques. The MF and HF notations are e.g., WWV, CHU, JJY; the LF stations are e.g., WWVB, MSF, DCF77.

TABLE 1. CHARACTERISTICS OF THE MAJOR T&F DISSEMINATION SYSTEMS

DISSEMINATION TECHNIQUES		24 HOUR SYNTONI- ZATION ACCURACY	SYNCHRONI- ZATION ACCURACY	AMBIGUITY	COVERAGE FOR STATED ACCURACY
VLF RADIO	GBR, NBA, OMEGA, ETC.	1×10^{-11}	ENVELOPE 1 - 10 ms	1 CYCLE	NEARLY GLOBAL
LF RADIO	STANDARD FREQUENCY BROADCAST (e.g., WWVB)	1×10^{-11} PHASE 24h	ENVELOPE 1 - 10 ms	YEAR	USA - LIMITED (WWVB)
	LORAN-C	5×10^{-12}	1 μ s (GND) 50 μ s (SKY)	TOC 15 MIN PHASE 10 μ s	SPECIAL AREAS
HF/MF RADIO	STANDARD FREQUENCY BROADCAST (e.g., WWV)	1×10^{-7}	1000 μ s	CODE - YEAR VOICE - 1 DAY TICK - 1 s	HEMISPHERE
TELE- VISION (VHF/HF RADIO)	PASSIVE LINE-10	1×10^{-11}	1 μ s (line of sight)	33 ms	NETWORK COVERAGE (Useful only in local areas in USA)
SATEL- LITE (UHF RADIO)	GOES	2×10^{-10}	50 μ s	1 YEAR	WESTERN HEMISPHERE
	TRANSIT	3×10^{-10}	30 μ s	15 MINS	GLOBAL
PORTABLE CLOCKS	PHYSICAL TRANSFER	1×10^{-13}	10 ns to 100 ns	N/A	LIMITED P- TRANSPORTATION

TABLE 2. INTERNATIONAL TIME AND FREQUENCY COMPARISON ($< 1 \mu$ s)

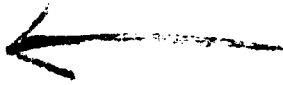
T/F Transfer Technique	Synchronization Accuracy	Time Stability	24 Hour Syntonization Accuracy	Coverage
(1) [†] GPS (Common-view)	10 ns	a few ns	$\lesssim 10^{-13}$	Global
(2) Shuttle (STIFT)	1 ns	0.001 ns	$\lesssim 10^{-14}$	To $\pm 57^\circ$ Latitude
(3) GOES (Trilateration)	10	a few ns	$< 10^{-13}$	All but near the poles
(4) LASSO	1 ns	0.1 ns	$\sim 10^{-14}$	Depends on Implementation
(5) [†] GPS	40 ns*	10 ns	$\sim 3 \times 10^{-13}$	Global
(6) [†] 2-Way (Communica- tion Satellite)	< 10 ns	$\lesssim 1$ ns	$\sim 10^{-14}$	All but near the poles
(7) [†] Portable Clock	10 ns to 100 ns	N/A	$\lesssim 10^{-12}$	Global (Best accuracy within reasonable driving vicinity of Air Ports)
(8) [†] Loran-C	500 ns	$\lesssim 40$ ns	$\lesssim 10^{-12}$	Excludes most of Asia and Southern Hemisphere

*This inaccuracy may increase if the GPS C/A code is deteriorated for strategic reasons.

[†]These techniques have been demonstrated.

Acknowledgments

The author deeply appreciates the assistance of Dr. Marc Weiss and Ms. Trudi Pepler in performing computations for the paper. The work of Mr. Dick Davis, Dr. Weiss, and Mr. Al Clements in developing the equipment and in data acquisition has been invaluable. The cooperation of the USNO staff in data acquisition and exchange has been most helpful. The editorial comments of Dr. Robert Kamper and Mr. Roger Beehler are also greatly appreciated.

9. Hafele, J. C., and Keating, R. E., "Around-the-World Atomic Clocks, Predicted Relativistic Time Gains." *Science*, 177 (4044), 166-168, 1972a. Hafele, J. C., and Keating, R. E., "Around-the-World Atomic Clocks Observed Relativistic Time Gains." *Science*, 177 (4044), 168-170, 1972b.
- 

References

1. Davis, D., Weiss, M., Clements, A., and Allan, D., "Unprecedented Syntonization and Synchronization Accuracy Via Simultaneous Viewing with GPS Receivers; Construction Characteristics of an NBS/GPS Receiver." *Proc. 13th Ann. Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, Dec 1-3, 1981.
2. Allan, D. W., Alley, C. O., Ashby, N., Decher, R., Vessot, R. F. C., Winkler, G. M. R., "Ultra-Accurate International Time and Frequency Comparison Via an Orbiting Hydrogen-maser Clock." *Third Symposium on Frequency Standards and Metrology. Journal de Physique*, Vol. 42, Colloque C-8, Supplement No. 12, Dec. 1981, p. 395-413.
3. Chi, A. R., "Satellite Time Transfer via TDRSS and Applications," *Proc. 11th Annual Precise Time and Time Interval Applications and Planning Meeting*, Washington, DC, 45-65, November, 1979.
4. Serene, B. and Albertinoli, P., "The LASSO Experiment," *Proc. 11th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, Washington, DC, 145-171, November, 1979.
5. Costain, C. C., Boulanger, J. -S., Daams, H., Hanson, D. W., Beehler, R. E., Clements, A. J., Davis, D. D., Klepczynski, W. J., Veenstra, L., Kaiser, J., Guinot, B., Azoubib, J., Parcelier, P., Freon, G., and Brunet, M., "Two-way Time Transfer Via Geostationary Satellites NRC/NBS, NRC/USNO and NBS/USNO via Hermes and NRC/LPTF (France) via Symphonie," *Proc. 11th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, Washington, DC, p. 499-520. November, 1979.
6. Detoma, E., and Leschiutta, S., "Two-way Sequential Time Synchronization: Preliminary Results from the SIRIO-1 Experiment." *Proc. 12th Ann. Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, Washington, DC. November, 1979.
7. Imae, M., Okazawa, H., Sato, T., Urazuka, M., Yoshimura, K., and Yasuda, Y., "Time Comparison Experiments with a Small K-band Antenna SSRA System via a Domestic Geostationary Satellite." *Digest Conference on Precision Electromagnetic Measurements*, June 1982 IEEE.
8. Allan, D. W., and Weiss, M. A., "Accurate Time and Frequency Transfer During Common-view of a GPS Satellite." *Proc. 34th Ann. Symp. on Freq. Control*, Philadelphia, PA, p. 334-347. May, 1980.

AN INTERNATIONAL TIME TRANSFER EXPERIMENT

C. Wardrip

NASA-Goddard Space Flight Center
Greenbelt, Maryland

J. Buisson, O. Oaks, M. Lister and S. Stebbins

Naval Research Laboratory
Washington, D.C.

B. Guinot, M. Granveaud, G. Freon and B. Dubois

Bureau International de L'Heure
Paris, France

W. Schluter and K. Nottarp

Institut Fur Angewandte Geodasie
Wetzell, Germany

V. Reinhardt, R. Kruger, P. Dachel, E. Detoma, J. Ingold, R. Roloff and T. Stalder

Bendix Field Engineering Corporation
Columbia, Maryland

Summary

This paper presents data obtained during an international time transfer experiment between the Bureau International de l'Heure (BIH), Paris, France; the Institut Fur Angewandte Geodasie (IFAG), Wetzell, Germany and the Bendix Field Engineering Corporation (BFEC), Columbia, Maryland utilizing GPS timing receivers developed for the NASA-Goddard Space Flight Center (GSFC) by the Naval Research Laboratory. The purpose of the experiment was to determine the accuracy of the GPS time transfer receivers for the NASA laser ranging network and to intercompare geographically separated atomic frequency standards.

Two NASA-GSFC hydrogen masers and a high performance cesium beam frequency standard in thermally controlled chambers were used for the frequency baseline at BFEC. The IFAG frequency baseline was established using an EFOS-1 hydrogen maser and two Oscilloquartz cesium beam frequency standards. The data from France is comprised of a high performance cesium beam frequency standard at BIH and a hydrogen maser at the Centre National d'Etudes des Telecommunications (CNT). Data is presented on the intercomparison of masers and cesiums using direct comparison to the Naval Observatory (USNO), BFEC, BIH, CNT and IFAG via portable clocks, Loran-C and TV Line 10. Data on the GPS NAVSTAR satellites, with typical rms values of less than 50 nanoseconds, is also presented. A discussion of the time transfer technique and frequency measurements is included.

INTRODUCTION

As an outgrowth of the Navigation Technology Satellite (NTS) timing receivers development (ref. 1) in 1977 by the Naval Research Laboratory (NRL) and the Goddard Space Flight Center (GSFC), a joint effort was started

in 1979 to develop Global Positioning System (GPS) timing receivers using signals radiated by the GPS satellites. In support of the GSFC Crustal Dynamics Program (ref. 2), the GPS timing receivers were designed for use in the GSFC Transportable Laser Ranging Network, which requires submicrosecond timing for correlation of highly accurate satellite tracking data with time.

In 1981 a prototype GPS receiver was completed and successfully tested at the NASA's Merritt Island Tracking Station at the Kennedy Space Center (ref. 3). During this time NASA was also involved in the SIRIO-2/LASSO experiment (ref. 4). It was planned to use the new GPS receivers in conjunction with the LASSO timing experiment to provide an independent measurement of the synchronization between participating LASSO laser ranging sites in Europe and the U.S. The expected accuracy of the LASSO was to be used to calibrate the GPS receivers and to determine their accuracy for use in the GSFC laser ranging network. The other intent of this LASSO/GPS experiment was to intercompare Hydrogen masers and other atomic standards over intercontinental distances.

Although the launch of SIRIO-2/LASSO was not successful, it was decided to continue with the GPS phase of the experiment and install GPS receiver at selected sites in Europe and the U.S. The participating agencies were the NASA Goddard Space Flight Center, the Naval Research Laboratory, the Bureau International de l'Heure and the Centre National d'Etudes des Telecommunications (Paris, France), the Institut fur Angewandte Geodasie (Wetzell, Federal Republic of Germany) and the Bendix Field Engineering Corporation. The participation of the U.S. Naval Observatory is acknowledged for providing the common time reference to which all observations were reduced.

AD P002454

The experiment was run for a period of seven months, from October 1982 through April 1983. Independent measurements were provided via portable clock trips in the U.S. and Europe and via Loran-C and TV line-10 in the U.S.

The NAVSTAR Global Positioning System (GPS)

Many references to the GPS signal structure and operation can be found in the literature (see references in ref. 3 and ref. 5). GPS comprises three segments. The Space Segment, the Control Segment and the User Segment. The phase III Space Segment will consist of a constellation of 18 to 24 satellites, six to eight in each of three orbital planes. The satellite orbits are nearly circular at an altitude of about 20,000 km and inclined 55° to the equator. The period is one-half of a sidereal day, resulting in a constant ground track, but with the satellite appearing 4 minutes earlier each day.

Each satellite transmits its own identification and orbital information continuously. The transmissions are spread spectrum signals, formed by adding the data to a direct sequence code, which is then biphasic modulated onto a carrier. At the present time, the control segment consists of a master control station and four monitor stations. The user segment consists of a variety of platforms containing GPS receivers, which track the satellite signals and process the data to determine position and/or time by simultaneous or sequential reception of at least four satellites.

Time Transfer Method

To transfer time via a GPS satellite, pseudo-range measurements are made consisting of the propagation delay of the received signal biased by the time difference between the satellite clock and the ground station reference clock (see fig. 1 and ref. 2 and 3). Data from the navigation message contain the satellite clock information and the satellite ephemeris, which allows one to compute the satellite position. Since the position of the satellite and of the ground station are known, the computed propagation delay can be subtracted from the pseudo-range and then corrected for the GPS time offset, to determine the final result of ground station time relative to GPS time, which can be referenced to the U.S. Naval Observatory.

The final results obtained from a single frequency receiver, such as the ones used in this experiment, will contain a small error due to the atmospheric delay, which may be modeled and corrected. If two ground stations clocks are synchronized to GPS time, the results can be subtracted to obtain the time difference between the ground station clocks. This can be done at any time, but the best results are obtained when data is taken simultaneously by each ground station from the same satellite (common view), since any error contributed by the satellite clock is cancelled when the data is subtracted.

GPS Time Transfer Receiver (TTR)

The GPS TTR (ref. 2) is a microcomputer based system which operates at the single L-band frequency of 1575 MHz. The receiver uses the C/A code only (1.023 MHz), tracking this code to within 3% of a chip (30 ns). The receiver has the capability to track satellites throughout their doppler range from horizon to horizon, and can track any GPS satellite by changing the receiver internal code. In this experiment the receivers were kept stationary, however, work is in progress to add position/navigation capabilities, so as to obtain time and position either on a stationary or on a moving platform. The block diagram in fig. 2

shows the GPS receiver configuration. Operator interface with the receiver is provided by a keyboard and CRT display. The time data is stored on disks and can also be outputted to an external printer/computer via a serial data interface.

General Organization of the International GPS Experiment

The general organization of the experiment is depicted in fig. 3. Clock comparisons at the respective sites were carried on via the NAVSTAR/GPS system using GPS satellites 3, 4, 5 and 6. Independent time comparisons were made at the beginning and at the end of the experiment via portable clock trips to participating facilities. Local TV links were used to intercompare site clocks when possible, namely between the USNO and BFEC and between BIH and CNT in Paris. In addition, regular clock trips (3 per week) were made between the USNO and BFEC.

BIH & CNT

The receiver was initially installed at the BIH on October 3, 1982. A malfunction occurred on Oct 30 and the receiver was returned to NRL for repair and reinstalled at BIH on January 5, 1983. The local time base at BIH for the experiment was a HP 5061 Cesium standard configured as shown in fig. 4.

On February 2 the receiver was installed at the Centre National d'Etudes des Telecommunications (CNT), also located in Paris, where a hydrogen maser was available as the local time reference and configured as shown in fig. 5. The hydrogen maser was replaced with a cesium standard near the end of the experiment.

Wettzell Laser Tracking and VLBI Site (Germany)

The time system for the GPS experiment at the Wettzell site was arranged as shown in fig. 6. The local time base consisted of an Oscilloquartz EFOS (Etalon de Frequence Oscilloquartz) hydrogen maser, that was continuously compared with an Oscilloquartz cesium standard. A Loran-C receiver provided an independent reference. Initial set-up and tracking began October 12, 1982, and the reference frequency standards for the receiver are as listed in the following table:

Dates	Type of Standard	
12 Oct. - 10 Nov. 1982	Cs	131
11 Nov. - 15 Dec. 1982	Cs	173
15 Dec. - 29 Apr. 1983	H-maser	EFOS

Bendix Field Engineering Corporation (BFEC)

The equipment configuration at BFEC is shown in fig. 7. The local time base was provided by an NP-2 hydrogen maser, with a HP-5061 (opt. 004) cesium standard as a back-up. Time comparisons to the USNO using the TV line-10 method were provided daily; in addition three portable clock trips to the USNO were made every week. The local time standards were kept in thermally controlled chambers.

Results of the Timing Experiment

The data presented were taken during the first 75 days of 1983 at BIH, CNT, IFAG and BFEC. At the end of January 1983 the receiver at BIH was taken to CNT, where data were taken until the end of April. Fig. 8 shows typical raw data behavior, showing the aging of the NP-2 hydrogen maser at BFEC versus UTC (USNO) compared via GPS satellites 3, 4, 5 and 6. The best

results are for satellites 3 and 5; the portable clock trips are in excellent agreement (within 50 ns) with the GPS synchronization results after a constant calibration offset (bias) was removed from the GPS data.

The following figure (fig. 9) shows the GPS synchronization data at BFEC referenced to the UTC (USNO) via TV-line 10 measurements; these were calibrated using portable clock trips, and the same constant calibration offset was removed. The data show an agreement of ± 50 ns (on the average) with the GPS measurements.

The data shown in fig. 10 represent the characteristic behaviour of the NP-2 hydrogen maser at BFEC as compared to UTC (USNO) via the GPS satellites. The measurements taken from satellites 3, 4, 5 and 6 are shown in this composite plot demonstrating the agreement of the time measurements as determined by individual satellites.

Fig. 11 through 14 shows the data taken at each location indicating the synchronization accuracy obtained via each satellite, with frequency offset and aging removed.

The additional data in fig. 15 was taken at BFEC from day 81 thru day 105 and demonstrates the accuracy and the reliability of the synchronization obtained via the GPS satellites over a period of three and a half months of continuous operation.

The results of this experiment are summarized in the table below, showing that the overall accuracy of the synchronization via the Global Positioning System is consistently better than 100 ns, and meet the synchronization requirement of the NASA laser ranging network.

International GPS Time Transfer Experiment Absolute Accuracy

Date	Station	UTC(USNO) - UTC(STATION) Via P.C. (us)	Via GPS TTR (us)	(PC-GPS) (us)
10/16/82	WET	-55.796	-55.663	-.133
4/16/83	WET	+ 8.880	+ 8.906	-.026
10/9/83	BIH	- .460	- .471	+.011
4/19/83	CNT	+ 2.020	+28.954	+.066
11/5/82	BFC*	+ .020	+ .042	-.022
3/23/83	BFC*	- .710	- .723	-.013

* Constant GPS receiver bias removed
P.C. = Portable Clock

References

1. Raymond, L., Oaks, J., Osborne, J., Whitworth, G., Buisson, J., Landis, P., Wardrip, C., Perry, J., "Navigation Technology Satellite (NTS) Low Cost Timing Receiver Development", Proceedings of the Eighth Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, November, 1976.
2. J. Oaks, J. Buisson, C. Wardrip - GPS time transfer receivers for the NASA transportable laser ranging network, NASA/GSFC X-814-82-6 (April 1982).
3. J. Oaks, A. Franks, S. Falvey, M. Lister, J. Buisson, S. Wardrip and H. Warren, Prototype design and initial test evaluation of a GPS time transfer receiver, NRL report 8608 (July 27, 1982).

4. S. C. Wardrip, E. Detoma, ESA/NASA SIRIO-2/LASSO time transfer experiment, 4th Annual Geodynamics Meeting, Washington, D.C. (January 1982).
5. The Institute of Navigation, Journal of the Institute of Navigation, Special issue dedicated to the GPS, Summer 1978.

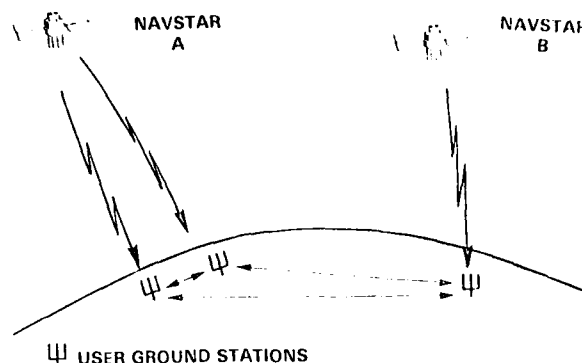


Figure 1.

NAVSTAR GPS Station Synchronization by Time Transfer

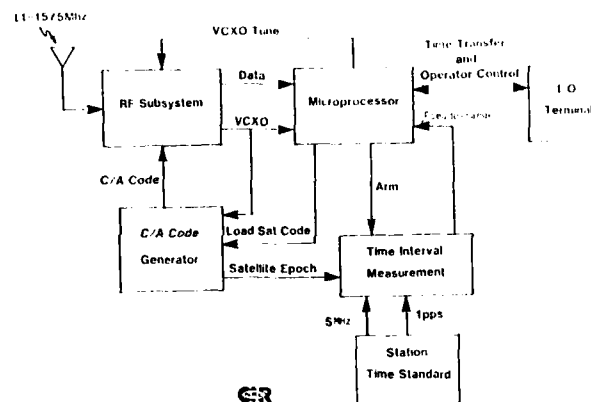


Figure 2.

GPS Time Transfer Receiver Block Diagram

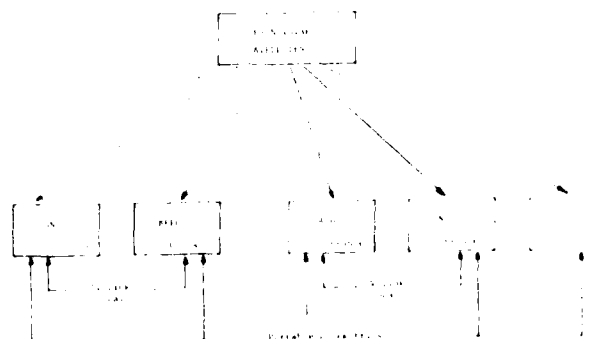


Figure 3.

General Organization of the GPS Experiment

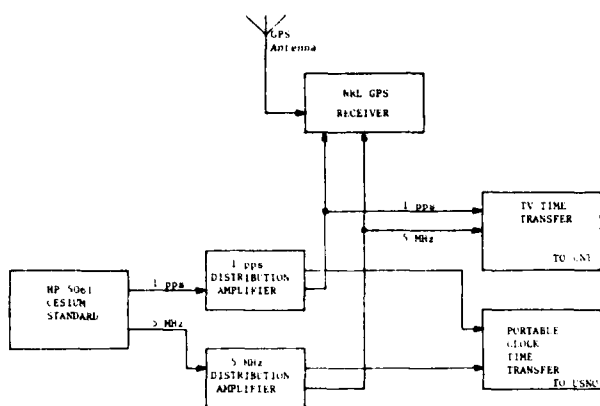


Figure 4. Bureau International De L'Heure (BIH)
GPS Experiment Timing System

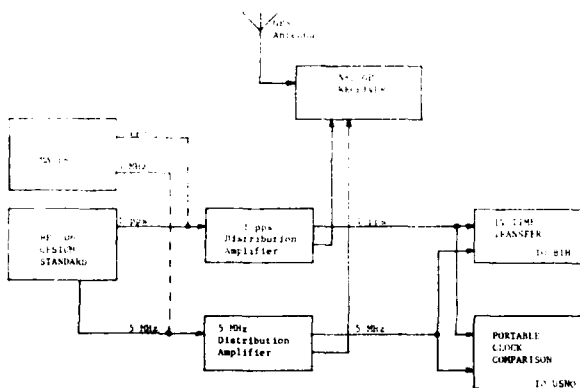


Figure 5.
Centre National D'Etudes Des Telecommunications (CNT)
GPS Experiment Timing System

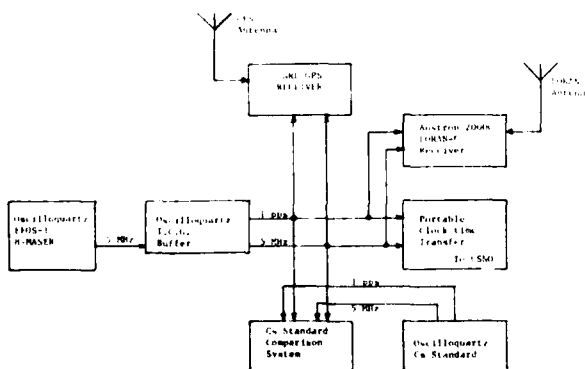


Figure 6.
Wettzell Laser Tracking and VLBI Site
GPS Experiment Timing System
NOTE: TCG = Time Code Generator

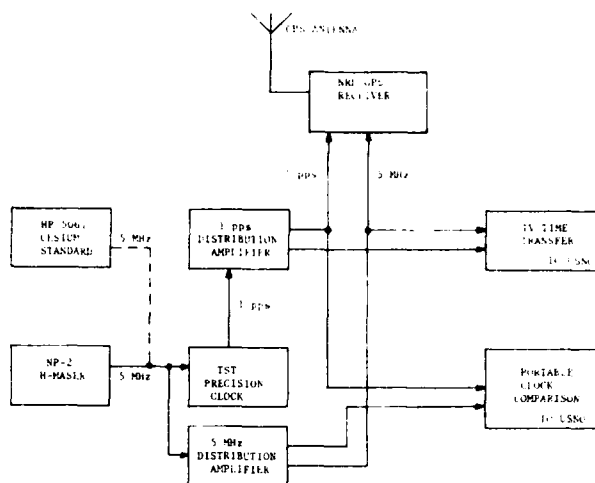


Figure 7.
Bendix Field Engineering Corporation (BFEC)
GPS Experiment Timing System

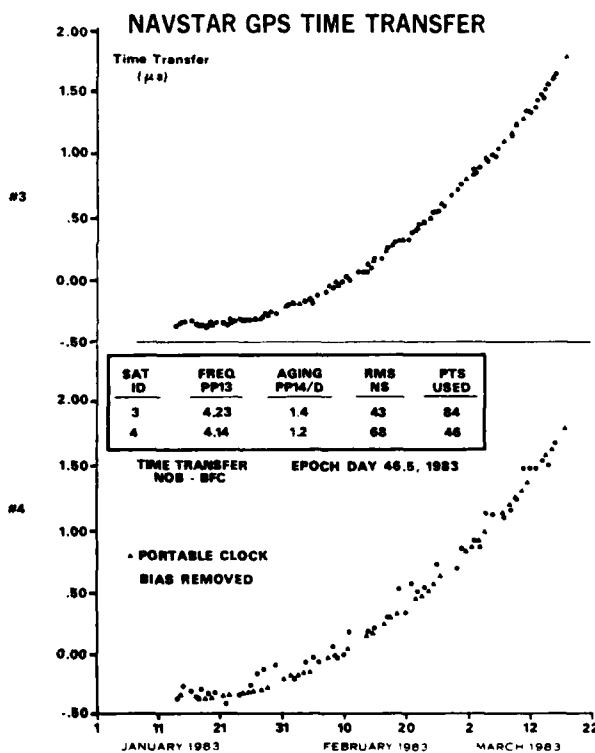


Figure 8a

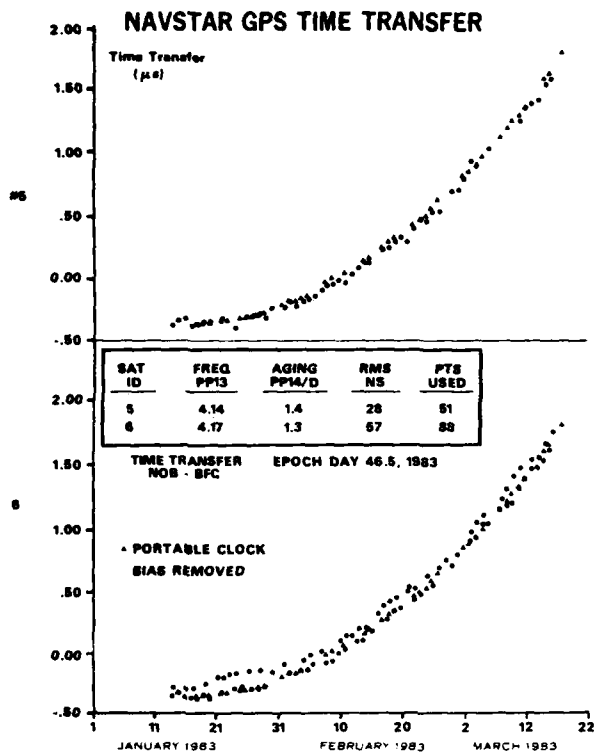


Figure 8b

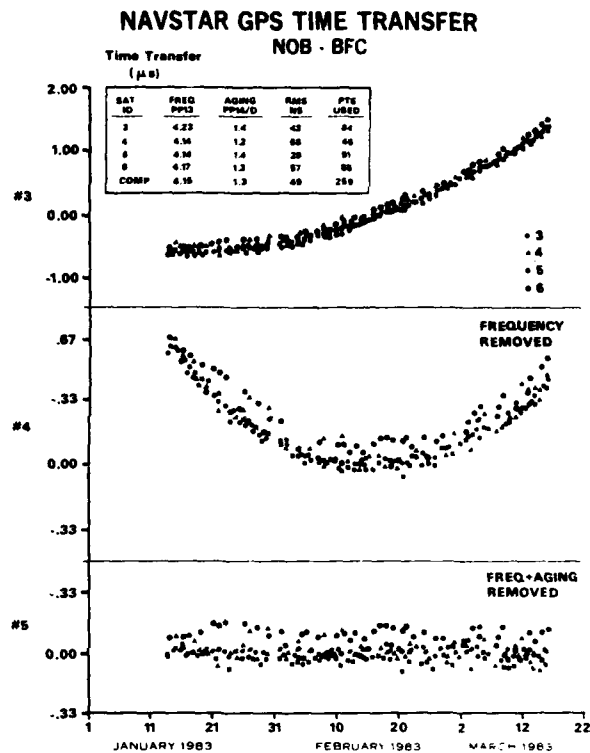


Figure 10

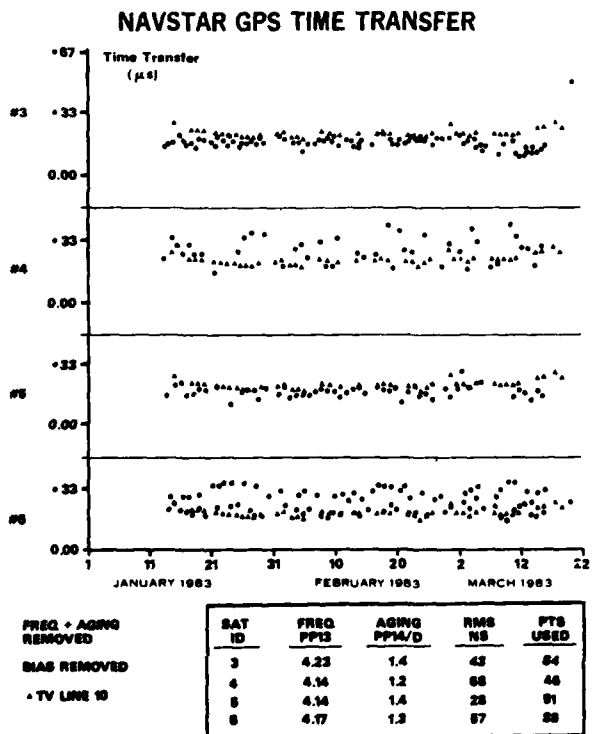


Figure 9. Time Transfer NOB - BFC

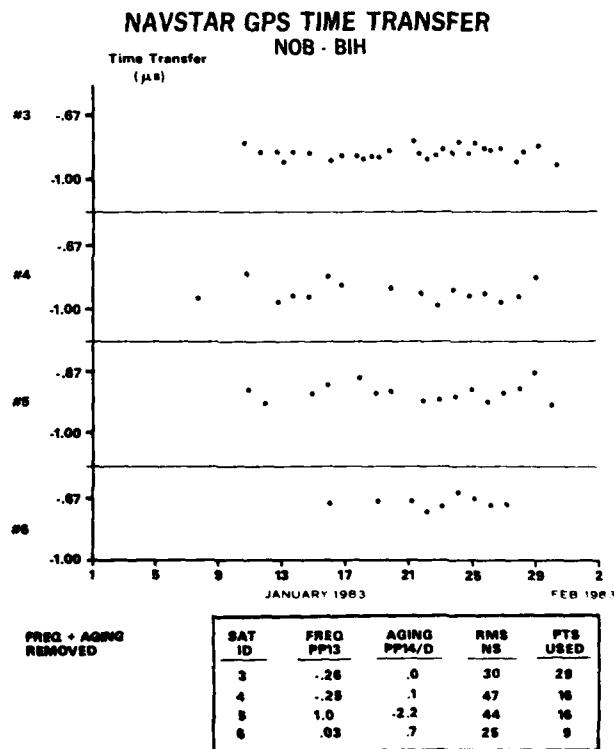


Figure 11

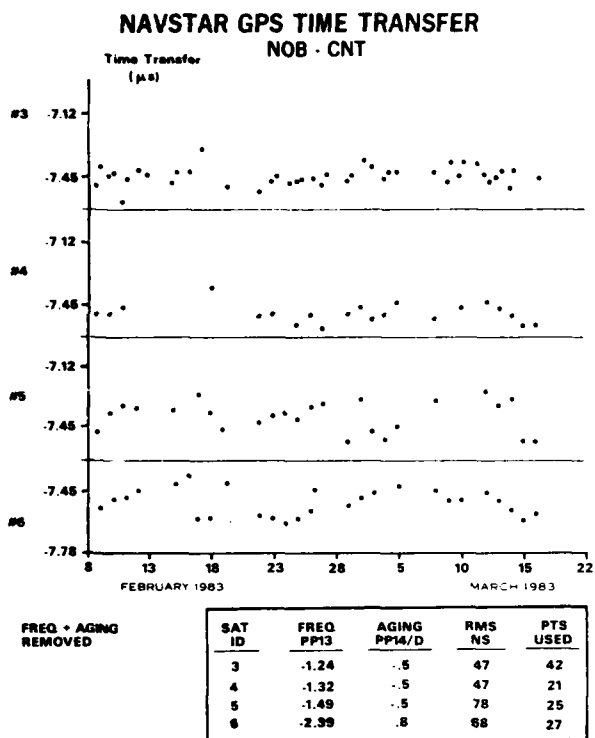


Figure 12

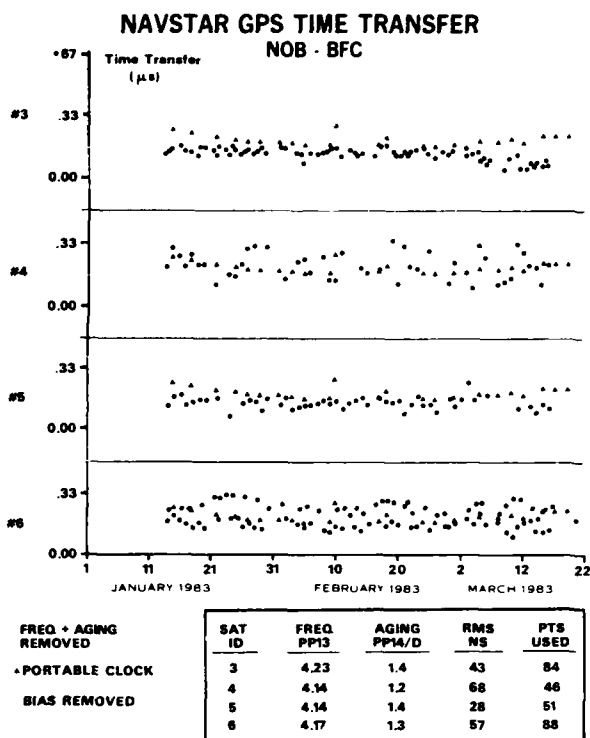


Figure 14

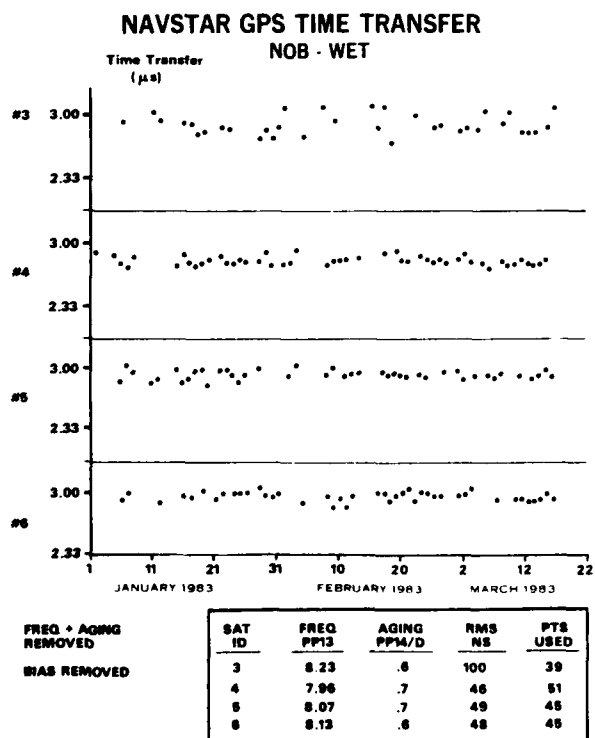


Figure 13

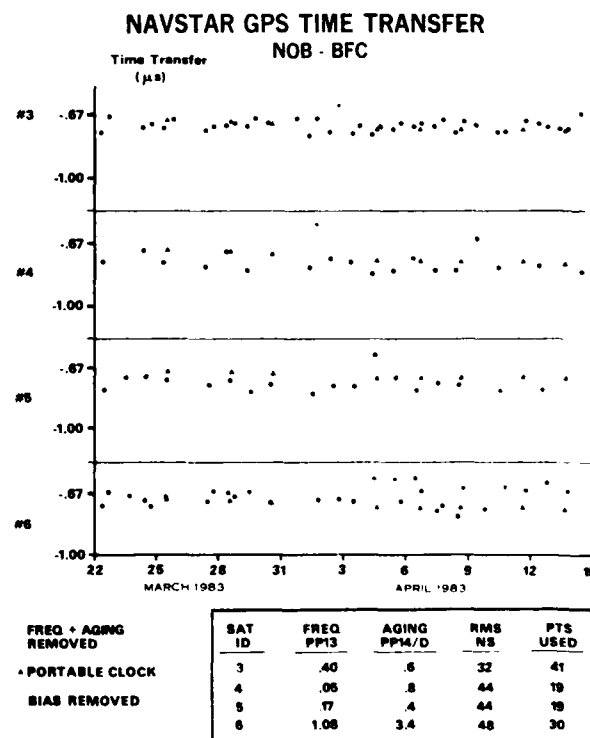


Figure 15

D. Kirchner

Dept. of Communications and Wave Propagation, Technical University of Graz, Austria

SUMMARY

After a short overview of the activities of those institutes which share the facilities of the Observatory Lustbühel, the timekeeping station and other systems of the observatory employed for time transfer purposes are described. Furthermore the results obtained with different frequency and time comparison methods and an overview of envisaged activities are given.

Introduction

The Observatory Lustbühel is situated on the top of a hill at the outskirts of the city of Graz. Figure 1 shows the geographic location of Graz, the second largest city of Austria, and the distances from Graz to other timekeeping laboratories, to the transmitters of the Norwegian and Mediterranean LORAN-C chains and to other VLF and LF transmitters.

The Observatory Lustbühel is jointly owned by four institutes of the two universities of Graz and is also the home of the Space Research Institute of the Austrian Academy of Sciences which acts as an interdisciplinary link between these institutes. The following Table lists three of the four institutes and indicates their principal research fields pursued at the observatory:

- Institute of Meteorology and Geophysics:
Satellite beacon measurements (differential-Doppler and Faraday-effect measurements)
Theory of Magnetospheric physics
- Institute of National Surveying and Photogrammetry:
Satellite geodesy: optical observations
Doppler observations
LASER observations
- Department of Communications and Wave Propagation:
Microwave propagation
Satellite communication
Timekeeping

Figure 2 shows the observatory with its two domes, one for astronomical observations and the other for the LASER telescope, and the different antennas on the roof. In the foreground at the left hand side one can see the pillar of the camera for optical observations and on the right hand side the satellite earth station.

Timekeeping

The timekeeping station serves two purposes: firstly it has to support all measurements carried out at the observatory with the required precise time and frequency information and second the installation is used to conduct experiments to study the performance of novel and more precise time comparison methods. For this purpose on the other hand, the facilities provided by the satellite earth station, the LASER ranging station and other installations are available.

Basic installation

All of the timekeeping equipment is situated in a special air conditioned room on the observatory's ground floor. The temperature is kept at $21 \pm 1^\circ \text{C}$ and the humidity at $45 \pm 5\%$. The equipment is operated via an uninterruptible power supply.

Figure 3 shows the main part of the equipment which is

housed in five standard 19" racks and two further ones for the computer and parts of its peripherals and for the GPS-receiving system. A simplified functional diagram of the set-up is shown in Figure 4. The system consists of the following main parts:

- Frequency and time standards
- Equipment used to compare the local standards with other standards around the world
- Measurement and data acquisition system. The system controller is also used to exchange data with other laboratories via a worldwide time sharing computer service.

Furthermore equipment exists for frequency and time distribution inside the observatory and to the other buildings. There are also devices to permanently measure temperature, humidity and barometric pressure. The standby power supply for the timekeeping station is housed separately in the observatory near the clock room.

The frequency standards are two Cesium-beam type standards model HP 5061A with clock options and separate standby power supplies. One of the standards uses an option 004 high performance Cesium-beam tube. One of the two serves as master and is used in connection with an ILC model 6459 high performance digital clock which provides system time.

Since the beginning of regular operation (1971) frequency and time comparisons are made by VLF/LF, television and LORAN-C timing techniques. Since 1980 also satellite time transfer techniques are employed.

The measurements are made automatically by an electronic counter with a single shot resolution of 2 ns which may, if required, be replaced by one of 20 ps. The counter, switches for signal routing, a digital clock and several other devices are all interconnected via the HP-Interface Bus (HP-IB, GP-IB, IEEE-488 interface or IEC-Bus) with an HP 9825 desktop computer as controller to form an automated measurement and data acquisition system¹. In order to standardize as much as possible RS-232C is the only other interface standard used besides the IEEE-488. All data are stored on magnetic tape cartridges and the most important ones are also recorded by strip chart recorders.

GPS-receiving equipment

Since the end of 1982 signals emitted by the satellites of the Global Positioning System (GPS) are used by means of a commercial receiver model TTS-502 of Stanford Telecommunications, Inc. (STI) for frequency and time comparisons. A detailed description of the GPS time transfer concept and the STI TTS-502 system is given elsewhere².

Figure 5 shows the GPS antenna and the pre-amplifier housing on the top of the building and Figure 6 the GPS-receiver together with the above mentioned system controller to which it is connected via an RS-232C interface. Figure 7 shows a schematic of the TTS-502 system and demonstrates the connection to the computer. The receiver/processor is interfaced to the system terminal via an RS-232C interface. The computer can be connected by a switch to the RS-232C printer port of the terminal or alternatively to the optional RS-232C output of the re-

ceiver/processor. In switch position A and running the appropriate programme in the computer, the data from the GPS-receiver are collected, preprocessed in the computer and stored on magnetic tape. In switch position B all information displayed on the system terminal is put out via the computer to a printer, e.g. to obtain a hardcopy of the data base actually used by the receiver.

The software used to incorporate the GPS-receiver in the automated timekeeping system is illustrated in Figure 8. One has to distinguish between the main programme (P&D) and the automatic GPS-measurement program (GPS-A) which is activated at certain times by the main programme and those which are activated by operator intervention using special function keys of the computer. The latter programmes are used to input GPS-measuring times in accordance with the receiving schedule previously fed into the TTS-502 via the system terminal at data base initialization, to exchange GPS-measuring results via the General Electric Mark III Computer Service, to produce hard copies of the TTS-502 system terminal display (GPS-B) and to display, print out and store GPS data at any time.

The automatic GPS measurement programme (GPS-A) runs at predefined times and performs the following tasks:

- Recording of raw values on magnetic tape. All data categories will be stored which have been selected at data base initialization of the TTS-502 system.
- Computation of a linear regression over each measurement interval.
- Storage of the parameters of the linear regression and other values in the format used by the U.S. Naval Observatory as described elsewhere³. In addition to these parameters also the tropospheric and ionospheric corrections and other parameters may be stored.
- Printout of the time-error and other parameters of interest.

During operation of this programme by operator intervention one can put out the raw values on the display of the computer or initiate a printout of the raw values in the TTS-502 data format on an external printer.

Satellite earth station, OTS-II experiment

Since the end of 1980 one year of time comparisons were carried out via the experimental European communication satellite OTS-II between the Observatory Lustbühl and the Van Swinden Laboratory Delft (Netherlands)⁴. Later on this experiment was continued on a broader basis and at present five European laboratories participate in this experiment. The method used is the well known television-time transfer technique with the only difference that the signals are coming from a geostationary satellite instead of from a ground based transmitter. The main problem with this method results from the fact that even a geostationary satellite moves and since one uses a one-way method an exact knowledge of the satellite position is mandatory as demonstrated in the following: For twenty measurements, one measurement every other second, one obtains a 1 σ standard deviation between 5 and 10 ns but poor orbit determinations may cause jumps of several hundred nanoseconds from one day to the next⁴.

In Graz the satellite earth station of the observatory is used for this experiment. The main characteristics of this station are:

- Antenna diameter: 3 m
- Antenna gain: 48 dB at 11 GHz
- Pre-amplifier: uncooled parametric amplifier or GaAs Fet low noise amplifier
- Down conversion: two stage: first IF (750 \pm 250) MHz, second IF 70 MHz
- System noise temperature: 360 K or 750 K
- Video signal S/N: 40 dB
- Transmitting power: up to 225 W

Since the Lustbühl satellite station is also equipped for transmission attempts have been made to measure the range to the satellite to improve the orbital parameters.

The positions of the participating stations have been established in a Doppler campaign carried out in 1979⁵.

LASER ranging station

The LASER station was set up between 1980 and 1982 mainly for geodetic applications. To take part in the LASSO experiment (LASER synchronization from stationary orbit) as a two-way station in addition to the short pulse low energy Nd:YAG LASER system a ruby LASER with higher energy however at the expense of pulse length has been installed⁶.

Both systems are mounted on the same LASER bench and allow to switch from one LASER to the other to make use of the same LASER telescope. At present the Nd:YAG LASER is in full operation and yields excellent results.

The most interesting features of the LASER station are⁷:

- Two LASER systems:

LASER	pulse width	energy per pulse	repetition rate
Nd:YAG	0.1 ns	0.1 J	2.5 to 10 Hz
ruby	3 to 6 ns	2.5 to 4 J	up to 0.25 Hz

- Optical switching between the two LASER's under computer control
- High precision mount: pointing accuracy <5 arc secs
- Receiving telescope: aperture 50 cm optimized for ruby and frequency-doubled Nd:YAG wavelengths
ISIT-TV-Camera
- Detection: single photon electron level
- Transmitting telescope: beam divergence variable from 5 to 200 arc secs (computer controlled)
- Aircraft detection system
- Minicomputer controlled
- Observation periods restricted to 0000-0600 local time by civil aviation authorities.

Analogous to the timekeeping station all instruments are connected via the IEEE-488 interface and, whenever

possible, use is made of commercially available instruments. Figure 9 shows the LASER bench and Figure 10 the computer, system terminal, monitor of the ISIT-Camera, measuring equipment and the mount controlling hardware. Figure 11 shows the receiving and transmitting telescope with the aircraft detection package above the receiving telescope and the ISIT-Camera below it.

Data exchange

For data exchange with other institutions the conventional telex network and the worldwide computer time sharing service Mark III of General Electric is used. At present the latter proves to be the most cost effective and time saving means of data exchange. The connection between Mark III and the measurement processor of the timekeeping station, which is used as an intelligent terminal by running a terminal emulation software is done by a 300 baud modem and dial-up telephone line. Data received via Mark III may be stored on magnetic tape. Furthermore a mailbox system is installed in Mark III by which immediate communication between the laboratories connected to Mark III is possible.

Results

The generated timescale is UTC(TUG) and its rate is identical with the rate of one of the two Cs-frequency standards.

Since 1977 the timekeeping station at the Observatory Lustbühl takes part in the work of the BIH on TAI, initially with one and later with two clocks. The weights given to these two clocks by BIH are shown in Figure 12. Figure 13 gives the relationship between UTC and UTC(TUG) obtained by LORAN-C measurements and clock transportations carried out by the U.S. Naval Observatory as published by the BIH in Circular D. The differences between UTC-UTC(TUG) as computed by the BIH and obtained from clock transportations are given in Table 1.

Year	1977	1978	1979	1981
Clock Tr.-BIH	0 ¹⁾ μ s	0.2 μ s	0.3 μ s	0.4 μ s

1) definition of origin

Table 1: Differences between the clock transportations and LORAN-C measurements (BIH Annual Report)

Furthermore from Figure 13 one can see that the clock is free running and time steps are introduced to keep the UTC(TUG) scale within about $\pm 5 \mu$ s of the UTC scale.

Since the beginning of 1983 time comparisons are made on a regular basis using the transmissions of the GPS satellites and the results are made available via Mark III. For a period of about four months (January 10, 1983 to May 3, 1983) Figure 14 gives the results of time comparisons between the U.S. Naval Observatory and the Observatory Lustbühl using the GPS satellites Navstar #4,5, and 6 (SV#8,5,9) in a common-view mode⁸. The USNO

measurement times and the values UTC(USNO)-GPS are made available by USNO via Mark III. The approximate coordinates of the USNO and the Observatory Lustbühl are 39° North Latitude, 77° West Longitude and 47° North Latitude and 15° East Longitude. The resulting baseline has a length of about 7000 km. Because the measurement times at USNO and TUG may differ by several minutes the USNO data are related to the TUG measurement times by using the intercept and slope of the time-errors reported by USNO. The receiving schedule is updated on a daily basis so that the satellites are always used at the same individual elevations and azimuths. Elevation and azimuths and other figures of interest are listed in Table 2.

From the beginning of the measurements to February 28, 1983 the TTS-502 was operated by the software originally supplied which could not make complete use of the new navigation message disseminated by the GPS satellites since January 15, 1983. Since February 28, 1983 the TTS-502 is operated by a new software which was supplied by STI as an upgrading kit for the receiver.

Therefore one has to distinguish data of three periods:

- January 10 to 11:
Tropospheric and ionospheric corrections applied
- January 12 to February 28:
No tropospheric and ionospheric corrections applied
- Since February 28:
Again tropospheric and ionospheric corrections applied

In order to demonstrate the magnitude of these corrections their ranges as applied by the receiver since January 28, are given in Table 2.

The new software applied on February 28 led to a jump of 87 ns ± 10 ns. This figure is computed by using linear regressions over 10 days of data before and after the jump. For further computations this step and a frequency offset of 3.7×10^{-13} between the two time scales have been removed. Furthermore the readings of January 10 and 11 were reduced by the ionospheric and tropospheric corrections. The residuals from a linear best fit through the modified data are shown in Figure 15 differentiating between the individual satellites.

Also shown in Figure 14 are the results of time comparisons between USNO and TUG using the LORAN-C transmitter Sylt of the Norwegian LORAN-C chain. To be able to distinguish between GPS and LORAN-C data to the latter a constant of 1 μ s is added. One can clearly see a seasonal effect in the LORAN-C data which could also be seen in comparing the OTS-II time transfer results with the corresponding LORAN-C data⁴. On April 29, 1983 a comparison with a portable clock of USNO was made which resulted in a time difference UTC(USNO)-UTC(TUG) = -0.4 μ s $\pm 0.2 \mu$ s. The corresponding values of the GPS and LORAN-C measurements as given by linear regressions over 10 days around April 29 are -0.37 μ s $\pm 0.01 \mu$ s and -0.46 μ s $\pm 0.1 \mu$ s, respectively. Taking into account that the IPPS fed into

Navstar#	SV#	Clock	Laboratory	Elevation	Azimuth	Data age	Troposph.C.	Ionosph.C.
4	8	Rb	USNO	50°	28°	<2 h	not available	
			TUG	43°	319°		10 - 11 ns	14 - 28 ns
5	5	Cs	USNO	62°	44°	<5 h	not available	
			TUG	36°	304°		12 ns	16 - 32 ns
6	9	Cs	USNO	37°	51°	<12 h	not available	
			TUG	60°	295°		8 ns	11 - 23 ns

Table 2: Elevation and azimuth at the measuring times at USNO and TUG; ionospheric and tropospheric corrections and other figures of interest for the used spacecraft.

the GPS-receiver as reference is delayed by 30 ns (6 m of coaxial cable) against UTC(TUG) one arrives at a zero difference between GPS and the portable clock measurement.

Plots of the square root of the Allan variance for the GPS, LORAN-C and the best period (20 days) of the OTS-II⁴ data are given in Figure 16. The stability curve given for GPS is calculated using the modified data of the whole measurement period shown in Figure 14 and using the data obtained by all three satellites. For LORAN-C two curves are given: the one with the lower stability applies to time transfers with USNO and that with the greater stability for time comparisons inside Europe. The best stability experienced with the OTS-II time comparisons are comparable to the stability obtained for the GPS measurements. Also included in Figure 15 are lines which indicate the necessary stability to obtain a time transfer accuracy of 10, 25, 50, 100 and 250 ns, respectively. In addition to that the typical frequency stability of a high performance Cs frequency standard of commercial type is included⁹. One can see that a 1 s accuracy of better than 25 ns is obtained and that for periods of some days the stability of the GPS measurements is greater than the stability of Cs clocks of standard performance.

Envisaged activities

Although in 1981 the responsible authorities were given a detailed research concept, no financial backing after 1983 is yet foreseen.

Assuming further support the following activities are planned:

- Further studies of the use of TV-signals transmitted by geostationary satellites for one-way time transfer measurements.
- Further attempts to improve the orbit determination by range measurements to the satellites.
- Use of the ISIT-Camera of the LASER ranging system to improve the position determination of satellites by optical observations.
- Two-way time transfer experiments using geostationary satellites.
- Upgrading of the timekeeping station: purchase of an additional clock and replacement of the desktop computer by a minicomputer.
- Use of the GPS receiving system to study the performance of the Global Positioning System for both time transfer and position determination.
- Study of LORAN-C groundwave propagation effects with the help of GPS.
- Participation in a future LASSO-type experiment.
- Participation in other continental and intercontinental time transfer experiments.

Acknowledgement

The useful discussions and comments of many colleagues and in particular of Dr. G.M.R. Winkler of the Time Service Division of the U.S. Naval Observatory and his kind support from the early beginning of the timekeeping activities at the Technical University Graz as well as the equipment supplied by the U.S. Naval Observatory are deeply appreciated.

The continuous encouragement of Prof. W. Riedler is

gratefully acknowledged.

The work was made possible by various grants of the Austrian Council for Scientific Research and the Austrian Academy of Sciences.

References

1. Kirchner, D.: Die Automatisierung einer Zeitstation mittels des IEC-Bus-Systems am Beispiel der Station des Observatoriums Lustbühl der Technischen Universität Graz, 10th International Congress of Chronometry (CIC), A2, pp. 111-117, 1979.
2. Van Dierendonck, A.J., Q.D. Hua, J.R. McLean and A.R. Denz: Time Transfer Using Navstar GPS, Proc. PTI Conf. 13th, pp. 389-417, 1981.
3. Putkovich, K.: USNO GPS Program, Proc. PTI Conf., 12th, pp. 387-413, 1980.
4. DeJong, G., R. Kaarls, D. Kirchner and H. Reßler: Time Comparisons Via OTS-II, Proc. PTI Conf., 13th, pp. 347-369, 1981.
5. Schlüter, W. and L. Amberg: Final Results of Doppler Point Positioning at Several Timekeeping Laboratories in Europe, Nachrichten aus den Karten und Vermessungswesen, Reihe II, H. 40, pp. 99-107, 1983.
6. Kirchner, D., P. Pesec, W. Riedler and K. Rinner: Proposal for the Participation in the SIRIO-2/LASSO-Experiment for International Clock Comparison, Proposal to the European Space Agency (ESA), 1979.
7. Kirchner, G.: The LASER Ranging System Graz-Lustbühl, Proc. of the Fourth International Workshop on Laser Ranging Instrumentation, Vol. II, pp. 463-467, 1981.
8. Allan, D.W. and M.A. Weiss: Accurate Time and Frequency Transfer During Common-View of a GPS Satellite, Proc. 34th Ann. Frequ. Control Symp., USAERADCOM, Ft. Monmouth, NJ 07703, May 1980.
9. Hellwig, H.: Microwave Time and Frequency Standards, Radio Sci. 14(4), pp. 561-572, 1979.

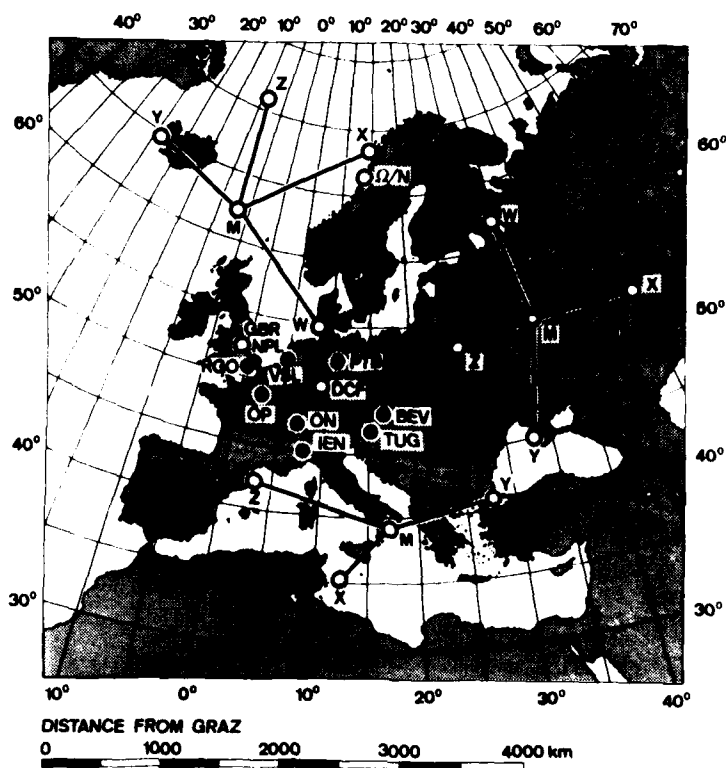


Fig. 1 Geographic location of Graz, other timekeeping laboratories and different VLF and LF transmitters



Fig. 2 Observatory Lustbühel with the satellite earth station

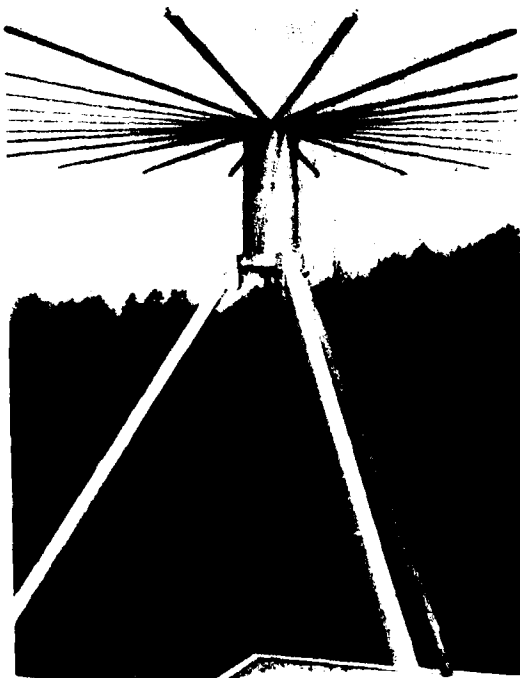


Fig. 5 GPS antenna with pre-amplifier housing



Fig. 6 TTS-502 receiving equipment and system controller with peripherals

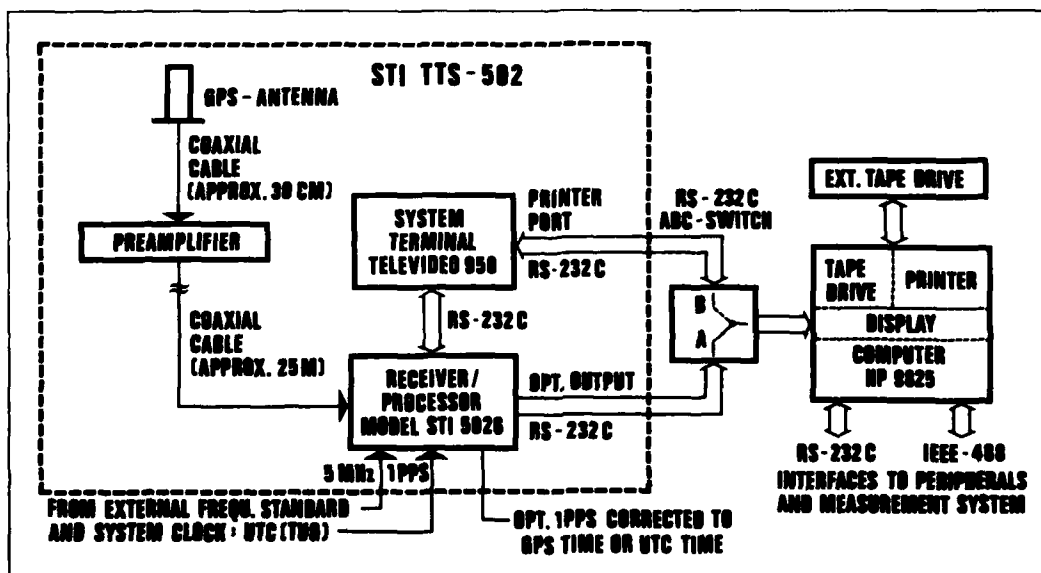


Fig. 7 Schematic of the TTS-502 and its connection to the system controller

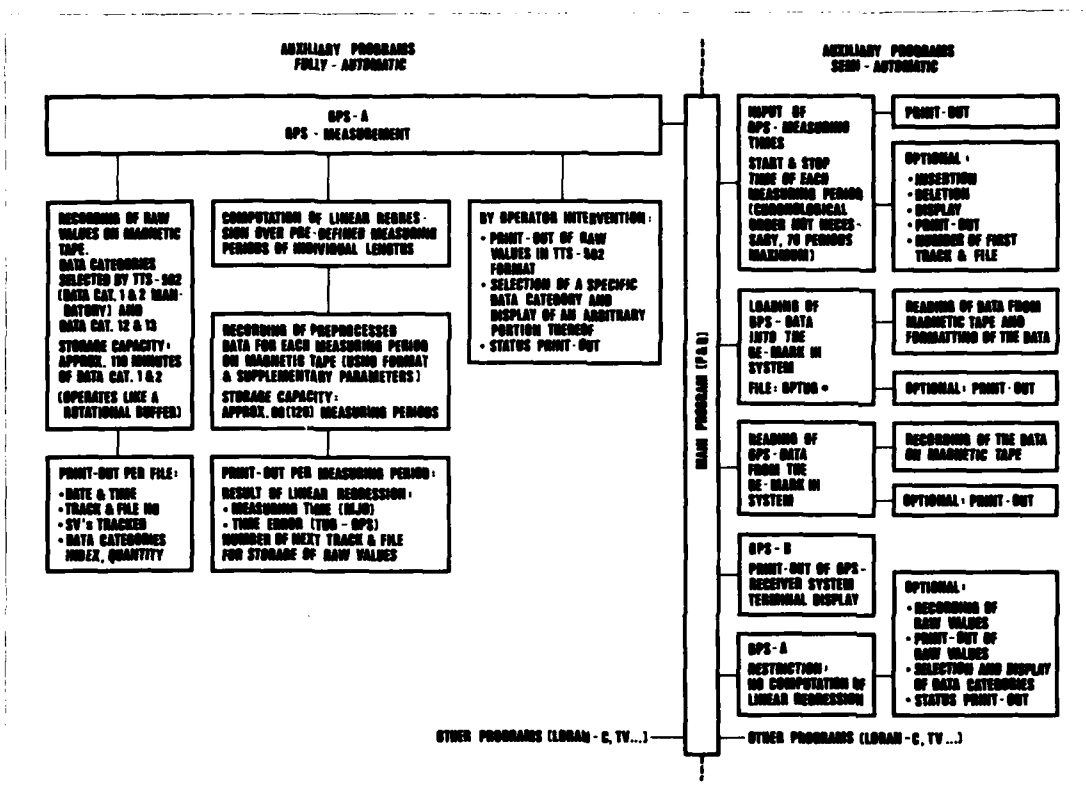


Fig. 8 Description of the software used to incorporate the TTS-502 into the automated measurement system

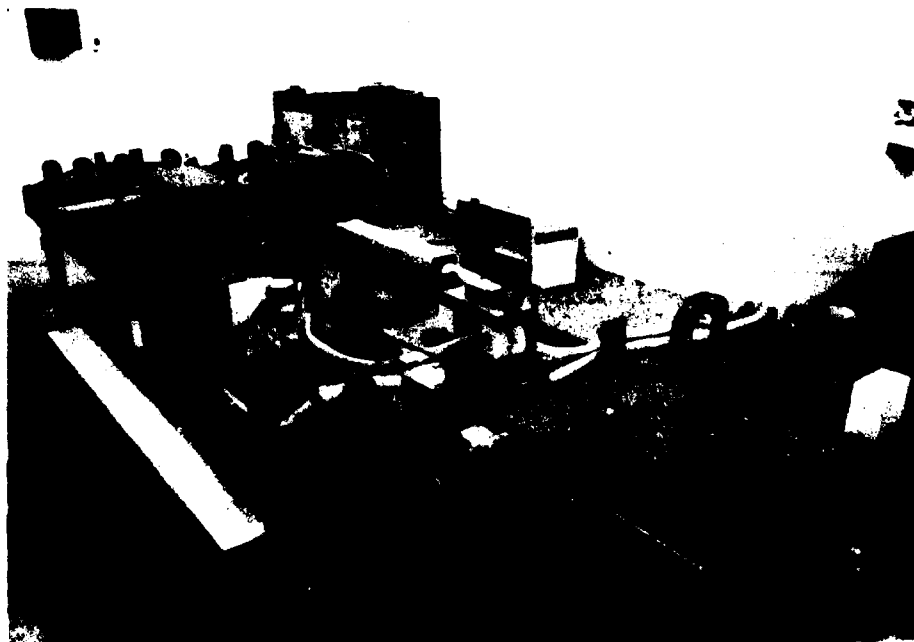


Fig. 9 LASER bench

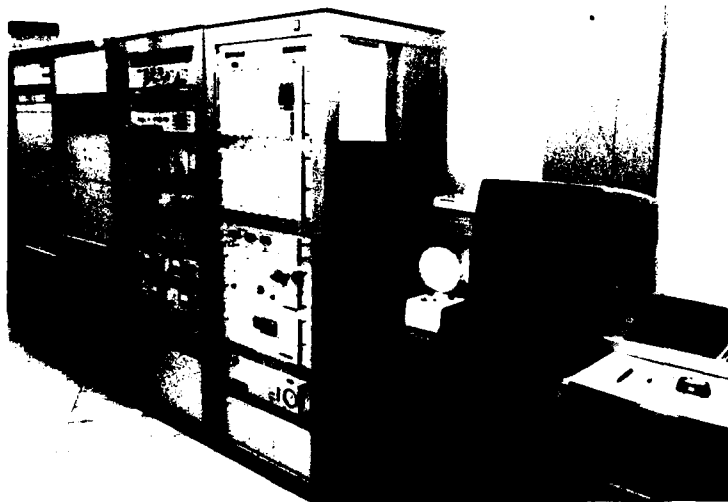


Fig. 10 Control equipment for the LASER station



Fig. 11 Mount and telescope system

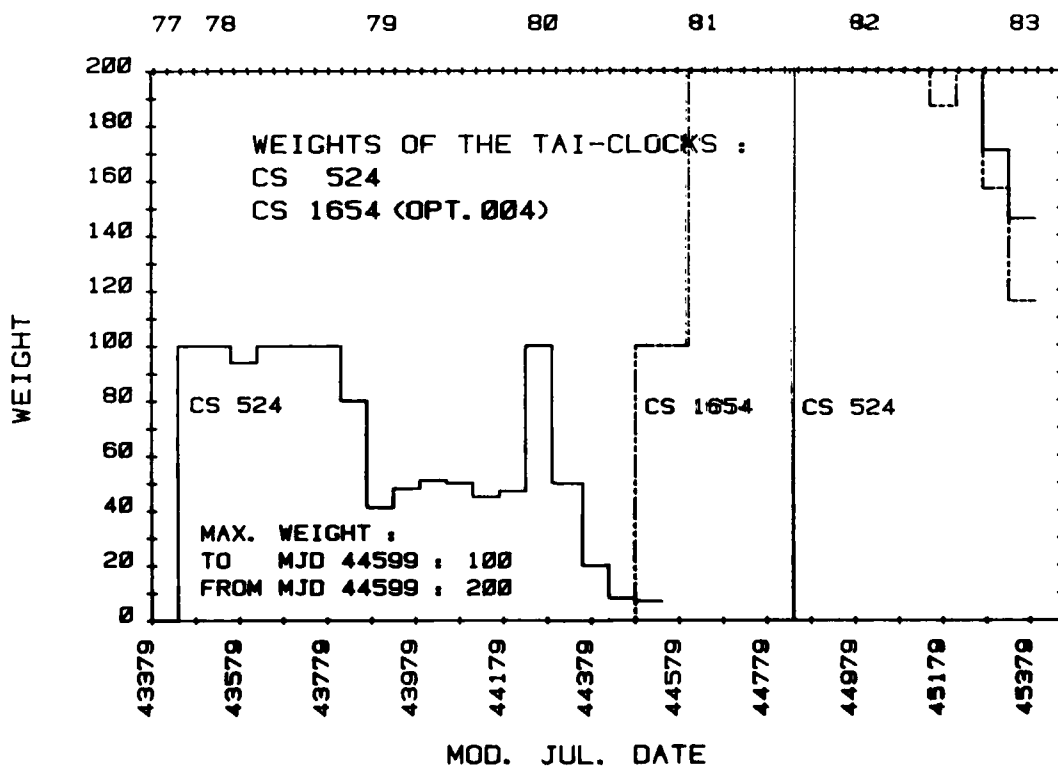


Fig. 12 Weights of the TAI clocks CS 524 and CS 1654(opt.004)

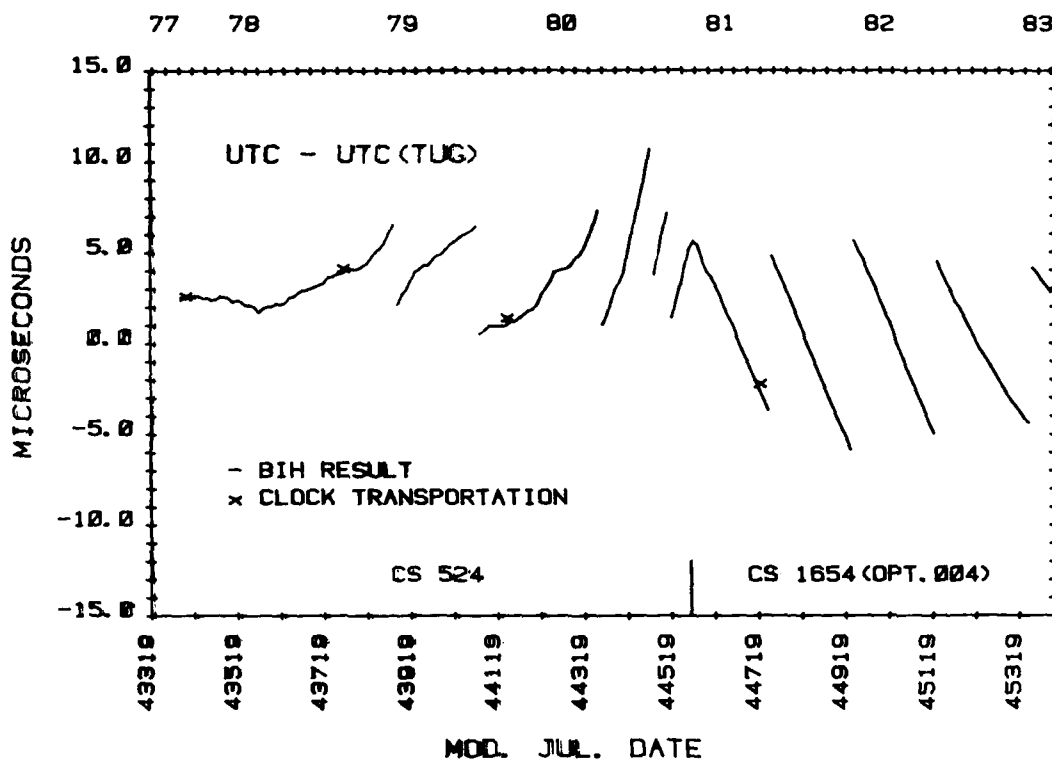


Fig. 13 Diagramme of the difference UTC - UTC(TUG) measured by LORAN-C and the results of portable clock visits carried out by USNO

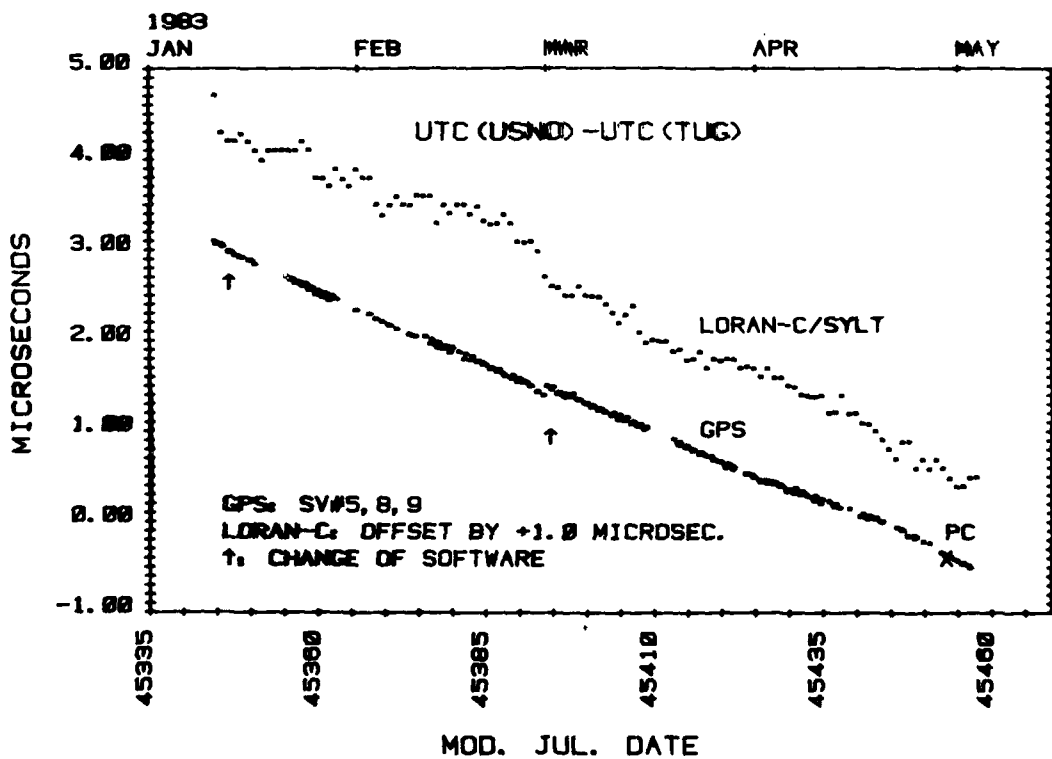


Fig. 14 UTC(USNO) - UTC(TUG) obtained by GPS and LORAN-C measurements and by clock transportation

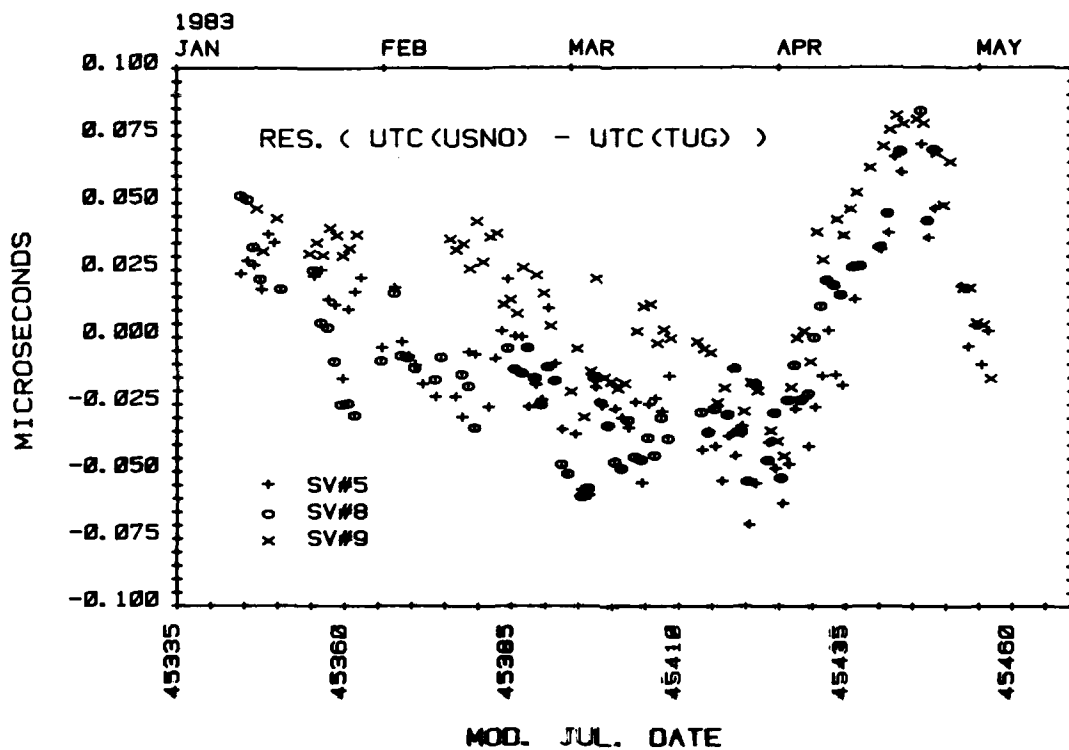


Fig. 15 Residuals from a linear best fit to the GPS data of Figure 14 after removing the time step of 87 ns on February 28, 1983 and subtracting the ionospheric and tropospheric corrections from the readings of January 10 and 11, 1983.

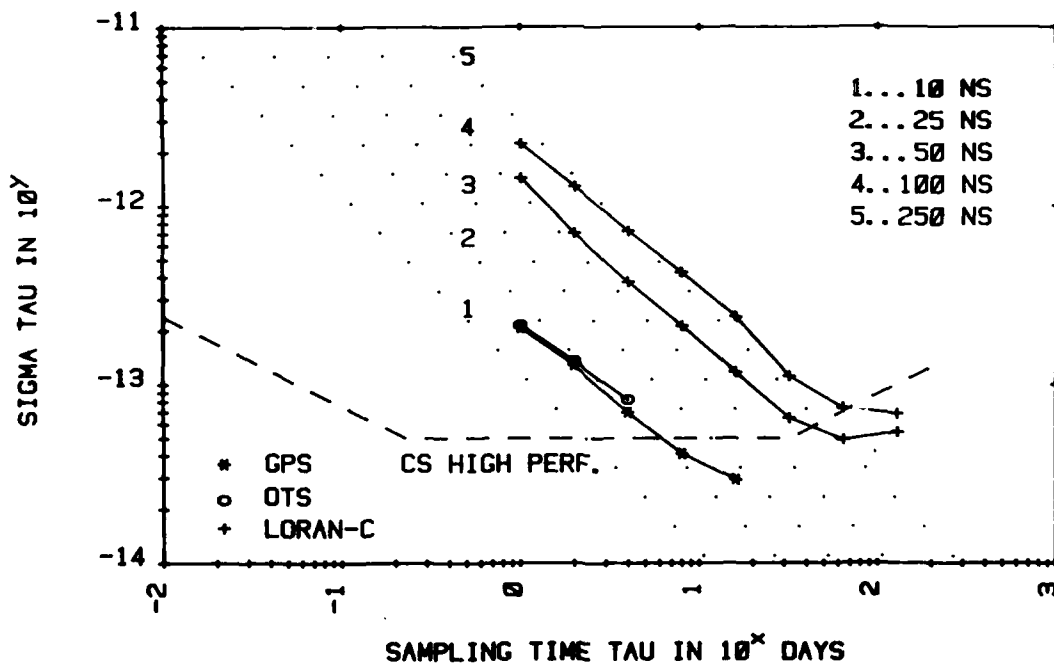


Fig. 16 Stability plot of the GPS, LORAN-C and OTS-II measurements

TIME DISSEMINATION FROM THE
NATIONAL RESEARCH COUNCIL OF CANADA

C.C. Costain, H. Daams, J.-S. Boulanger and R.J. Douglas
National Research Council Canada

Summary

CHU Canada broadcasts continuously on 3.330 MHz, 7.335 MHz and 14.670 MHz with powers of 3 kW, 10 kW and 3 kW respectively. Voice announcements of the time are made in French and English each minute, and an FSK time code is broadcast in the 31st to 39th seconds pulses. The NRC time signals are broadcast Canada-wide each day by the Canadian Broadcasting Corporation at 12.00 noon ET on their French network and at 1.00 p.m. ET on their English network. Talking clocks in French and English can be accessed on the DDD telephone network. The FSK code is also available on the DDD telephone network, with triply redundant code generators and logic circuits to ensure an error free code. Prototype remote clocks have been developed which measure and correct for the delay from the remote location, ensuring millisecond accuracy even via satellite. The two-way satellite time transfer experiment using the *Symphonie* satellite terminated July 1, 1983 after 4 years with France and two years with Germany. Two satellite ground stations with 3 m antennas and 1 W power have been installed near the NRC time laboratory. Two-way transfer using low power CW tones via Telesat Canada Anik satellites in the 6/4 GHz commercial band has achieved subnanosecond precision. Tests of prn code transfer are planned.

* * * *

This report on Time Dissemination from NRCC will include all of the systems mentioned in the summary, but the weighting of the satellite portion will be in the inverse proportion to that in the summary. However, it is important to remember that there are perhaps a million customers for 1 s accuracy for every one at 1 μ s accuracy. For the general public, second accuracy is required, although now with crystal watches they want 0.1 s accuracy. They don't need it, but they want it, and we must take that into account. For operational networks, such as TV, millisecond precision is desirable, and for navigation and communication networks, microsecond precision, and more and more, nanosecond precision, is required.

The NRC time signals are broadcast Canada-wide daily on the AM stations of the Canadian Broadcasting Corporation, at 12:00 noon ET on the French network, and at 1:00 pm ET on the English network. Talking clocks in French and English at the NRC laboratories can be accessed on the DDD telephone network.

CHU Canada broadcasts continuously on 3.330 MHz, 7.335 MHz and 14.670 MHz with powers of 3 kW, 10 kW, and 3 kW respectively. These odd frequencies were chosen over fifty years ago, for practical communication considerations, at the top

of the 80, 40 and 20 meter amateur bands. Voice announcements are made in French and English each minute and an FSK time code is broadcast in the 31st to 39th seconds pulses.

The FSK code is also available on the DDD telephone network, with triply redundant code generators and logic circuits to ensure an error free code. Prototype remote clocks have been developed which measure and correct for the delay from the remote location, ensuring millisecond accuracy even via satellite. We hope these clocks will find their way into most cable TV system, where an automatic call/week would ensure time display of 0.1 s accuracy. This would make precise time continuously available to a very large fraction of the Canadian public. We are also trying to persuade TV stations to display graphic clocks, ticking off perhaps 10 seconds to the hour, as is common in many European TV networks.

Four years ago¹ we reported here on the two-way time transfer between France and Canada using the *Symphonie* satellite. This experiment terminated July 1, 1982, after 4 years, and for the last two years Canada/France, Canada/Germany and France/Germany transfers were made once or twice a week. The results were used by the BIH in constructing TAI, but with an experimental satellite the experiment was bound to end. The results of the transfer between NRC and PTB in Braunschweig are given in Figure 1. The time scales in both laboratories are controlled by primary cesium clocks.

To have a practical operational time transfer network, we must have a much more economical system, with small low power terminals at the laboratories, and using commercial satellites. We have been experimenting with the system the radio astronomers used² to measure the phase difference between the local oscillators at VLBI stations. In this system pairs of low power CW tones are exchanged between the stations, and the difference frequency between CW tones is recovered to effect the phase or time transfer.

Two satellite ground stations with 3 m antennas and 1 W power have been installed near the NRC time laboratory for operation in the 6/4 GHz commercial satellite band. In the first experiments the Anik A1 satellite of Telesat Canada was used. Tones with the frequencies $f_0 \pm 16$ MHz and $f_0 \pm 16.5$ MHz were transmitted, and the difference frequencies of 32 MHz and 33 MHz recovered. The 1 MHz difference was also recovered to identify the cycle of the 32 MHz. The results of one run are shown in Figure 2. The rms deviation from the 32 MHz signal is 1.07 ns, and the 10 point running average is well within 1 ns.

Recently we have been using a simpler format with only two tones, $f_0 \pm 0.5$ MHz, to recover 1 MHz directly, and I wish to discuss these experiments in more detail.

As Anik A1 has died, we were assigned channel 7, 6185 MHz, on Anik A3. In Figure 3 a simplified diagram of the transmitter is given. All signals are derived from a 5 MHz cesium standard. One reason for showing our transmitter and receiver circuits is to emphasize that commercial up-converters and down-converters that we have tried and have seen advertised have too much residual FM for the very narrow band receivers that we must use to recover the low level signals. These are more than 50 dB down from the signals used in the Symphonie experiment. In the lower left of Figure 3 the spectrum analyzer output of our transmission is shown. The carrier is suppressed by 30 dB from the $1/6 W$ in the ± 0.5 MHz tones, and the strongest overtone is 15 dB down.

In Figure 4 the portion enclosed by the dotted line is a commercial frequency agile receiver, but we found it necessary to synthesize all local oscillator frequencies, to reduce the initial 8 kHz residual FM to a fraction of 1 Hz.

In Figure 5 the details of the narrow band detector are given. The final 5 MHz phase lock loop is a third order loop with an effective bandwidth of about 0.1 Hz.

Typical results, that is, the best results, are given in Figure 6. The rms deviation is 0.4 ns. With 10 s averaging of the phase lock loop, the results are very similar to the 10 s average of the 32 MHz signals obtained earlier.

However, we do experience interference from other satellite or ground services at times. Figure 7, with results from one hour earlier than those in Figure 6, shows such interfering noise, which is certainly to be expected at the very low levels at which we are operating.

Indeed, we understand that in Washington operation in the 6/4 GHz band is impossible, but the 14/12 GHz band is practical and it is available commercially. We will obtain 14/12 GHz hardware for one ground station, and expect about 15 dB improvement from the increased antenna gain.

It is considered that our present 1 MHz experiment was a test of a PRN code system, with none of the bits missing. With commercial geostationary satellites, the station keeping is such that very little Doppler shift is expected, and path delays can be accurately predicted, so that little search capability is required for a PRN code system. We must, at an early date, decide what we want by way of hardware. I understand that Comsat and the USNO are carrying out some modem tests at the moment. Always an optimist, I hope that in a year we can link NRC, USNO and NBS with the beginning of an operational network that could put nanosecond accuracy and subnanosecond precision anywhere in the world served by geostationary commercial satellites, at a cost of \$40,000 per station.

Acknowledgement

We are very much indebted to Telesat Canada for giving us access to the Anik A1 and Anik A3 satellites, and for the advice of their staff in the course of the experiment.

References

- 1 C.C. Costain, J.-S. Boulanger, H. Daams, L.G. Miller, National Research Council, G. Freon, P. Parcelier, Laboratoire Primaire du Temps et des Fréquences, M. Brunet, Centre National d'Etudes Spatiales, J. Azoubib, B. Guinot, Bureau International de l'Heure, "Two-Way Time Transfers Between National Research Council (Ottawa) and Paris Observatory Via the Symphonie Satellite", Proc. 33rd Ann. Symp. on Freq. Control, pp 473-476 (1979).
- 2 S.H. Knowles, W.B. Waltman, J.L. Yen, J. Galt, D.N. Fort, W.H. Cannon, D. Davidson, W. Petrachenko, and J. Popelar, "A phase coherent link via synchronous satellite developed for very long baseline radio interferometry", Radio Sci., 17, 1661-1670, (1982).

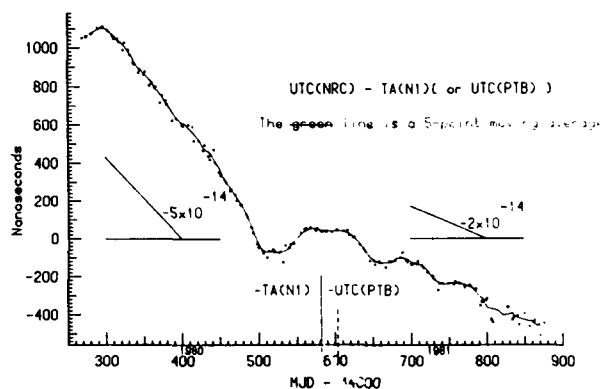


Figure 1.
Comparison Between NCR And PTB Via Symphonie

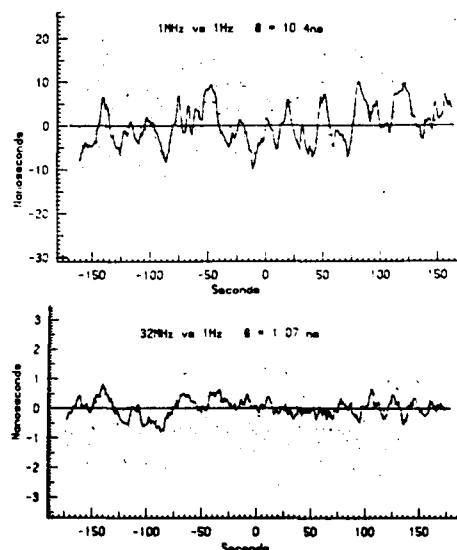
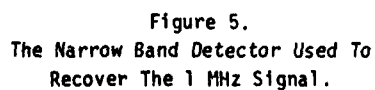
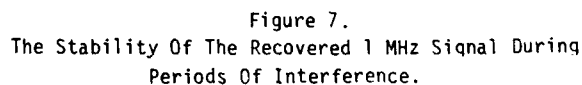
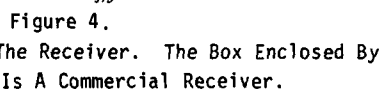
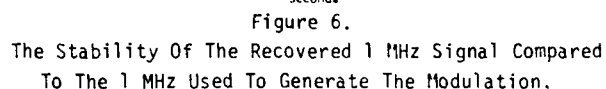
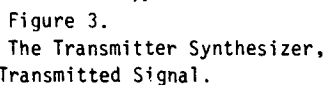


Figure 2.
The 10 Point Running Average And Standard Deviation Of
The Recovered 1 MHz and 32 MHz Signals Via Anik A1.



R. S. Withers and P. V. Wright*
 Lincoln Laboratory, Massachusetts Institute of Technology
 Lexington, Massachusetts 02173

Summary

Transversal filters have been realized as miniature superconductive tapped electromagnetic delay lines. The low loss of superconductors permits the use of a compact stripline structure which offers useful delay at microwave frequencies. Taps are realized as backward-wave couplers or as impedance discontinuities. Chirp filters are formed by cascading such couplers of increasing length. Pulse expansion and compression have been demonstrated over a 2.6-GHz bandwidth. By weighting the taps, side-lobe levels have been reduced to 25 dB below the peak output. A coupled-mode analysis has been employed for predicting the response of this class of filters and has achieved very good agreement with experiment. The analysis also predicts that, if coupling were increased to reduce insertion loss below about 10 dB, significant phase distortion results with a commensurate degradation of side-lobe levels. An algorithm has been developed whereby the tap positions are adjusted so as to predistort the phase response in a way which compensates for the distortion caused by the heavy tap weighting. Subsequent analysis of the predistorted device demonstrates that 41-dB relative side-lobe levels can be achieved in a Hamming-weighted filter with 3 dB insertion loss. The technique is applicable to acoustic and optical grating filters.

Introduction

Transversal filters, also known as Kallman or finite-impulse-response filters, are widely employed as matched filters for analog signal processing. More recently surface-acoustic-wave^{1,2} and acousto-optic³ technologies have been developed and used to fabricate transversal filters with bandwidths as great as 1 GHz.

Linear frequency-modulated (LFM) filters, also known as chirp filters, are commonly constructed as transversal filters and are valuable components in high-performance radar,^{4,5,6} spread-spectrum communications, and spectral analysis systems. These filters are characterized by impulse responses which have an instantaneous frequency which increases or decreases linearly in time (called "up" and "down chirps" respectively), and consequently have a linear group-delay-vs-frequency characteristic.

A technology must offer three essential functions if a transversal filter is to be realized: signal delay, tapping of the signal stream at specified delays and with specified weights, and summation of the tapped signals. We have exploited the low loss of superconductive transmission lines to realize transversal filters as miniature tapped electromagnetic delay lines.^{7,8} Electromagnetic delay lines offer inherently wide bandwidth (up to a few tens of gigahertz) but only several nanoseconds of delay per meter. To achieve reasonable delays (10-1000 ns) in a small volume, the transmission lines must be reduced in size. This results in an increased transmission loss and precludes the use of normal conductors.

Employing microfabrication techniques to form miniature niobium striplines on low-loss dielectric

substrates, chirp filters with bandwidths as great as 3.6 GHz and dispersion as great as 64 ns have been produced in a compact planar form. In this paper are presented the basic structure of these tapped delay lines, a brief discussion of their analysis, synthesis, and fabrication, a presentation of achieved device performance, and a new technique for the design of high-performance filters with low insertion loss.

Device Design, Modeling, and Fabrication

Design

Stripline is used because of its superior isolation, delay, and dispersion characteristics compared with microstrip or coplanar constructions of equal substrate area. The stripline pattern is defined in a single layer of superconductive thin film surrounded by two layers of low-loss dielectric, both of which are coated with superconductive ground planes. Two distinct coupling configurations, labeled proximity-tapped and reflectively tapped, have been exploited.

Proximity-Tapped Delay Lines (PTDL). In this configuration, the input signal is launched as a forward-propagating wave on one of a pair of coupled lines. At specified points along the line pair, a fraction of the signal energy is coupled onto a backward-propagating wave on the second line. Backward-wave couplers¹⁰ are employed to effect this coupling. The strength of each coupler is determined by the spacing of the lines and its peak-response frequency is determined by the line length between transitions in line spacing.

Suppose that an array of these couplers were cascaded end-to-end as shown in Figure 1(a). Suppose further that the length of the couplers were a linear function of length along the line. The resulting structure has a local resonant frequency which is a linear function of delay. In other words, it is a chirp filter with a linear group-delay-vs-frequency (or quadratic phase) relationship.¹¹ The strength of the couplers is controlled by varying the line spacing s to give a desired amplitude weighting to the response.

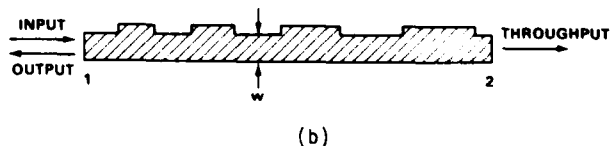
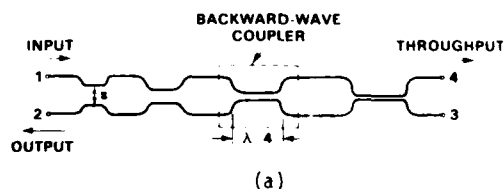


Figure 1(a) Chirp filter formed by cascading backward-wave couplers. (b) Reflectively tapped chirp filter.

In the time domain, each coupler represents two taps, one with a positive impulse response and one with a negative, corresponding to the two transitions in

**This work supported by the Department of the Army.

*Present address: RF Monolithics, 4441 Sigma Rd.,
 Dallas, TX 75234

AD-A136 673

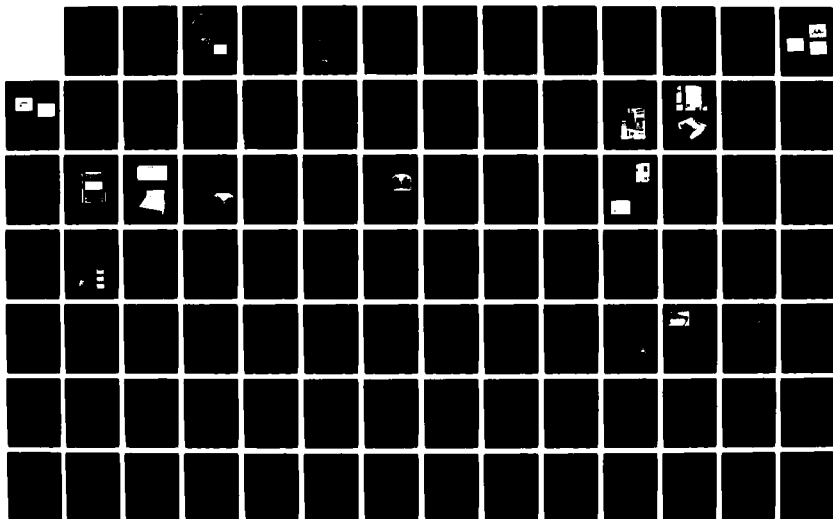
PROCEEDINGS OF THE ANNUAL SYMPOSIUM ON FREQUENCY
CONTROL (37TH) 1-3 JUNE 1..(U) ARMY ELECTRONICS
RESEARCH AND DEVELOPMENT COMMAND FORT MONMOUTH 1983

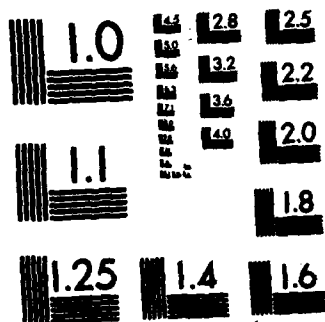
2/6

UNCLASSIFIED

F/G 9/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

line spacing s . The realization of a canonical tapped delay line as a cascade of backward-wave couplers is straightforward. There is a direct analogy between this stripline filter and normal-incidence acoustic or optical grating filters.

It is noteworthy that both ends of the line pair are accessible. Thus, in the case of a chirp filter, both up-chirp and down-chirp filter performance may be obtained from a single device. In general, for lightly tapped lines (≥ 10 dB insertion loss), for which multiple internal reflections are negligible, the impulse response from one terminal pair is the time-reverse of that from the other. The device is thus its own matched filter, a convenient advantage for some applications.

Reflectively Tapped Delay Lines (RTDL). Taps can also be effected by steps in line characteristic impedance from a reference impedance Z_r to a tap impedance Z_t and back. In this case, coupling occurs between forward- and backward-propagating waves on a single line. The impedance changes are most easily produced by changing the line width w (Figure 1(b)). The tap positions along the line are the same as those along the line pair of the PTDL, and weighting can easily be achieved by varying Z_t .

The advantage of the RTDL is that a single line, not a line pair, is used; it thus offers twice the dispersion of a PTDL on the same substrate. It is also immune to the directivity problems of proximity couplers on anisotropic substrates.^{8,12} On the other hand, the RTDL is much more susceptible to reflection-producing impedance discontinuities such as those caused by photolithographic defects, substrate thickness variations, and voids between the dielectric layers. The spurious reflections are directly coupled into the output of the RTDL but are to first order isolated from the PTDL output.

Modeling

From the previous section one can see how a chirp filter with prescribed tap positions and weights is constructed. It is also useful to determine the expected frequency-domain characteristics of a single device and the performance of a pair of devices as matched filters.

Coupling-of-modes theory¹³ has been applied to analyze these devices. The coupled-mode equations are

$$\frac{dR}{dz} = K(z) S(z) \quad (1a)$$

and

$$\frac{dS}{dz} = K^*(z) R(z) \quad (1b)$$

where R is the complex amplitude of the forward-propagating wave input at port 1, S is the amplitude of the backward-propagating output wave, z denotes length along the grating, K is the complex coupling factor, including the dephasing between the input wave at frequency f and the locally resonant frequency of the grating, and K^* is its complex conjugate.

For a linear chirp filter, the dephasing (angle of K) is a linear function of the input frequency and a quadratic function of z . In the weak-coupled limit, in which the input wave amplitude R remains constant in z , and for a flat amplitude weighting (constant $|K|$) Eq. (1b) may be expressed as a Fresnel integral, for which series expansions and tabulations are available.¹⁴

In the general case of arbitrary amplitude weighting and coupling strength, Eq. (1a) and (1b) must be integrated numerically. This is facilitated by the transformation of the two equations to a single Riccati equation, a nonlinear first-order equation. As the variations of R , S , and $|K|$ are gradual on the scale of changes in the dephasing term, it is possible to approximate this equation by Fresnel integrals in a piecewise fashion.

Computations have been performed in this manner to model the superconductive chirp filters, although the model is just as applicable to normal-incidence acoustic and optical gratings. It is noteworthy that phase and amplitude distortions caused by strong coupling are included to arbitrary accuracy.

By integrating Eqs. (1a) and (1b) at a set of frequencies across the band of interest, the transfer function of the device (amplitude and phase vs frequency) is obtained. Such calculations are compared with experimental results later in this paper.

Fabrication

From the specified device parameters (center frequency, bandwidth, insertion loss, and amplitude weighting function), a synthesis routine is used to generate the desired tap positions, tap weights, and the resulting line separations or widths. The proximity-tapped delay line shown schematically in Figure 1(a) would be difficult to fabricate, being typically 1-mm wide and 3-m long. To utilize round substrates with a minimum number of small-radius bends, the delay line is wound up into a quadruple-spiral configuration.

Computer-aided design techniques are used to accurately approximate the spiral delay line with several thousand rectangles for subsequent photomask generation. Reflectively tapped lines are designed by similar techniques and laid out as a double spiral.

Niobium films are deposited on the dielectric wafers to a thickness of 3000 Å by RF sputtering and patterned by reactive ion etching.¹⁵

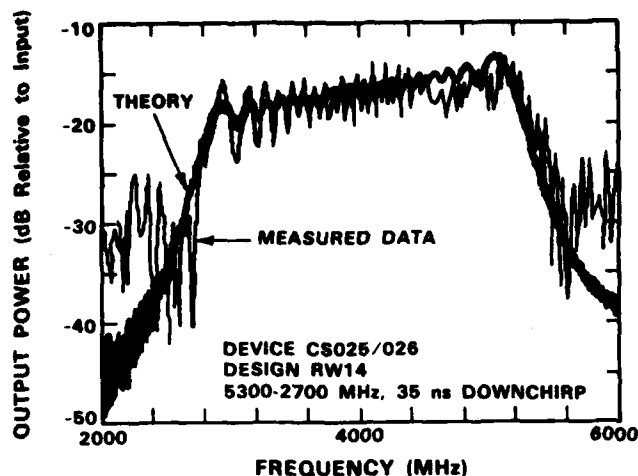
In the package, the second wafer with its ground plane is held in place against the patterned wafer by an array of springs, preventing air gaps which would cause impedance discontinuities. The RF terminals, brought to bonding pads at the wafer edge, are connected with short strips of 0.5-mm-wide ribbon to SMA connectors brought through the back of the package.

Performance of Proximity-Tapped Delay Lines

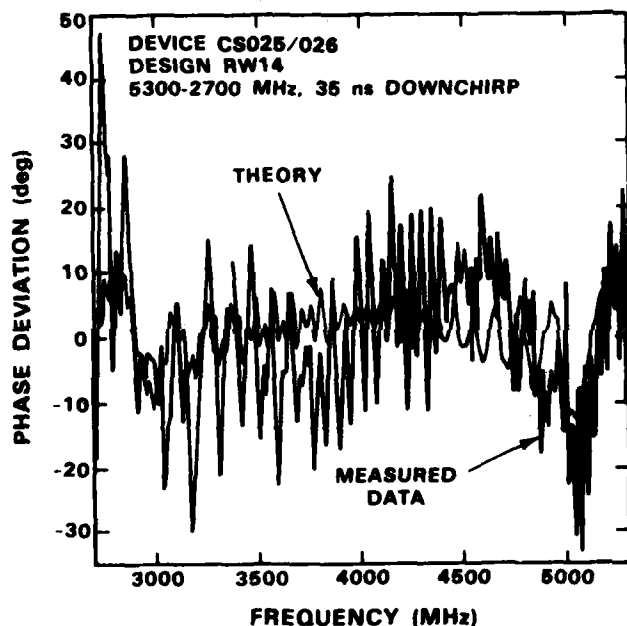
Chirp filters have been fabricated on 5-cm-diameter, 125- μ m-thick sapphire. One particular delay line pattern was designed to have a dispersion of 35 ns and a bandwidth of 2.6 GHz centered on 4.0 GHz.⁸ The design is for a predicted tap-pair coupling coefficient of 10^{-2} , i.e., all couplers are 40 dB in strength. (Such a device with all taps of the same strength is called "linear-weighted". The increasing number of taps per unit length as the high-frequency end of the device is approached results in an insertion loss which decreases linearly with frequency.)

Figure 2(a) shows the measured and predicted amplitude of the down-chirp pair. The predictions were made with the theory described previously and are based on the physical design parameters of the device. Except for the fine-scale deviations (attributed to spurious reflections), the agreement across the

passband (2.7 to 5.3 GHz) is very good, theory and experiment being in agreement within about 2 dB.



(a)



(b)

Figure 2(a) Predicted and measured amplitude of a proximity-tapped superconductive delay line; (b) predicted and measured deviation from quadratic phase for this device.

The measured phase spans a range of more than 25000 degrees across the passband and is essentially a quadratic function, as expected. The measured and theoretical deviations from a best-fit quadratic are plotted in Figure 2(b). The difference between measured and predicted response is dominated by distortions which probably result from spurious reflections. The rms value of the deviation of the measured response from quadratic phase is 12.00. The theoretical deviation from quadratic phase is dominated by the Fresnel ripple resulting from the asymmetry of the integral near the band edges.

The least-squares-fit quadratic coefficient of phase for the measured data is 2.361×10^{-3} deg/MHz², which implies a chirp slope of 76.24 MHz/ns. The

up-chirp phase was also measured and gave a least-squares value implying a chirp slope of 76.40 MHz/ns, which is excellent agreement and indicates well-matched filters. The design value is 74.2 MHz/ns for both up and down chirp. This error of less than 3% could be caused by a 3% overestimation of the effective dielectric constant of the substrate.

Silicon appears to be an excellent dielectric substrate for this application.¹² Its most desirable property is its dielectric isotropy, which eliminates the limit on time-bandwidth product imposed by anisotropy, and the advantages offered by the profuse technology behind this material cannot be overstated.

Proximity-tapped delay lines have been made on 5-cm-diameter, 125- μ m-thick wafers of 150 Ω -cm resistivity, (111)-cut n-type silicon, utilizing the same niobium technology described previously.

One design is for a Hamming-weighted chirp filter with a 2322-MHz bandwidth centered at 4 GHz, dispersion of 37.5 ns, and an insertion loss at center frequency of 10 dB.⁷ The results of pulse-compression testing of this device are shown in Figure 3. A 10-V, 135-ps impulse is applied to the up-chirp end of the device (port 4 in Figure 1(a)). The resulting expanded pulse (obtained at port 3) is amplified, time-gated, delayed, and applied to the down-chirp end of the filter (port 1). The matched filter response (obtained at port 2) is shown in Figure 3. This compressed pulse exhibits 25-dB relative side-lobe levels and a central lobe full-width of 2.2 ns. This width is somewhat wider than the 1.3 ns of a Hamming-weighted pulse of this bandwidth because the signal was weighted upon both expansion and compression. The theoretically predicted side-lobe level of 41 dB is not achieved because of substrate thickness and linewidth variations which produce spurious reflections and deviations of coupling from design values.

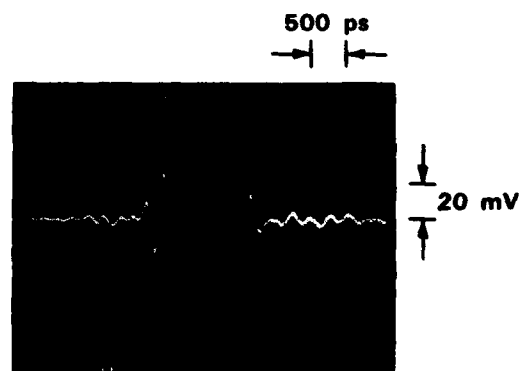


Figure 3 Measured compressed pulse response of a Hamming-weighted niobium-on-silicon proximity-tapped delay line.

Phase Compensation for Low-Insertion-Loss Filters

Error in Uncompensated Filters

Tapped-delay-line filters are most straightforwardly designed in the weakly tapped limit. In this limit, which is valid for insertion losses of 10 dB and higher, the input signal propagates essentially undepleted and, as a corollary, multiple internal reflections are negligible. It is noteworthy

that it is not the weight of the individual taps, but the product of weight and number of taps contributing at a given frequency, which determines the validity of the weakly tapped assumptions. In this limit there is a direct mapping of individual tap weights and positions along the line to the amplitude and time-dependence of the impulse response. In the case of a linear-chirp filter, the tap spatial periodicity is a linear function of distance along the line or, equivalently, the tap position is a quadratic function of tap number. In the frequency domain, the line has a local resonant frequency which is a linear function of position and, consequently, a linear group-delay-vs.-frequency characteristic.

For more tightly coupled lines (≤ 10 dB insertion loss), distortions in both amplitude and phase occur. The amplitude of the response at a given frequency is no longer proportional to the product of the tap weight and the number of taps resonant at that frequency but to the hyperbolic tangent of this product. This saturation effect can be compensated analytically¹⁶ and has in fact been incorporated into the device analysis and synthesis procedures used to design the filter of Figure 3. Phase errors are more difficult to handle and cannot be treated analytically in closed form for a chirp filter. Heuristically, the effect of tight coupling is to deplete the input wave and thereby shift the effective center of reflection from the exact point of synchronism toward the input, thereby reducing the group delay by an amount dependent on the local tap strength.

In Figure 4 is shown the phase error (deviation from linear) of a pulse-compression system consisting of cascaded up- and down-chirp filters, each with 37.5-ns dispersion and 2.3-GHz bandwidth (the same as the filter of Figure 3) and a 3-dB insertion loss at center frequency. Each filter is weighted with the square root of the Hamming function so that the system response is Hamming weighted.⁵

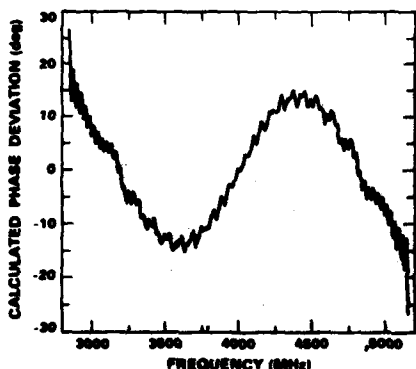


Figure 4 Calculated phase error of a matched pair of chirp filters with 3-dB insertion loss, 37.5-ns dispersion, 2.3-GHz bandwidth, Hamming weighting, and no phase compensation.

The phase error of Figure 4 was calculated using a method somewhat different from the coupling-of-modes solution used to obtain the curves of Figure 2. In anticipation of the need to analyze the nonlinearly chirped phase-compensated gratings, an integral formulation was developed¹⁷ which is in many ways similar to the WKBJ approximation¹⁸ and is much more

efficient numerically than the coupling-of-modes solution.

The phase error for the filter pair of Figure 4 is clearly significant, having a slow variation of $\pm 10^\circ$ across the central two-thirds of the band. The detrimental effects of this error are evident in Figure 5, which displays the calculated compressed pulse envelope of the filter pair. A trailing side lobe of 21 dB relative magnitude is evident. An ideal Hamming-weighted system would yield a peak side-lobe level of 42.8 dB.

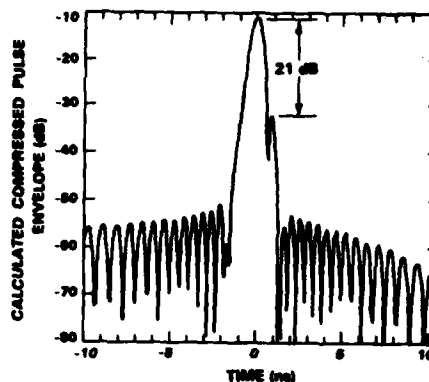


Figure 5 Calculated compressed pulse response envelope of the filter pair of Figure 4.

Predistortion Algorithm

In order to compensate for the phase errors caused by tight coupling, a technique has been developed whereby the tap positions are adjusted (i.e., the design is "predistorted") in a manner based on the calculated phase error (e.g. Figure 4) of the undistorted device. The predistorted filter is then analyzed using the integral (WKBJ) formulation to verify the efficacy of the adjustment.

The phase error of a single uncompensated device, $\phi_e(f)$, is first calculated. This is converted to a group delay error by the familiar relation

$$\tau_e(f) = \frac{1}{2\pi} \frac{d\phi_e(f)}{df} \quad (2)$$

where ϕ_e is expressed in radians.

The tap spacing is uniquely described by the grating frequency function $f_g(z)$ which, for an uncompensated linear-chirp filter, is

$$f_g(z) = f_c + B \frac{z - \frac{L}{2}}{L} \quad (3)$$

where f_c is the center frequency of the device, B is the bandwidth, z is the position along the line measured from the input, and L is the line length. We wish to perturb this in such a way that the group delay is predistorted by an amount equal and opposite to $\tau_e(f)$. Invoking the space-frequency relationship of the chirp filter, we achieve this to first order by setting

$$f_g(z) = f_c + B \frac{z - \frac{L}{2}}{L} - \frac{B}{f_c} \tau_e(f) \quad (4)$$

$$f = f_c + B \frac{z - \frac{L}{2}}{L}$$

where T is the device dispersion. Of course, $T = 2L/v_p$ where v_p is the phase velocity.

The phase error $\phi_e(f)$ is obtained by calculation at a number of discrete frequency points. For use in the predistortion algorithm, Eq. (4), it is convenient to have an analytic form. One useful form is the Fourier sine series

$$\phi_e(f) = \sum_{n=1}^k A_n \sin 2\pi \left(\frac{f - f_c}{B} \right) \quad (5)$$

This series does not include the even (in frequency around f_c) components of the phase error but, as is evident in Figure 4, the even errors of the up- and down-chirp filters cancel.

Phase Error in Compensated Filters

The predistortion algorithm (Eq. (4)) was applied to the devices of Figure (4) and (5), using the series (Eq. (5)) with 8 terms to approximate the phase error. The compensated devices were then analyzed; the resulting system phase error is shown in Figure 6. This clearly demonstrates an order-of-magnitude improvement over the uncompensated device.

The compressed pulse response of the compensated system is shown in Figure 7. The relative side-lobe level is 41 dB, within 2 dB of the ideal.

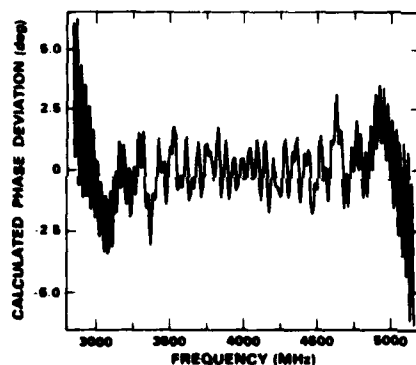


Figure 6 Calculated phase error of the matched pair of chirp filters of Figure 4 after phase compensation.

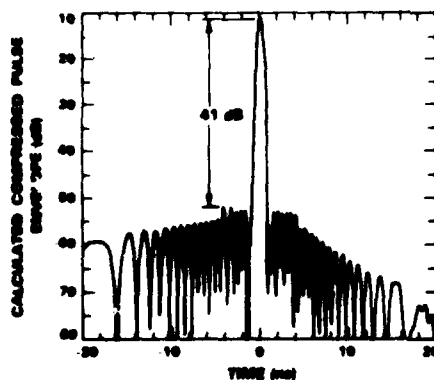


Figure 7 Calculated compressed pulse response of the compensated filter pair of Figure 6.

Application

The compensated grating function (Eq. (4)) can be readily implemented in the pattern-generating routine used to fabricate proximity-tapped and reflectively tapped delay lines. The logical candidate for strong tapping and compensation is the RTDL, because of its high tap-weight-independent spurious reflection levels.

It should be noted that a compensated filter cannot be used as both up- and down-chirp filters, as can an uncompensated filter. The compensated device is not its own matched filter; two separate devices are needed.

An uncompensated RTDL with 64-ns dispersion, 3.6-GHz bandwidth, and 3-dB insertion loss has been fabricated and tested. However, spurious reflection levels are still too high to discern the predicted 21-dB side lobe. Improvement of substrate uniformity and device packaging is necessary to benefit from phase compensation of this class of devices.

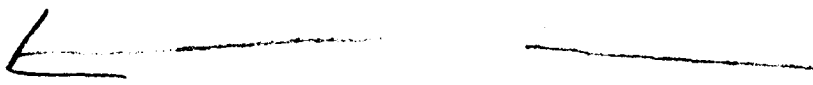
The predistortion algorithm is directly applicable to normal-incidence acoustic and optical grating filters and could be adapted to oblique-incidence filters.

Acknowledgements

We wish to acknowledge the following contributors: A. C. Anderson, J. B. Green, and S. A. Reible for technical collaboration; R. W. Ralston and E. Stern for technical guidance; J. H. Holtham for scientific programming and, in collaboration with G. Fitch, the spiral mask designs; W. Macropoulos, R. L. Slattery, C. M. Vanaria, J. Hamer, E. M. Macedo, and P. R. Phinney for fabrication and packaging; S. S. Cupola, W. C. Kernan and G. V. O'Malley for RF testing and cryogenic support; and S. Lavache for manuscript preparation.

REFERENCES

1. Special Issue on Surface-Acoustic-Wave Device Applications, IEEE Trans. Sonics Ultrason., vol. SU-28, no. 3, pp. 117-228, May 1981.
2. Special Issue on Surface-Acoustic-Wave Devices, Proc. IEEE, vol. 64, no. 5, May 1976.
3. Devices and Systems for Optical Signal Processing, vol. 218. Bellingham, WA: SPIE, 1980.
4. J. R. Klauder, A. C. Price, S. Darlington and W. J. Albersheim, "The Theory and Design of Chirp Radars," Bell Sys. Tech. J., vol. 39, no. 4, pp. 745-808, July 1960.
5. C. E. Cook and M. Bernfeld, Radar Signals. New York: Academic Press, 1967.
6. R. C. Williamson, "Properties and Applications of Reflective-Array Devices," Proc. IEEE, vol. 64, no. 5, pp. 702-710, May 1976.
7. J. T. Lynch, R. S. Withers, A. C. Anderson, P. V. Wright and S. A. Reible, "Multi-GHz-Bandwidth Linear-FM Filters Using Superconductive Stripline," to be pub., Appl. Phys. Lett., 1983.
8. R. S. Withers, A. C. Anderson, P. V. Wright, and S. A. Reible, "Superconductive Tapped Delay Lines for Microwave Analog Signal Processing," IEEE Trans. Magn., vol. MAG-19, no. 3, pp. 480-484, May 1983.

9. R. L. Kautz, "Miniaturization of Normal-State and Superconducting Striplines," J. Res. Nat. Bur. Stand., vol. 84, no. 3, p. 247, May 1979.
 10. B. M. Oliver, "Directional Electromagnetic Couplers," Proc. IRE, vol. 42, no. 11, pp. 1686-1692, Nov. 1954.
 11. B. Loesch, E. M. Hofstetter and J. P. Perry, "A Technique for Synthesizing Signals and Their Matched Filters," Tech. Rept. No. ESD-TR-69-414. Lexington, MA: Massachusetts Institute of Technology, Lincoln Laboratory, 1969.
 12. A. C. Anderson, R. S. Withers and S. A. Reible, "Substrates for Superconductive Analog Signal Processing," IEEE Trans. Magn., vol. MAG-19, no. 3, pp. 485-489, May 1983.
 13. P. V. Wright and H. A. Haus, "A Closed-form Analysis of Reflective-Array Gratings," in 1980 Ultrasonics Symposium Proceedings. New York: IEEE, 1980, pp. 282-287.
 14. M. Abramowitz and I. A. Stegun, Handbook of Mathematical Functions. Washington, DC: National Bureau of Standards, 1972.
 15. S. A. Reible, "Reactive Ion Etching in the Fabrication of Niobium Tunnel Junctions," IEEE Trans. Magn., vol. MAG-17, no. 1, pp. 303-306, Jan. 1979.
 16. P. V. Wright, "Weighting Synthesis to Account for Tight Coupling," unpublished, May 1982.
 17. P. V. Wright and R. S. Withers, "Phase Compensation Technique for Strongly Reflecting Gratings," unpublished, 1983.
 18. P. M. Morse and H. Feshbach, Methods of Theoretical Physics. New York: McGraw Hill, 1953, pp. 1092-1105.
- 

LOW PHASE NOISE MULTIPLE FREQUENCY MICROWAVE SOURCE*

J.T. Haynes, H. Salvo, R.A. Moore
Systems Development Division
Westinghouse Defense and Electronics Center
Baltimore, Maryland 21203

B.R. McAvey
Research and Development Laboratories
Westinghouse Electric Corporation
Pittsburgh, Pennsylvania 15235

ABSTRACT

Use of high overtone bulk acoustic resonators (HBAR) has made possible development of low phase noise multiple frequency microwave sources operating directly at microwave frequencies. Most recently a low noise source has been developed at L-band which provides signals at approximately 5-MHz intervals and achieves phase noise suppression equivalent to sources based on low frequency quartz crystal stabilization and multiplication. The direct microwave HBAR multiple frequency source requires a fraction of the hardware required to achieve the same phase noise suppression for multiple microwave frequency operation by any other means.

Stabilization is based on exploiting the high overtone resonance of crystals such as YAG, sapphire, lithium niobate or lithium tantalate for which intrinsic loss indicates a potential Q approximately ten times that of quartz. Resonators of this type have been fabricated as high as 10 GHz. For the 1.5-GHz band compressional mode, loaded Q's of over 50,000 have been achieved in several samples. The resonators consist of deposited film transducers on the resonator crystal. Single reverberation is confined to the region under the transducer. The crystal is mounted rigidly to its housing which minimizes effects of external vibration on frequency stability. Measured vibration sensitivity is described in a companion paper.

The source consists of a low noise voltage controlled oscillator (VCO) stabilized in an automatic frequency control (AFC) loop. The HBAR is the frequency determining element in the AFC discriminator. Measured phase noise agrees well with predicted performance. Details of noise performance will be provided in the paper. Multiple frequency operation can easily be obtained with the addition of digital control circuits to preposition the VCO on any of the HBAR's responses. Thus, a low phase noise electronically controlled multiple frequency source is obtained with minimal hardware.

INTRODUCTION

The high overtone bulk acoustic resonator (HBAR) provides the basis for low noise oscillators stabilized directly at microwave frequencies. The HBAR's high Q, closely spaced, periodic resonances provide stabilization for multiple frequency oscillators. The phase noise performance is equivalent to sources phase locked to low frequency crystal standards, but requires only a fraction of the hardware.

The HBAR has been described in previous papers by Moore, et al.^{1,2} so that a description of its operation will be limited to briefly reviewing its dominant mode structure as shown in figure 1. The HBAR might be considered as an acoustic analog of the optical Fabry-Perot interferometer. Geometrically it is more like present day microwave delay

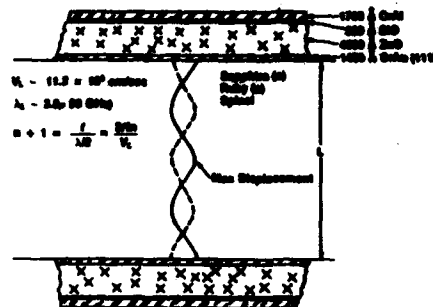


Figure 1. Geometry of HBAR Showing Substrate and Transducer Detail

lines than other acoustic resonators. A transmission resonator is illustrated in figure 1 which is the form we have found most useful. The main feature of the geometry, the separate transducer coupling to the resonator body, provides the following key operational advantages:

- Coupling to high order overtones
- Availability of higher Q nonpiezoelectric substrates
- Confinement of acoustic activity allowing rigid mounting and lower sensitivity to vibration.

The coupling to high order overtones allows useful resonator coupling directly at microwave frequencies and provides a large number of

equally spaced overtones in the operating band. We are using overtone levels of several hundred (as opposed to up to the fifth or seventh with traditional quartz resonators) which, for the resonator described in this paper provides an overtone spacing of approximately 5 MHz. Even closer overtone spacing could be provided but would limit allowable bandwidth for the oscillator feedback circuit and limit noise performance. The rigid mounting of the resonator has led to almost two orders of magnitude suppression in vibration sensitivity which is described in a paper in this conference by Rossman and Haynes.³

The relationship of Q of the HBAR to other resonators is illustrated in figure 2 in terms of FQ product as a function of frequency. Because Q's of most resonators fall off as reciprocal frequency, FQ product serves as a good first order basis of comparison. The sloping lines on figure 2 are of constant FQ product. Quartz is characterized most closely by $FQ = 10^{11}$ with examples indicated slightly above this value for frequencies up to approximately 200 MHz. By contrast above 1 GHz the HBAR Q's achieved are well above the $FQ=10^{11}$ line with examples of almost $9(10)^{11}$ at L-band and $5(10)^{12}$ at X-band. In this paper we are reporting on the resonator with Q's over 50,000 in the 1.5- to 2-GHz frequency band.

* Supported in part by contract F19628-82-C-0066, RADC/ESD, Hanscom AFB, Mass. 01731

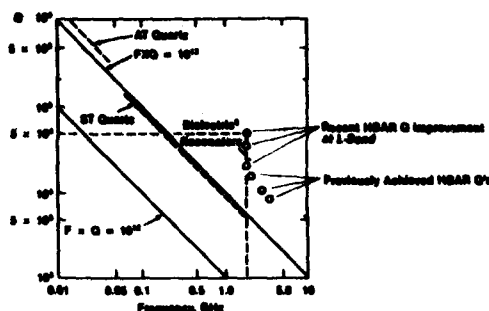


Figure 2. FQ Product Chart Showing Relative Q's of Recently Fabricated HBAR's To Both Crystal Resonators and Dielectric Resonators

The advantageous use of higher Q media is best illustrated in figure 3 which illustrates the available Q's of materials for the frequency range 1 to 10 GHz. It should be noted that there are several materials with Q's approximately an order of magnitude greater than quartz (SiO_2). Not only does the separate transduction allow nonpiezoelectric substrates, it also allows the use of a nonpiezoelectric crystalline orientation of piezoelectric materials. Greatest success in the present program has been with cube edge and diagonal cut YAG.

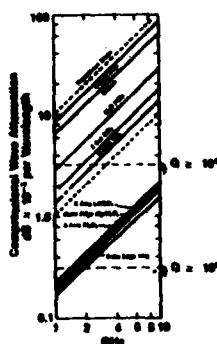


Figure 3. Available Q of Typical Materials

THE RESONATOR

The most recent results and highest Q's being reported on this paper were the result of designs using a transmission line type model of the resonator. The model, based on the Mason Transducer equivalent circuit, takes into account film thicknesses, acoustic impedances, layer propagation attenuation, and aperture size. This model does not include effects of the diffraction of the acoustic wave and therefore predicts two values of the aperture which give the same values of insertion loss and Q, each at a different frequency. This is basically a statement that the resonator can be uncoupled both by making the aperture too large or too small. This symmetry is broken in the actual device by diffraction.

To investigate this effect, resonators are identical except for aperture size where fabricated on the same substrate. The transducers for these experiments were simple circular spots having apertures of $2.0 \times 10^{-8} \text{ m}^2$ to $2.0 \times 10^{-5} \text{ m}^2$ in 14 logarithmic steps. Experimental resonators on (1,0,0) YAG and (1,1,1) YAG have been fabricated and tested. Results indicate that the Q's increase with increasing aperture size, however, the insertion loss is also increasing. By the time the aperture has reached $3.55 \times 10^{-7} \text{ m}^2$ the insertion loss is so large that resonance peaks are not observable. The insertion loss vs aperture size curve, shown as figure 4 for (1,0,0) YAG and figure 5 for (1,1,1) YAG, exhibits a minimum near 3.55×10^{-8} for the set of film thicknesses used.

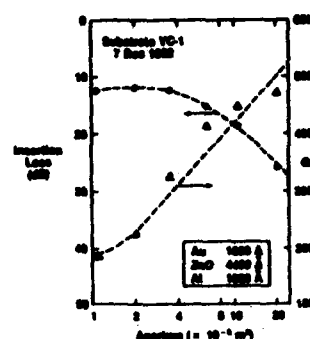


Figure 4. Insertion Loss and Q vs Aperture for (1,0,0) YAG

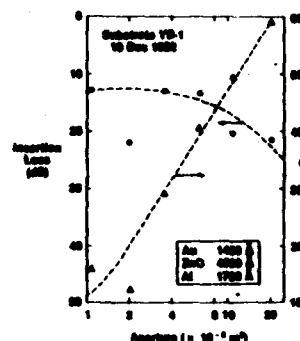


Figure 5. Insertion Loss and Q vs Aperture for (1,1,1) YAG

Additionally comparing figures 4 and 5, the experiment indicates that (1,1,1) YAG may be a better substrate material than (1,0,0) YAG. The resonators fabricated on (1,1,1) YAG exhibit lower insertion loss for the same Q as the (1,0,0) YAG. This is surprising in that (1,0,0) YAG is the lower loss material. It is theorized that the anisotropy of the (1,1,1) direction is aiding the resonance.

Based on the experiment, we have selected for the resonator design a circular spot transducer with an aperture of $6.3 \times 10^{-8} \text{ m}^2$ which provides the best tradeoff between lowest insertion loss and highest Q. The piezoelectric layer is zinc oxide 4500Å thick. The bottom contact is chrome-gold 1200Å thick and the top contact is chrome-aluminum 1800Å thick.

The substrates are prepared from a boule of laser quality YAG. After X-ray alignment, the boule is sliced into substrates of thickness appropriate for the required peak frequency spacing. The faces are polished using a technique believed to be unique in the industry. The process provides exceptional results for hard, lower loss, media for use at higher microwave frequencies.

The transducers are fabricated using standard integrated circuit photolithography. The bottom contact of the active transducer aperture is prepared by the flash evaporation of chromium-gold. The chromium adhesion layer of 50Å to 100Å is evaporated first with a short interval of coevaporation with the gold. This procedure is known to produce ordered gold films.

The zinc oxide is deposited using an RF diode sputtering system. Sputtering is from a zinc oxide target in a 25% O_2 to 75% Ar atmosphere. The thickness is monitored using an interferometric laser monitor. The top electrode is deposited using a rejection process. After coating the substrate with resist and developing the pattern, a chromium-aluminum layer is evaporated on the surface. 200Å of chromium is used and the thickness of the aluminum is determined by the trans-

ducer design. Using acetone and an ultrasonic cleaner the unwanted chromium-aluminum is removed.

The high-overtone bulk acoustic resonator chip is shown at the top of the exploded view of the mounting scheme in figure 6. After pretesting to ensure the required Q and insertion loss, the resonator substrate is cut to 0.125 inch by 0.200 inch. The YAG substrates are about 0.034-inch thick. The ground planes of the top and bottom transducers extend to the outer edges of the substrate as shown in the figure. The transducer top contact of aluminum is connected to the package terminals by epoxy reinforced gold wire ball bonds. The substrate is bonded to a supporting metal preform using conductive epoxy which in turn is epoxied to the grounding base of the resonator package. The top ground plane of the substrate is grounded by wraparound tabs shown in position in the figure. The gold wire contact leads are epoxied to the microstrip line terminals of the package. The preform epoxy mounting provides a robust support for the resonator body.

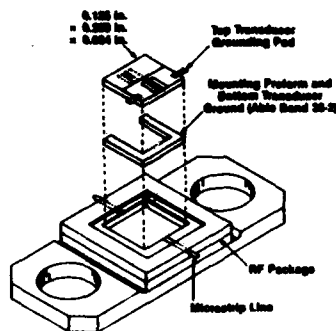


Figure 6. High Q Bulk Mode Resonator

A typical measure resonator response is shown in figure 7. To date the best resonators exhibit Q of greater than 50,000 with insertion losses of less than 10 dB. Some resonators with higher Q's, near 65,000, have been demonstrated by the increased insertion loss.

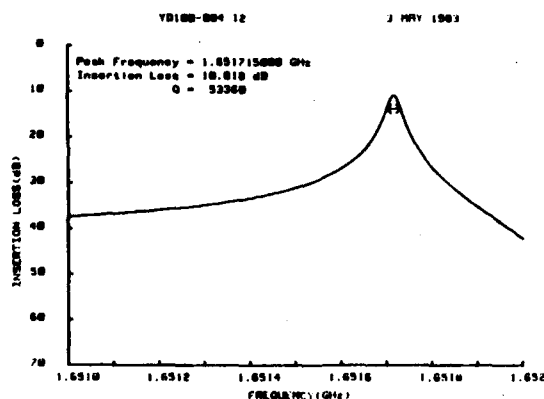


Figure 7. Typical High Overtone Bulk Acoustic Resonator Response

SOURCE STABILIZATION

To fully exploit the HBAR's unique characteristics, an automatic frequency control (AFC) loop technique is used as shown in figure 8. The resonator is used as the frequency determining element of the frequency discriminator. The output of a low noise voltage controlled

oscillator (VCO) is applied to the discriminator. At the discriminator output — mixer output — an error voltage appears that is proportional to the frequency error between the VCO and the particular selected HBAR resonance. This voltage is then fed back to the VCO through a loop video amplifier. To take full advantage of the resonator's filtering properties the output is taken after the resonator.

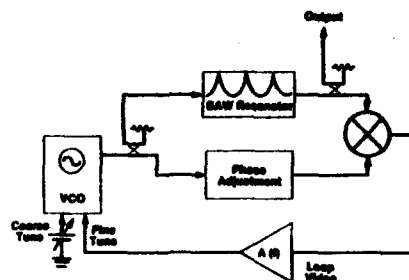


Figure 8. Simplified Diagram of AFC Stabilized Source

Multiple frequency operation is achieved relatively simply with the addition of hardware to allow the loop to lock on any of the HBAR's responses within the VCO's tuning band. This implementation is shown in figure 9. The principle addition is a digital controller which accepts a user input and addresses a digital-to-analog converter to preposition the VCO on the proper HBAR response. I and Q mixers provide amplitude and phase data to the controller so that it may adjust the phase shifter for quadrature signals at the mixers at each frequency. Once positioned the loop locks through the loop video amplifier as previously described.

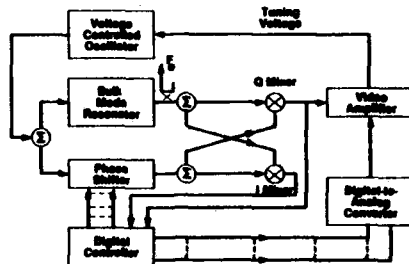


Figure 9. Diagram of Multiple Frequency Source

PHASE NOISE PERFORMANCE

Short term stability is limited by resonator Q, video amplifier noise, and VCO AM noise at modulation rates below a few kilohertz. Above this, phase noise is a function of VCO phase noise as modified by loop gains and resonator filtering.

Quantitatively the video amplifier contribution is:

$$L(f) = \frac{F_0^2 e_0^2}{8Q^2 E^2 f_A^2} \left[\left(\frac{f_A}{f} \right)^2 + \left(\frac{f}{f_A} \right)^2 \right]$$

- Where:
- $L(f)$ = Single sideband phase noise (1/Hz)
 - $\frac{F_0^2}{e_0^2}$ = Output frequency
 - e_0^2 = Video amplifier noise voltage floor (V^2/Hz)
 - Q = Resonator Q
 - E = Mixer phase sensitivity (V/Rad)
 - f_A = Video amplifier 1/f noise break frequency
 - f = Modulation frequency

AM noise on the VCO and its power amplifier is converted to phase noise by the mixer LO port acting as an inefficient AM detector. The contribution to phase noise is:

$$L(f) = \frac{F_o^2 T^2 M(f_o)}{2Q^2 f^3}$$

Where: 1 = Fractional mixer video voltage change per fractional LO power change

$M(f_o)$ = 1/f AM noise spectral density at frequency f_o (1/Hz)

Above a few kilohertz phase noise performance is determined by first calculating the suppression of VCO open loop phase noise by the AFC loop. Noise is then filtered further by taking the output after the resonator. This effect along with the above mentioned contributions are displayed in figure 10. Parameters associated with the L-band source under development were used.

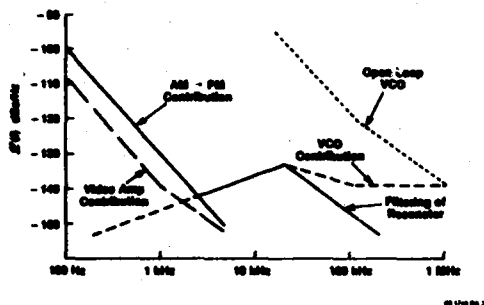


Figure 10. Predicted L-Band Noise Performance

Thus far a single frequency source at L-band has been fabricated and tested. Measured performance not only agrees reasonably well with predicted performance but compares favorably with typical L-band radar requirements. Figure 11 shows measured and predicted phase noise along with a typical requirement. The test set floor is included for analysis.

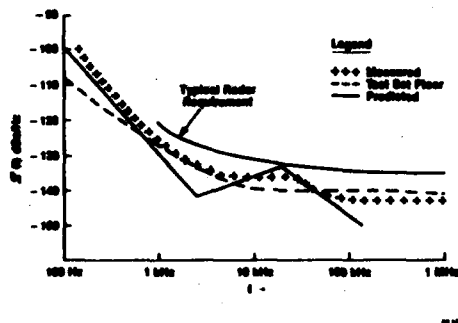


Figure 11. Measured L-Band Phase Noise Performance

In the 100 Hz to 1 kHz region one should note that although there is an approximate 6-dB discrepancy, the predicted and measured curves are both proportional to $1/f^3$. In this region VCO AM noise conversion to phase noise dominates and is proportional to $1/f^3$. VCO AM noise generally is difficult to measure and the prediction is based on a typical value which might reasonably be worse by 6 dB. In the 1 kHz to 4 kHz region the measurement was test set limited as it was above 50 kHz. The measured data does begin to follow the filter response of the resonator above 20 kHz. The test set floor is a measurement of two identical oscillators in the test set. Measurements 3 dB below this floor, in the region above 100 kHz, indicate that the HBAR stabilized source is at least 10 dB below the noise level of a single test set oscillator. It should be noted that in order to use an existing noise test set, the frequency of the HBAR source was doubled.

CONCLUSIONS

A source making use of the high overtone bulk acoustic resonator has been described. Resonators achieving Q's over 50,000 at 1.5 to 2 GHz, greater Q than by any other means at those frequencies, were described. A source with noise performance equivalent to those for which the frequency is multiplied from low frequency crystal reference oscillators is reported. Much lower hardware levels are required. With improvements in HBAR technology, even superior performance may be possible.

REFERENCES

1. R.A. Moore, J.T. Haynes, B.R. McAvoy, "High Overtone Bulk Resonator Stabilized Microwave Sources," IEEE Ultrasonics Symposium, Cat # 81CH1689-9, Vol., pp 414-424, 1981.
2. R.A. Moore, J. Gosdell, A. Zahorchak, R.A. Sundelin, W. Hopwood, J.T. Haynes, B.R. McAvoy, J. Murphy, "High Q Bulk Acoustic Resonators for Direct Microwave Oscillator Stabilization," 34th Annual Frequency Control Symposium, Ft. Monmouth, N.J., 07703, pp 243-251, May 1980.
3. H. Rossman, J.T. Haynes, "Determination of Acceleration Sensitivity of Bulk Mode Resonator Plate," 37th Annual Frequency Control Symposium.

ACKNOWLEDGEMENT

The authors wish to acknowledge D.C. Cooke and M.S. Buchalter for their aid in measuring phase noise.

SYNCHRONIZATION AND TRACKING WITH SYNCHRONOUS OSCILLATORS

Vanki Ramagala
Technical Staff Specialist
Fairchild Communications & Electronics Company
Germantown, Maryland 20874

Summary

A Synchronous Oscillator (SO) is a coherent synchronization and tracking network in which an external signal forces the oscillator to be synchronized in frequency with a final phase difference determined by the initial frequency difference between the oscillator natural frequency and the external frequency. A SO acquires, tracks, filters, divides and improves the signal-to-noise ratio of the input signal in a single process. Most of the analyses in the literature, since the original work by Van der Pol, are based upon forced oscillation (FO) principle. In this operation the oscillator ceases to oscillate as soon as the external signal is applied to it.¹⁻⁴

However, in the SO, the oscillations continue to exist after the application of the external signal. A SO has a very high skirt selectivity, approaching those of crystal filters ($>60\text{dB/OCT}$) and a linear phase within its tracking band. Also, a SO can be synchronized by an input signal-to-noise as low as -10dB , compared to $+5\text{dB}$ and higher in phase-lock loops. A SO constitutes a novel network, not realizable by other means, at least in the same simplicity and quality. It is also a new class Van der Pol oscillator. It is a nonlinear, oscillatory network which is used in linear signal processing, such as filtering, synchronization, tracking and improving the signal-to-noise ratio. SO has also a self-adjusting feature to maintain constant output signal-to-noise ratio even if input signal-to-noise ratio deteriorates.

Introduction

Synchronization and improvements in synchronization techniques have made it possible to advance the state of the art of communications systems and to produce efficient digital coherent communications networks.

The concept of synchronization goes back to the 17th century, when Huygens reported that two clocks slightly out of step with each other became synchronized when hung on a thin wooden board.

The ability to predict the occurrence of certain events offers biological organisms advantages for survival. It is not surprising that a clock seems to exist in almost any organism, ranging from humans and other mammals to insects and plants. There is also evidence that a given organism may contain more than one independent clock, each clock controlling a different function, especially in humans and other highly evolved animals. Seasonal occurrences, and high or low tides caused by the gravity of the moon and the sun can also be attributed to synchronization.

In electronics, when an external frequency f is properly applied to the base of a transistor, oscillating with a frequency f_0 , beats of two frequencies result. When frequency f approaches f_0 , the beats suddenly disappear, and only frequency f remains.

General Considerations

Since the publication of the first paper on synchronization, or frequency of entrainment, by Van der Pol in the late 1920's,² many articles have appeared that analyze the properties and behavior of Van der Pol, or injection oscillators.¹⁻⁴ In spite of the large number of published articles, there is a lack of reporting on applications and practical circuits. There has been not serious attempts in their use as inadequate solutions of nonlinear differential equations did not reveal the real capabilities of Van der Pol oscillators under various operating conditions. When the input energy, the efficiency and the gain of the Van der Pol oscillator are changed an entirely different synchronization and tracking network results. A highly nonlinear and unstable network is used in a linear signal processing.⁸⁻¹⁰ SOs are novel networks.

An SO can be considered a coherent acquisition and tracking network, in which the external signal forces the SO to be synchronized with the external signal in frequency and to have a final phase difference that is determined by the initial difference between the SO natural frequency and the external control frequency. It is well known from the theory of oscillations that there is a negative incremental resistance at the feedback point of the oscillator amplifier, and this negative incremental resistance cancels all the AC losses in the circuit. The effect of the external signal is to increase or decrease the incremental negative resistance, thereby to serve as an energy source to the free-running oscillator, and consequently to establish a new steady-state synchronization condition. In an FO the incremental negative resistance is cancelled and the oscillations cease to exist. In a FO the oscillator becomes an amplifier, whereas in an SO the oscillations continue to exist after the application of the input synchronization signal. This feature constitutes the important difference between the FO and SO. Therefore, in a SO the nonlinearities are preserved.

Phase stability is always associated with amplitude instability. The difference between the linear and nonlinear cases is that the instability is permanent in the linear case, whereas in the nonlinear case, the instability decreases with the increasing amplitude and vanishes for a fixed value of amplitude, at which the oscillation becomes stationary.

An important feature in the operation of an SO is the tracking range as a function of input power and the associated phase differences between the input control frequency and the output frequency. When the input or control signal level drops, the bandwidth of the SO decreases, and as a result, the output signal-to-noise ratio remains fairly constant. The energy or power of the input signal is one of the three parameters that determine the overall tracking range of the SO. The gain of the oscillator active element and the quality factor Q of the tank circuit are also determining factors of the track-in and pull-in ranges of the SO.

An SO is not only a synchronous tracking network but is also a linear-phase bandpass filter, with the bandwidth determined by the input driving signal, as well as by other factors. An SO is also a frequency-dividing network. In a carrier recovery network (CRN), the SO can perform all three of these functions simultaneously and improve the S/N by 30dB, as well.

Another important feature of an SO is its capability for storing input frequency information for several cycles of the input signal when this is turned off. This storage capability is especially useful in clock recovery networks (CLRN), in which the input data consist of continuous ones or zeros for extended time durations wherein no synchronization signal is provided. The SO continues to oscillate at the frequency of the external signal until the time constants of the circuit take over. The operating frequency of SO is limited only by the characteristics of the circuit elements.^{8,9}

The performance of an SO is much more rewarding and sophisticated than a Phase-Lock Loop (PLL). A SO is a multifunctional network. For example in a QPSK modem for carrier recovery a SO replaces a second order phase-lock loop and two frequency lock loops. A PLL consists mainly of a phase detector, a low-pass filter, an operational amplifier, and a voltage-controlled crystal oscillator (VCXO). The tracking is performed with a DC signal, and frequency information of the input signal is interrupted. The nonlinearities associated with the PLL elements reduce the accuracy of the system. Also, the low-pass filter, the amplifier, and the VCXO are frequency limited, and a PLL has a poor acquisition speed. In an SO, the synchronization occurs directly without conversion to a DC or a low-frequency signal. The frequency information between the input control signal and the output frequency is not lost; therefore, the SO is a coherent synchronization system. In the CRN of a QPSK modem, synchronization, tracking, filtering, and frequency division can be performed simultaneously in the same network. Experiments indicate that the tracking range and the noise rejection (synchronization gain) product for an SO is higher than for other synchronization and tracking networks.

An SO can divide and multiply frequencies by integers and certain noninteger numbers, with noise present. This capability is unique, especially with noise and is not achieved in a single process by any other means.

Because SOs can divide frequencies, there is a valid argument for the possibility of an SO functioning in a parametric mode, with division by 2 being a special case.⁴ For example, the synchronization signal is at its lowest, or has the lowest potential, when the external signal reinforces it. This relation indicates that the external energy is supplied to the SO when the SO energy is minimum, as in a parametric operation. Indeed, mathematically, the SO and parametric functions have the same differential equation, with the driving or control signal being twice the resonant frequency of the tank circuit. The manifestations of the two phenomena, subharmonic resonance (driving with $\sin 2\omega t$) and the parametric resonance, are different. For subharmonic resonance, the synchronization is directly produced by the external periodic excitation. For parametric resonance, it is caused by the periodic variation of one of the parameters in the absence of any direct external excitation.

Experiments indicate that an SO can be synchronized only when the waveform is at its minimum potential. Moreover, there is no mixing process in an SO to produce transients, nor are there loop delays. Therefore, because of all three considerations, unlike PLL, there can be no false lock with an SO.

There has not been serious attempts in their use as inadequate solutions of nonlinear differential equations did not reduce the real capabilities of the Van der Pol oscillators. Therefore, SOs should be distinguished from Van der Pol oscillators or injection oscillators.

The performance of an SO is determined by its quality factor, which, for a given driving energy, can be defined as

$$\text{Quality factor} = \frac{\text{Tracking range} \times \text{Synchronization Gain}}{\text{Time} \times \text{Bandwidth}}$$

This product should be maximized for carrier recovery. Also the skirt selectivity should be as steep as possible. In an SO, unlike passive circuits, the bandwidth (tracking range) and the noise rejection capacity can be increased at the same time, and both can be optimized to a high degree independently from each other, to deliver a highly efficient system. An optimization curve is shown in Figure 1. A similar situation exists in a chirp signal; where

Time x Bandwidth product is maximized.

There are three important differences between a SO and a FO:

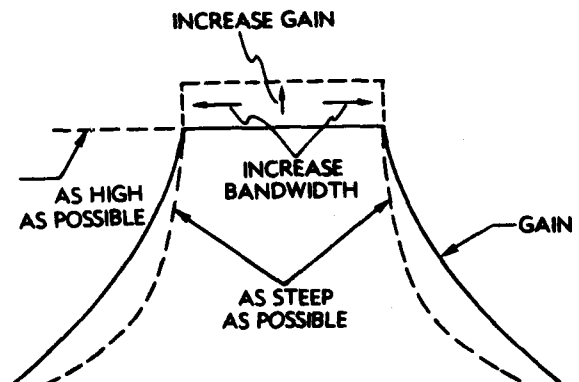
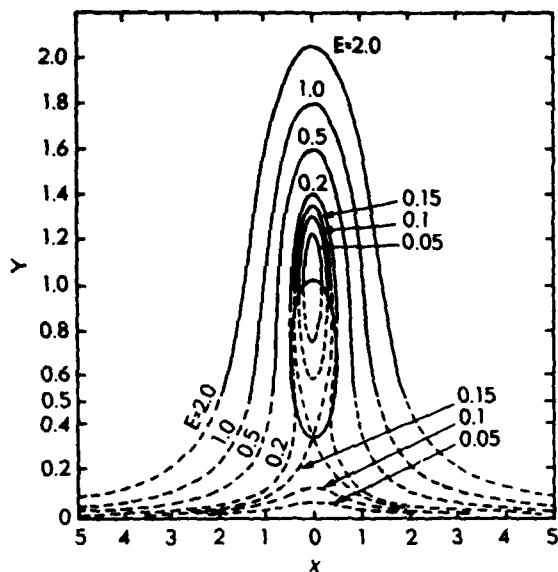


Figure 1
Quality Factor of A Synchronous Oscillator

1. In the FO mode the phase shift between the input and the output is 180° , whereas in the phase-locking mode the phase shift remains 360° after synchronization.
2. In the FO mode the synchronization curves resemble those of tuned resonant circuits, (See Figure 2) whereas the synchronization curves for SOs are rectangular in shape as shown in Figure 3. In FO the oscillator becomes an amplifier.
3. In FO decreasing the input causes the output of the SO to decrease, whereas in the SO decreasing the input, increases the output as shown in Figures 2 and 3.



x = AMOUNT OF DETUNING
Y = RESULTANT AMPLITUDE
E = EXTERNAL APPLIED VOLTAGE

Figure 2
Synchronization Curves For Van der Pol Oscillators

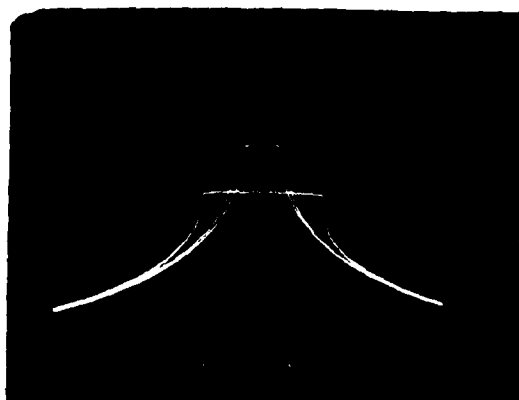


Figure 3
Gain-Synchronization Curves For Synchronous Oscillators

The classical Van der Pol oscillator equation¹⁻⁴

$$\ddot{v} + (\gamma v^2 - \alpha)\dot{v} + \omega_0 v = \omega_0^2 E_1 \sin \omega t$$

is solved by assuming a solution of the form

$$a_1(t) \sin \omega_1 t + a_2(t) \cos \omega_1 t$$

This form of solution linearities a highly nonlinear differential equation and the unique performance and functional properties of the oscillator are lost.

Although Dewan's¹ classification and concepts are correct, his results do not differ very much from the results of the classical solutions.

Gain and Phase Synchronization Curves

The gain-synchronization curves of a 90MHz SO and a 560MHz SO used as carrier recovery networks are shown in Figures 4 and 5. They have been displayed on HP 8505 Network Analyzer. Figure 5 compares also the selectivity of a three section filter to that of a SO. The skirt selectivities of Figures 4 and 5 are over 60dB/OCT, approaching those of crystal filters. The gain-synchronization curves of FOs, including Dewan's¹ resemble the curves of resonant circuits.

A SO can be synchronized reliably by a signal as low as -80dBm and a signal-to-noise ratio as low as -10dB. The SOs of Figures 4 and 5 are synchronized by

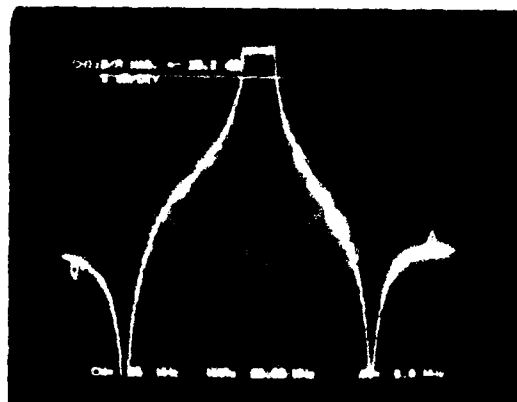


Figure 4
Gain-Synchronization Curve For
90MHz Synchronization Oscillator

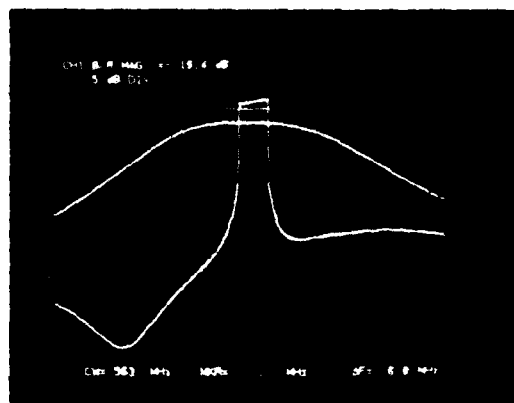


Figure 5
Gain-Synchronization Curve of A PLSO
Compared To A Three Section Filter

an input signal level of -20dBm, and have a synchronization bandwidth of approximately 1MHz. The skirt selectivity of an SO remains in the order of 12dB/OCT, and the signal-to-noise sensitivity is over 0dB. In phase-lock loop, for example, the skirt selectivity is only 6dB/OCT, and requires a signal-to-noise ratio higher than +5dB for synchronization. Also, a SO can acquire frequency synchronization within few cycles. Although phase settling requires more time it remains at least one order of magnitude lower than phase-lock loops, especially when there is frequency offset. For example, with a frequency offset 25KHz and tracking bandwidth of 60KHz, a phase-lock loop in the absence of noise can be synchronized in 10μseconds⁵, whereas

SO with a signal-to-noise ratio of -2dB can be synchronized in less than 500ns . The synchronization time approaches $1\mu\text{s}$ when the input signal-to-noise ratio becomes -5dB . A phase synchronization time curve is shown in Figure 6 for a 560MHz carrier recovery network with an input signal-to-noise ratio of -5dB . In a preamble used in a modem there are 48 bits of 1's and 128 bits of 1's and 0's for carrier and clock synchronization respectively. The start of the 1's and 0's sequence is shown by the second pulse, which settles $1.5\mu\text{s}$ sooner than the total time allocated for this purpose which is $2.8\mu\text{s}$, as shown in Figure 6. The pulse shows that a phase change occurred at that instant frequency is in synchronism and lock.

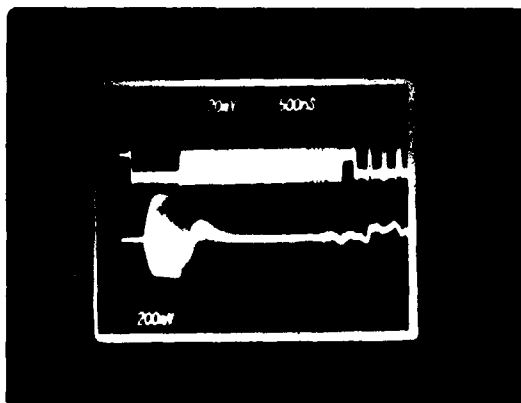


Figure 6
Phase Acquisition Curve of A Synchronous Oscillator in a Modem

The basic SO network which produces the gain-synchronization curves of Figures 3, 4, and 5 is shown in Figure 7. The series L_1C_1 resonant circuit produces the dips (zeros) in Figure 4 and 5 and their location are important in determining the steepness of the skirt selectivity of the gain-synchronization curves. Figure 3 shows also the adaptive filtering property of the SO. When the input signal-to-noise ratio is reduced the bandwidth becomes narrower and the gain increases to keep the output signal-to-noise ratio almost constant. The dips in Figures 4 and 5 before and after the synchronization band are due to zeros of the SO which occur before the poles. Similar dips exist also in crystal and elliptic filters.

The phase shift of a SO is linear within the synchronization band as shown in Figure 8. The total Phase shift of a SO within its synchronization band is always 180° . The phase shift at the edges of the synchronization band can be made to drop faster to reduce the phase shift per unit frequency offset at the around of the center frequency. Increasing the selectivity of the gain-synchronization curve increases the phase shift at the edges of the synchronization band to reduce the phase shift within the synchronization band. Figure 8, for example, has only 120° phase shift within the synchronization, instead of 180° .

Synchronous Oscillator as a Network Element

Figure 8 shows that the gain and phase synchronization curves violate Bode's phase-attenuation integral theorems.⁶ According to this theorem sharp changes in amplitude produce sharp and nonlinear changes in phase at and around the synchronization

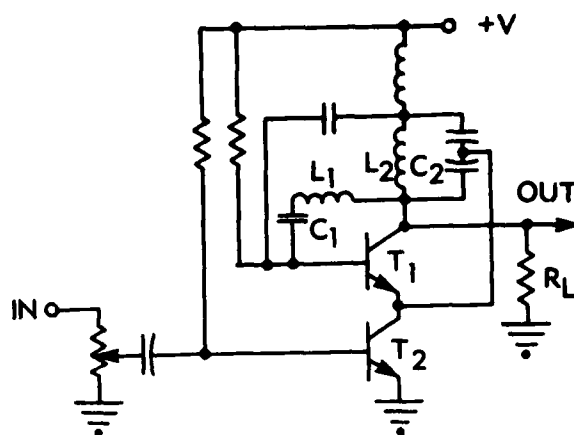


Figure 7
A Synchronous Oscillator Circuit Diagram



Figure 8
Gain And Phase Synchronization Curves of A Synchronous Oscillator

band. As a matter of fact the phase nonlinearities extend several octaves beyond the abrupt amplitude changes. Therefore, Figure 8 indicates an important difference between a linear minimum or nonminimum phase shift network and the nonlinear SO network. In addition, the gain synchronization curves of Figures 3, 4, and 8 by themselves approach the non-realizability criterion of linear networks (Paley-Wiener criterion) due to their sharp skirt selectivities and their flat synchronization band.⁷

Synchronous Oscillators as Carrier and Clock Recovery Networks

A SO can synchronize, track, filter, divide and improve the signal-to-noise ratio by as much as 30dB in a single process.⁸⁻¹⁰ For example, in a four phase shift keying modem, the SO replaces one second order phase-lock loop and two frequency lock loops, in the original carrier recovery network of a QPSK modem.

SOs have been used successfully in modems at various bit rates as clock and carrier recovery networks. The results were excellent and superior to any other clock and recovery network. For example, the bit error rate with recovered clock and carrier differed only by 0.2dB from the hard wired cases in burst mode operation in a 120Mbit/sec modem. Other results were equally impressive. SOs can perform also functions

not possible by other means, at least in the same simplicity and quality SOs do.

Broadbanding SO

The three possible ways to increase the tracking range of an SO:

- to increase the S/N of the input signal,
- to decrease the effective Q of the SO, and
- to decrease the gain of the oscillator transistor.

The increase in input signal level introduces high distortion and jitter, while lowering the gain of the oscillator transistor may suppress the oscillations and decreasing Q lowers the immunity to noise. Therefore, once the above parameters are optimized for a given requirement, additional adjustments in the SO parameters will degrade the performance of the SO. The method analyzed below increases the tracking range without affecting the other performance features of the SO.

The circuit in question of a 560MHz SO is shown in Figure 7, where a series L_1C_1 network is added to the basic circuit. The equivalent circuit of the series and parallel LC networks looking away from the base of T_1 is shown in Figure 9.

We start the analysis by assuming a single series resonant circuit shown in Figure 10.

The impedance of Figure 10 is given by

$$\begin{aligned} Z_1(\omega) &= sL_1 + \frac{1}{sC_1} + R/L \\ &= \frac{s^2L_1C_1 + 1 + sC_1R_L}{sC_1} \end{aligned} \quad (2)$$

Let

$$\omega_0^2 = 1/L_1C_1$$

and

$$Q_1 = \omega_0 L_1 / R_L$$

we can write equation (2) as

$$Z_1(\omega) = Q \frac{-\omega^2 + \omega_0^2 + j\omega\omega_0}{j\omega\omega_0} R_L \quad (3)$$

We are interested in the variation of reactance near the resonance frequency. The reactance in this case is given by

$$X_1(\omega) = Q_1 R_L \frac{(\omega - \omega_0)(\omega + \omega_0)}{\omega\omega_0} \quad (4)$$

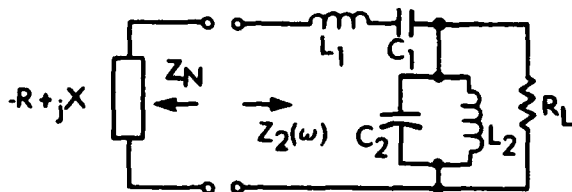


Figure 9
Equivalent Resonant Circuit of A
Synchronous Oscillator

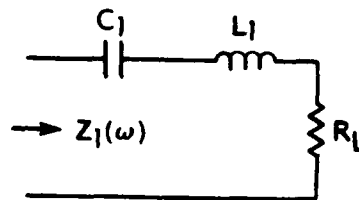


Figure 10
A Series Equivalent Tuned Circuit - L_1C_1

Equation (4) can be written as

$$X_1(\omega) = Q_1 R_L \frac{2\Delta\omega}{\omega_0} \quad (5)$$

We can write equation (5) as

$$\frac{\Delta\omega}{\omega_0} = \frac{X_1(\omega)}{2Q_1 R_L} \quad (6)$$

or

$$\frac{\Delta f}{f} = \frac{X_1(\omega)}{2Q_1 R_L}$$

We add a parallel resonant circuit to the existing series resonant circuit, and we have the circuit shown in Figure 9 looking into $Z_2(\omega)$. Z_N is the equivalent circuit looking into the base of T_1 . The reactance of series and parallel circuits near resonance varies in the same amount but in opposite directions. That is, the reactance of a series resonant circuit increases while that of a parallel resonance circuit decreases away from resonance.

Therefore, equation (6) can be written for Figure 9 as

$$\frac{\Delta\omega}{\omega_0} = \frac{X_2(\omega)}{2(Q_1 - Q_2) R_L} \quad (7)$$

By using a combination of a series and a parallel resonant circuit in an SO, it is possible to increase the tracking range while keeping individual Q's high. This is possible as both Q's are multiplied by the gain of the oscillating transistor. With $Q_1 + Q_2$,

the tracking range becomes wide while keeping the steep skirt selectivities. The circuits of Figures 3, 4, 5 and 8 use the L_1C_1 series resonant circuits.

Without the L_1C_1 networks the gain-synchronization curves can not achieve high selectivity and high gain. Note the sharp skirt selectivities with high gain. These assure high noise rejection capability with tracking range varying from 1 to 2.6 MHz. Note the dips on both sides of the tracking range, in Figure 4. This type of tracking performance is characteristic of two resonant circuits, the dips occurring due to series feedback resonance. In Figures 3 and 8 the dips are further away from the center frequency and can not be seen.

Although simple linear analysis shows that the transfer function of an SO resembles that of a second order PLL,³ experimental results and the present analysis indicate that the SO outperforms in all aspects a second order PLL. In addition a

SO is a multifunctional network which can replace several phase-lock and frequency-lock loops. SOs are also much simpler networks.

References

- 1 Edmond M. Dewan, "Harmonic Entrainment of Van der Pol Oscillations: Phase-Locking and Asynchronous Quenching," IEEE Transactions on Automatic Control, Vol. AC-17, No. 5, October 1972, pp. 655-663.
- 2 Van der Pol, "Forced Oscillations in a Circuit with Nonlinear Resistance," Phil. Mag., Vol. 3, January 1927, pp. 65-80.
- 3 R. Adler, "A Study of Locking Phenomena in Oscillators," Proceedings of the IRE, June 1946, pp. 351-357.
- 4 Hayaska Takeshi and Sokamoto Kazuo, "A Study of Injection Locking and Negative Resistance Amplifier," Transactions of the IECE of Japan, Vol. E61, No. 12, December 1978, p. 984.
- 5 Gardner, "Phaselock Techniques" John Wiley and Sons, Inc. 1966, page 46.
- 6 H. W. Bode, Network Analysis and Feedback Amplifier Design, D. Van Nostrand Company, Inc., Princeton, N. J., 1945.
- 7 Kuo, Network Analysis and Synthesis, John Wiley and Sons, Inc. 1962, Chapter 9.
- 8 Vasil Uzunoglu, "Carrier Recovery Networks for QPSK Modems Employing Synchronous Oscillators, U.S. Patent 4,355,404, October 19, 1982.
- 9 Vasil Uzunoglu, "Universal Clock Recovery Network for QPSK Modems," U.S. Patent 4,272,067, June 16, 1981.
- 10 Vasil Uzunoglu, "Division by Non-Integer Numbers Using Synchronous Oscillators, U.S. Patent 4,356,456, October 26, 1982.



FREQUENCY (STANDARD) COMBINER SELECTOR

Victor S. Reinhardt
Bendix Field Engineering Corporation
Columbia, Maryland

Raymond J. Costlow
Applied Physics Laboratory
Laurel, Maryland

Summary

This paper describes the Frequency Standard Combiner Selector (FCS) which was being developed jointly by the Bendix Field Engineering Corporation and the Applied Physics Laboratory for the NASA Space Tracking Data Network (STDN) headquartered at Goddard Space Flight Center (GSFC). The purpose of the FCS is to provide a redundant high performance frequency source for a redundant timing system.

The FCS is designed to take maximum advantage of an ensemble of frequency standards at a station by:

1. automatically combining their frequency outputs into a single output with the average properties of the whole ensemble,
2. automatically eliminating faulty members of the ensemble,
3. automatically keeping the FCS output on time and on frequency with respect to UTC regardless of the long term behavior of the ensemble,
4. being single point failure protected so a single point failure in the FCS itself will not disrupt operation.

This paper discusses in detail the design of the FCS and how its functions are implemented by the various subsystems. The heart of the FCS is a digitally implemented phase lock loop based on a multiple mixer phase comparison system, a digitally controlled VXCO, and a microprocessor. The digital loop allows one to phase lock the VXCO to the average phase of the ensemble of standards even though the various standards operate at different frequencies and even though the VXCO output frequency is purposely offset from the average frequency of the ensemble. This offsetability feature allows the FCS to use a second very long time constant (1 week) phase lock loop to adjust the offset frequency so the timing system output is kept on time and on frequency with respect to a 1 pps Universal Coordinated Time (UTC) reference provided by a suitable timing receiver. The second loop also allows the microprocessor to determine the frequencies of individual input standards against UTC. This allows the FCS to readjust the VXCO offset frequency when there is an input standard failure so that the FCS output remains on frequency as well as on time with respect to UTC.

The theoretical and experimental behavior of the digital loops used in the FCS are presented and other embodiments of the FCS concept are discussed. A subset system of the FCS called the Frequency (Standard) Selector which only switches redundant output between the various input standards in case of input standard failure is also discussed.

Description of the FCS Hardware

Figure 1 shows a simplified block diagram of the FCS and Figure 2 shows an overall view of the rack containing the FCS next to a teletype and a PDP-11

computer used as a back up to the LSI-11 controlling the FCS. A digitally controlled crystal oscillator (DXCO) composed of an 18 bit digital to analog converter and a high stability VXCO is used to provide the 5 MHz outputs of the FCS under normal conditions. A phase comparison system is used to monitor the difference in phase between the 5 MHz output of the DXCO with a resolution as high as 0.05 ps. This information is used by an LSI-11 microprocessor to phase lock the DXCO to the average phase of the atomic standards. In this phase lock mode, the fractional frequency of the DXCO can be arbitrarily offset with respect to the average frequency of the atomic standards with a range of $\pm 1 \times 10^{-7}$. This effectively turns the DXCO into a low noise synthesizer.

In practice, the frequency resolution of this synthesizer is virtually infinite. The processor controls the DXCO frequency with a resolution of 8×10^{-13} per bit. But the phase lock loop allows the processor to change the frequency of the DXCO ten times a second so that the loop can adjust the phase of the DXCO in 0.08 ps steps. This effectively turns the DXCO into a synthesizer with a frequency tracking resolution of:

$$\frac{\delta f}{f} = \frac{8 \times 10^{-14}}{\tau}$$

where τ is the averaging time in seconds.

The frequency offsetability of the phase lock loop is used to keep the DXCO output on time with UTC regardless of the frequencies of the input standards. An event clock monitors the difference in time between the local timing system 1 pps output driven by the FCS and UTC information provided via LORAN-C or some other time source. The proper frequency offset to keep this time difference zero is maintained either manually or in a secondary loop.

Error detection and single point failure protection are accomplished in a triple redundant subsystem called the Frequency (Standard) Selector (FS). The FS constantly monitors the health of the input standards, the DXCO, and other system elements. If an input standard fails, it is thrown out of the phase lock loop by the processor. Errors that will cause loop failure, including processor failure, cause the FS to transfer the 5 MHz outputs from the DXCO to the first standard. This is accomplished in 1 μ s so an LC filter can maintain the 5 MHz outputs during the switching process keeping the time error caused by this type of failure down to ± 100 ns. A subsequent failure in the first input standard then causes the FS to similarly switch all outputs to the second input standard. The FS is described in more detail later in the paper.

A teletype and digitally controlled multichannel strip chart recorder constantly provide hard copy documentation of system operation. This documentation can also be sent to remote locations via several RS-232 interfaces in the FCS.

The controller of the FCS is a Digital Equipment Corporation LSI-11 microcomputer which is interfaced to most of the digital hardware of the FCS through an IEEE standard CAMAC crate and data bus. The CAMAC crate containing the LSI-11, the 18 bit digital to analog converter, and a multichannel event clock (to be described later) is shown in Figure 3. Not shown is the 16 channel recorder interface which also fits in the CAMAC crate. The CAMAC crate was used to interface the hardware rather than the standard LSI-11 data bus so the modules developed for the FCS could be used with other computers. The four RS-232 interfaces are part of the LSI-11. They can be used to remotely control the FCS as well as for transmitting monitoring data.

The phase comparison system is a multichannel generalization of a dual mixer system¹. The system utilizes an 8 channel phase comparator, a multichannel event clock, and a crystal local oscillator (LO). The 8 channel phase comparator is shown in Figure 4. The 8 channel phase comparator is made to be used with an LO which is offset from the frequency standards to be compared by frequencies up to 10 Hz. The 8 channel comparator consists of a buffered distribution system for the LO and 8 buffered single channel phase comparators which produce TTL square waves synchronized to the zero crossings of the beat frequencies between the LO signal and the each input frequency standard signal. At 5 MHz, the 8 channel comparator provides greater than 100 db of isolation between pairs of input standards and between the input standards and the LO. At 5 MHz with a 1 Hz beat, the noise floor of the phase comparator is 0.05 ps. The performance of the phase comparator as a function of averaging time is shown in Figure 5. Even though the 8 channel phase comparator is designed for 5 MHz, it is usable from over 50 MHz to under 100 kHz. The theory behind the operation of the phase comparator is described in the next section.

The multichannel event clock records the epochs of either the positive or negative edges of the TTL signals coming from the 8 channel phase comparator and the 1 pps UTC reference. Several different multichannel clocks have been developed with resolutions ranging from 200 ns to 12.5 ns and with 6 to 16 channels. If higher resolution is desired, a laboratory time interval counter can be interfaced via an IEEE 488 interface connected to either the CAMAC crate or directly to the microcomputer.

The LO consists of a Frequency and Time Systems FTS-1000 crystal oscillator with a nominal frequency of 4,999,990 Hz. Other LO's with only a 1 Hz offset from 5 MHz have been successfully used in the system. The LO is left free running so the phase comparison system will operate regardless of the status of any of the atomic input standards or even the VXCO. The operation of the phase comparison system with a free running LO is described in detail in the next section.

Theory of the Phase Comparison System

Before describing the theory of the phase comparison system, it is helpful to review some pertinent frequency standard theory. The output of a frequency standard can be described by:

$$V = A \sin [2\pi f_0 + \phi(t)] \quad (1)$$

where f_0 is its nominal or ideal frequency and $\phi(t)$ describes all the phase deviations from ideal behavior.

One can show that if this frequency standard is used to drive a clock that, at any instant, the error in this clock's reading is given by:

$$x(t) = \frac{\phi(t)}{2\pi f_0} \quad (2)$$

x is called the normalized phase error or clock reading error. For simplicity in this paper, x will be called the clock error.

By taking the time derivative of ϕ one obtains the instantaneous angular frequency offset from $2\pi f_0$:

$$\frac{d\phi}{dt} = \dot{\phi} = 2\pi\delta f(t)$$

Dividing this by $2\pi f_0$ yields the instantaneous fractional frequency error:

$$y = \frac{\delta f}{f_0} = \frac{\dot{\phi}}{2\pi f_0} = \frac{dx}{dt} = \dot{x} \quad (3)$$

That is, the instantaneous fractional frequency offset, y , is the time derivative of the clock error, x . Averaging y over some time τ :

$$\bar{y}(t + \tau, t) = \frac{1}{\tau} \int_t^{t+\tau} y(t') dt' \quad (4)$$

yields the important relation:

$$\bar{y}(t + \tau, t) = \frac{x(t + \tau) - x(t)}{\tau} \quad (5)$$

Now, within this framework, the phase comparison system can be described.

The phase comparison system is a generalization of the dual mixer phase comparison technique¹. A simplified schematic diagram which describes the dual mixer technique is shown in Figure 6. The multiple mixer system utilizes more than two mixers connected to the transfer or local oscillator so any pair of frequency standards can be intercompared as in the dual mixer system. In the actual system, the mixer is replaced by a single channel phase comparator consisting of two buffer amplifiers to provide isolation, a mixer to produce a beat frequency signal, and a low noise zero crossing detector with a narrow band front end amplifier to produce fast rise time square waves synchronized to the beat frequency. The noise bandwidth of the system is defined by the bandwidth of the front end amplifier.

To understand the detailed operation of the phase comparison system, consider an input on channel i of the form:

$$V_i = A_i \sin (2\pi f_0 t + \phi_i(t))$$

and a signal from the local oscillator of the form:

$$V_L = A_L \sin (2 f_0 t + \phi_L(t))$$

where $f_0 = 5$ MHz. Notice that all the phase deviations in these signals from that of an ideal 5 MHz oscillator have been put into ϕ_i and ϕ_L respectively (included in ϕ_L is the 10 Hz offset from 5 MHz).

Each single channel phase comparator outputs a TTL square wave whose edges correspond to the zero crossings of the beat frequency out of the double balanced mixer. The mixer for channel i outputs the beat frequency between the local oscillator and the channel i input. Since the local oscillator is lower in frequency than all the channel i inputs, the i th mixer output is of the form:

$$V_M = A_M g(\phi_L(t) - \phi_i(t))$$

where g is some sine-like function whose only important properties for this discussion are that:

$$g(n\pi) = 0$$

and that the slope of g is positive for:

$$\phi_L - \phi_i = 2n\pi$$

The zero crossing detectors of the i th channel turn the function, g , into a square wave whose positive edge occurs at:

$$\phi_L(t_{ni}) - \phi_i(t_{ni}) = 2n\pi \quad (6)$$

where t_{ni} is the epoch of the n th positive edge of the channel i beat frequency square wave and where the phases ϕ_L and ϕ_i have been defined to make $\phi_L - \phi_i$ zero at the epoch of the first positive edge under consideration, t_{oi} .

Equation 6 implicitly relates the phase deviations between the channel i oscillator and LO to the epoch of the i th channel zero crossing. To obtain the clock errors explicitly in terms of the zero crossing epochs, we proceed as follows. Using (2), (6) becomes:

$$x_L(t_{ni}) - x_i(t_{ni}) = \frac{n}{f_0}$$

Taking the difference between two channels, gives:

$$x_i(t_{ni}) - x_j(t_{mj}) = x_L(t_{ni}) - x_L(t_{mj}) - \frac{n-m}{f_0}$$

Using (5), for \bar{y}_L and \bar{y}_j , the fractional frequency offset of the LO and CS_j respectively, finally yields:

$$x_i(t_{ni}) - x_j(t_{mj}) + \frac{n-m}{f_0} = \quad (7)$$

$$[\bar{y}_L(t_{ni}, t_{mj}) - \bar{y}_j(t_{ni}, t_{mj})](t_{ni} - t_{mj})$$

This is the equation we are after since it relates the difference in clock error between two inputs measured at the same time to the t_{ni} 's which are measured by the multichannel event clock.

Equation (7) states that, to determine the difference in clock error between any two channels, only the relative frequency between the LO and one of the oscillator channels concerned must be known. In other words, only the relative frequency and not the phase of the LO affects the phase measurement. In practice, one need only estimate $\bar{y}_L - \bar{y}_j$ periodically to obtain sufficient accuracy for the phase measurement. The error this will introduce is given by:

$$\delta x = \delta \bar{y} (t_{ni} - t_{mj})$$

Since, in general, we can make $t_{ni} - t_{mj}$ 100 ms by choosing n and m , x will be less than 10^{-13} s if $\bar{y}_L - \bar{y}_j$ is measured to 10^{-12} . The LO is a low noise crystal oscillator whose stability from 1 s to 100 s is better than 10^{-12} and, in general, all the atomic standards' frequencies will be known to 10^{-12} . We can therefore use any atomic standard input to estimate $\bar{y}_L - \bar{y}_j$ and can take up to 100 s to make the measurement to have an accuracy of 10^{-13} s for differential phase measurements.

Through (7) we can also determine how frequency instabilities in the LO determine the system measurement noise for clock error. Defining the clock error jitter variance from measurement to measurement as:

$$\sigma_{mx}^2 = \frac{1}{2} \langle [x_{ni} - x_{n+1,i} - x_{mj} - x_{m+1,j}]^2 \rangle$$

in the absence of any noise from the input oscillators where $\langle \dots \rangle$ indicates an infinite time average, one can show from (7) that:

$$\sigma_{mx}^2 = \sigma_x^2 (2, t_{ni} - t_{mj}, t_{ni} - t_{n+1,i}) \quad (8)$$

where $\sigma_x^2(2, \tau, T)$ is the Allan variance³ of x of the LO for averaging time τ and dead time $T - \tau$. (σ_{mx}^2 has been defined to make the system noise equal to the Allan variance of the LO. If the more conventional σ_x^2 were used to define the system noise, the system noise equation would not be so simple.) In analyzing (8), it is important to consider whether $t_{ni} - t_{mj}$ is greater or less than t_c , the correlation time of the low pass filters in the zero crossing detectors (which determine the noise bandwidth). Assuming that $t_c \ll t_{ni} - t_{mj} \leq 1$ s, the Allan variance above is approximately a constant. For the LO used, it is about 1 ps. Thus for $t_{ni} - t_{mj} \gg t_c$, the phase noise of the LO over the beat period determines the system measurement noise. For $t_{ni} - t_{mj} \ll t_c$, the correlation of the low pass filter causes a reduction in system noise given approximately by:

$$\sigma_{mx} = \frac{|t_{ni} - t_{mj}|}{t_c} \cdot \sigma_x$$

where σ_x is the LO phase noise for $|t_{ni} - t_{mj}| \gg t_c$.

Since the measurement system is measuring independent frequency standards, $|t_{ni} - t_{mj}|$ will assume all

values between 0 and the period of the beat frequency between the LO and the frequency standards. This means that if τ_c is less than the beat frequency (which is true in our case), the LO phase noise will limit the phase resolution of the multiple mixer measurement system. A way to overcome this for long term measurements is to average the results of N measurements of x_{ni} after correcting each with an independent measurement of $\bar{y}_L - \bar{y}_i$. Since each measurement is individually corrected for changes in LO frequency, the contribution of the LO to the measurement noise will act like white phase noise because the long term correlations from the LO will have been taken out by the correction process. Therefore, the contribution of the LO noise in the average of x_{ni} will be reduced by the square root of N .

If the multichannel event clock has a resolution, R , the smallest phase change that can be measured is $\bar{y}_L R$. For $\bar{y}_L = 2 \times 10^{-6}$ (10 Hz offset), and $R = 2 \times 10^{-7}$ s, the phase measurement resolution is 0.4 ps.

Digital Phase Lock Loop

First consider locking the DXCO to a single reference input. Let y be the fractional frequency output of the DXCO. If the DXCO control has a sensitivity of S (df/f per bit) and the control register value is Y , then:

$$y = S Y + y_f$$

where y_f describes the free running behavior of the DXCO. Let x_{ni} be the clock error between the DXCO and the reference oscillator:

$$x_{ni} = x_o(t_{no}) - x_i(t_{ni})$$

where channel o is the DXCO channel. Then using (7):

$$x_{ni} = R (T_{no} - T_{ni}) (\bar{y}_L - \bar{y}_i)$$

where R is the resolution of the event clock in seconds per bit and T_{ni} is the epoch of the n^{th} beat zero crossing for channel i in bits. Since we are locking on only one reference, we can assume it is on frequency, that is, $\bar{y}_i = 0$.

A first order phase lock loop is given by the equation:

$$Y_n = -(B/S) x_{ni}$$

which for times long compared with the beat period, τ_o , is equivalent to:

$$\frac{dx}{dt} = -Bx + y_f$$

This has the solution:

$$x = x_o e^{-Bt} + y_f/B$$

where B^{-1} is the time constant of the phase lock loop.

When phase locking to a cesium standard, B^{-1} should be about 100 s because the DXCO is less noisy than a cesium frequency standard up to about a 100 s averaging time. This causes problems due to the presence of y_f , however. The DXCO has a rated frequency drift of less than 10^{-10} per day or about 10^{-15} per second. Since this is a slow drift:

$$\frac{dx}{dt} = B^{-1} \frac{dy_f}{dt}$$

so the DXCO would run at a fractional frequency offset from the input standard of about 10^{-13} due to the free running DXCO frequency drift. Such an uncontrolled offset in the phase locked DXCO is unacceptable.

To overcome this problem, a second order loop is used. The DXCO register is incremented at each beat period, τ_o , by:

$$\Delta Y_n = -(B/S) \Delta x_{ni} - (C\tau_o/S) x_{ni}$$

where the operation Δ is defined as:

$$\Delta a_n = a_n - a_{n-1}$$

This is equivalent to:

$$\frac{\Delta Y_n}{\tau_o} = -(B/S) \frac{\Delta x_{ni}}{\tau_o} - (C/S) x_{ni}$$

and for times long compared to τ_o this becomes:

$$\frac{d^2 x}{dt^2} = -B \frac{dx}{dt} - Cx + D \quad (9)$$

where D is the drift rate of the DXCO. Equation (9), in general, has the solution:

$$x(t) = a_1 e^{-r_1 t} + a_2 e^{-r_2 t} + \frac{D}{C} \quad (10)$$

where:

$$r_1 = B/2 + \frac{1}{2} \sqrt{B^2 - 4C}$$

$$r_2 = B/2 - \frac{1}{2} \sqrt{B^2 - 4C}$$

and where a_1 and a_2 are constants determined by $x(0)$ and $\dot{x}(0)$. For the phase lock loop to be stable, we must have:

$$B^2 \geq 4C$$

$B^2 = 4C$ defines the critically damped solution:

$$x(t) = (a_1 t + a_2) e^{-r_c t} + \frac{D}{C} \quad (11)$$

where:

$$r_c = B/2 \text{ and } C = \frac{B^2}{4}$$

In this case, a frequency drift will only produce a phase offset:

$$x = D/C$$

For critical damping, $r_c^{-1} = 100$ s, and $D = 10^{-10}$ /day, this equation yields $x = 10$ ps. This is below the phase jitter of the average phase of even 10 cesium standards so DXCO frequency drift should not cause problems with a second order loop.

Now consider a second order phase lock loop which tracks the DXCO to the average phase of several atomic standards (CS_i) with an adjustable fractional frequency

offset, y_0 . Again, the phase difference between the DXCO and CS_i is:

$$x_{ni} = R(T_{no} - T_{ni})(\bar{y}_L - \bar{y}_i)$$

where again R is the resolution of the clock, and T_{ni} is the i th channel beat zero crossing epoch in bits.

In a practical phase lock loop, a problem in computing x_{ni} occurs. Because the frequencies of the input standards are in general different, $T_{no} - T_{ni}$ will diverge in time (even if y_0 were set to zero). This means that an infinite memory would be required to store all the past values of T_{ni} required to compute $T_{no} - T_{ni}$ at any given instant. To avoid this problem, using (7) again, we can rewrite x_{ni} as:

$$x_{ni} = R(T_{no} - T_{mi})(\bar{y}_L - \bar{y}_i) - \frac{n-m}{f_0} \quad (12)$$

This allows us to use the closest T_{mi} to T_{no} so we don't have to store past values. From (12), we can obtain the difference between the DXCO phase and the average phase of N cesiums:

$$x_n = \frac{1}{N} \sum_{i=1}^N [R(T_{no} - T_{mi})(\bar{y}_L - \bar{y}_i) - \frac{n-m(i)}{f_0}] \quad (13)$$

where the fact that m is a function of i is explicitly indicated.

To lock the DXCO to the average phase of the cesiums with a fractional frequency offset y_0 , we can define a variable:

$$x_n = y_0(T_{no} - T_{00}) - \frac{x_n}{R}$$

where the units of x_n are in bits. Since the event clock is driven by 5 MHz from the DXCO, the DXCO output defines the local time scale against which measurements are being made. We therefore can consider that:

$$R = \frac{1}{f_0}$$

Using this and (13) we obtain:

$$x_n = \frac{1}{N} \sum_{i=1}^N [(T_{mi} - T_{no})(\bar{y}_L - \bar{y}_i) + J_{ni}] \quad (14)$$

$$+ y_0(T_{no} - T_{00})$$

where:

$$J_{ni} = n - m(i)$$

A second order loop in the variable x_n is defined by:

$$\Delta Y_n = \left(\frac{B}{Sf_0}\right) \Delta x_n + \left(\frac{C\tau_0}{Sf_0}\right) x_n \quad (15)$$

where ΔY_n is the amount the DXCO control register is incremented. In the actual algorithm, ΔY_n is limited to:

$$-L \leq \Delta Y_n \leq L$$

in order to keep the Y register from being changed drastically by a bad data point. This allows the loop to function in high noise environments. One can show that, by judicious choice of L , the loop can be made to function smoothly with bad data occurring 50 percent of the time or more and that the loop can function without degraded performance with bad data occurring up to 10 to 20 percent of the time (if the bad data points are randomly distributed).

Because the CS_i have different frequencies, J_{ni} and $y_0(T_{no} - T_{00})$ will individually grow without limit n goes to infinity. This will not cause computational problems in computing x_n , however, if we rewrite (14) as:

$$x_n = \frac{1}{N} \sum_{i=1}^N (T_{mi} - T_{no})(\bar{y}_L - \bar{y}_i) + K_n \quad (16)$$

where:

$$K_n = \frac{1}{N} \sum_{i=1}^N J_{ni} + y_0(T_{no} - T_{00}) \quad (17)$$

K_n must be bounded as n goes to infinity because all the other terms in (16) are bounded (The phase lock loop causes x_n to be bounded.) K_n can be computed by setting it initially to zero and computing increments to add with:

$$\Delta K_n = \frac{1}{N} \sum_{i=1}^N \Delta J_{ni} + y_0 \Delta T_{no}. \quad (18)$$

The ΔJ_{ni} can be computed by two methods. The simplest method is to use the definition of the J_{ni} and increment ΔJ_{ni} for each DXCO zero crossing and decrement ΔJ_{ni} for each CS_i zero crossing. The disadvantage of this is that should a noise pulse cause a false T_{ni} or should a zero crossing be skipped, the DXCO will shift in phase by 200 ns/N.

To avoid such permanent phase shifts in the DXCO, another method for calculating the ΔJ_{ni} can be used as follows. With the convention that T_{mi} is the first channel i event after T_{no} occurs, we can define:

$$DT_{ni} = T_{mi} - T_{no}$$

DT_{ni} will range between zero and approximately 0.1 s. If the \bar{y}_i and y_0 are not too large, DT_{ni} will change slowly with n until 0 or τ_0 is reached. Then, the value of DT_{ni} will suddenly change. Thus, to compute ΔJ_{ni} , one can use the algorithm:

1. If: $DT_{ni} - DT_{n-1,i} > \frac{\tau_0}{2N}$
2. then: $\Delta J_{ni} = -1$
3. If: $DT_{ni} - DT_{n-1,i} < -\frac{\tau_0}{2N}$
4. then: $\Delta J_{ni} = +1$

5. Otherwise: $\Delta J_{n1} = 0$

The advantage of this method is that if there is a bad J_{n1} or a skipped zero crossing, ΔJ_{n1} will switch between +1 and -1 or vice versa on successive calculations leaving no permanent phase shift.

\bar{y}_L and the \bar{y}_i can be calculated from the event clock data. Once the DXCO is locked, it defines local time, and by definition is at frequency f_0 . Therefore, \bar{y}_L and the \bar{y}_i can be estimated by:

$$\bar{y}_{nL} = \frac{\bar{f}_{nL} - f_0}{f_0} = \frac{m}{T_{n+m,0} - T_{n,0}} \quad (19)$$

and:

$$\bar{y}_{nL} - \bar{y}_{ni} = \frac{\bar{f}_{nL} - \bar{f}_{ni}}{f_0} = \frac{m}{T_{n+m,i} - T_{ni}} \quad (20)$$

Because the LO is a free running crystal oscillator, \bar{y}_L must be constantly updated to correct for crystal drift. Since only $\bar{y}_L - \bar{y}_i$ must be known, to calculate X_n , (20) is all that is needed for loop operation. To make the estimate error on the order of 10^{-12} or less, but often enough to correct for LO changes, m should be between 10 and 1000 (1-100 second averages). (19) is used to determine \bar{y}_i from (20) for diagnostic purposes.

FCS Performance

The performance of the phase comparison system and the 5 MHz distribution amplifiers used in the FCS have been presented elsewhere² so this data will not be presented here. Figures 7, 8 and 9 show the critically damped second order phase lock loop performance for 1 s, 10 s, 100 s time constants respectively using two Hewlett Packard high performance cesium frequency standards as input standards. Besides showing the parameters discussed in the previous sections, the figures also show the 1 pps from one of the cesiums monitored by the clock. This shows whether the loop jumps cycles of 5 MHz. In the figures, X_e is X_n and VXCO is Y_n from the theory section converted to time units and fractional frequency change in the VXCO respectively. Notice that in the 10 s loop there is bunching of the VXCO (Y variable) noise distribution and that in the 100 s loop there is a sudden spike in X_e accompanied by a jump in VXCO representing about a 4×10^{-12} jump in the DXCO crystal. The spikes in the 100 s loop occurred only occasionally, with periods between the spikes ranging from a few hours to a few days. Notice also that in the 100 s loop the 1 pps did not jump. This shows that the loop does not jump cycles even under a severe disturbance. The authors feel that both the 10 s bunching in the VXCO data and the 100 s spikes are due to vibration of the AT cut crystal used in the DXCO by fans in the rack holding the FCS. Discussions with others in the crystal manufacturing business^{4,5} indicate that the crystal may have exhibited these occasional jumps because it was cut too near the anomalous resistance point. Since we did not specify to Frequency and Time Systems that the crystal oscillator should exhibit smooth clock behavior and since the crystal oscillator meets all manufacturers specifications, we can only

fault ourselves in this matter. We are also confident that these skips would disappear if a properly specified crystal were used as the DXCO.

Before we could cure these problems, however, work on the FCS was stopped for the reasons having nothing to do with the FCS itself. The FCS was being developed to provide a redundant high performance frequency source for the next generation timing systems for NASA's Spaceflight Tracking and Data Network (STDN) stations run by Goddard Space Flight Center (GSFC). Development work on the FCS was stopped before the FCS was completed because NASA Headquarters decided to consolidate STDN and the Deep Space Network (DSN) run by the Jet Propulsion Laboratory (JPL). JPL had been developing a Coherent Reference Generator (CRS) similar to the FCS for their next generation timing systems. When the consolidation was announced, GSFC and JPL worked out a combined effort to produce a single next generation timing system for the combined DSN-STDN network. It was decided that, for the combined system, the JPL CRS and the GSFC timing system would be used. Therefore all work on the FCS was discontinued and only the Frequency (Standard) Selector (FS) alone was produced for the new DSN-STDN the timing systems.

Frequency (Standard) Selector

The purpose of the Frequency Standard Selector (FS) is to provide a user with a triply redundant source of 5 MHz and 10 MHz derived from 3 independent 5 MHz frequency standards designated the primary, secondary and tertiary frequency standards. Unlike the FCS, the FS will not combine the multiple frequency standard inputs into a single output and will not keep the output on frequency with respect to UTC. A block diagram of the FS is shown in Figure 10. Both the primary and secondary frequency standard channels each have an internal RF detector and 2 user tailorable fault monitors. When an RF failure or a fault condition on either of the fault monitor lines occurs for the primary standard, the FS switches all outputs to the secondary standard. If the secondary standard then has an RF failure or causes a fault condition on either of the 2 secondary fault monitors, the FS will switch all outputs to the tertiary standard. All fault sensing is latched so any change initiated on a fault will not be restored until the latches are reset by push button or remote TTL command. To minimize the time error occurring from switch-over, LC filters are used to hold up the amplitude of the 5 MHz signals until logic circuits and RF detectors can react. A relatively abrupt change of up to ± 100 ns will occur in the output signals at switch-over, however, because of the differing phases of the 5 MHz signals from the 3 frequency standards. The fault monitors can be tailored by the user via DIP kluge plugs to sense, as fault conditions, either a positive or negative change in a voltage ($>2V$), a resistance (also used for contact closure), or a logic level.

The Frequency Standard Selector achieves redundancy by complete triplication of all circuitry. After a matrix of passive power splitters to create 9 isolated RF sources from the 3 RF inputs, all circuitry in each of the 3 sections is completely independent. The unit is designed to be completely protected against all single point failures and to allow repair of any modular unit without interrupting power. To aid in fault diagnosis, each section also monitors the voltages coming from the power supplies of the other two sections. For convenience, the user can interchange frequency standards 1 or 2 as the primary and secondary standards from the front logic panel. One can also select an alternate switch-over sequence from the primary directly to the tertiary standard.

All control switches and LED status monitors in the FS are connected to remote TTL connectors so a remote computer can both monitor the performance and remotely control the FS. This was done both to allow the combined DSN-STDN network to operate in a computer controlled environment and to allow the FS to be turned into the FCS by adding appropriate modules.

Some pictures of the FS are shown in Figures 11, 12, and 13. Figure 11 shows a front view of the FS showing the triple power supply and a diagnostic monitor panel. Figures 12 and 13 show the triply redundant digital and RF sections respectively.

CONCLUSIONS

Even though sudden frequency jumps in the VXCO caused occasional 0.5 ns spikes to appear in the DXCO output when a 100 second time constant was used, the FCS concept basically proved successful; even under the severe stress caused by these VXCO jumps the digital phase lock loop did not break lock or jump cycles. With some further work to reduce the vibration environment of the VXCO and to find a more smoothly behaving VXCO, it is felt that these occasional spikes could be eliminated. The FCS concept also goes beyond the specific hardware used. It is felt that the digital loop described here can be successfully implemented with IEEE 488 standard test equipment and a calculator if the calculator has sufficient speed to implement the loop within the time constraints imposed by the beat period and the loop time constant.

For long term time keeping purposes, the loop can be implemented utilizing a phase microstepper rather than a DXCO (as in the U. S. Naval Observatory master clock system). In this configuration the output would be derived directly from a single input, but the reliability problems caused by this could be eliminated by using the Frequency Standard Selector to switch inputs in case of master standard failure.

ACKNOWLEDGMENTS

The authors would like to acknowledge Bob Bush and Harry Smith of Bendix Field Engineering Corporation for designing and building several of the digital modules in the FCS. The authors would also like to acknowledge Robert Kruger of Bendix Field Engineering for building the Frequency Selector.

REFERENCES

1. D. W. Allan and H. Daams, "Picosecond Time Difference Measurement System," 29th Annual Symposium on Frequency Control (Atlantic City, New Jersey, 1975).
2. V. Reinhardt et. al., "A Modular Multiple Use System for Precise Time and Frequency Measurement Distribution," Proceedings of the Tenth Annual Precise Time and Time Interval Planning Meeting (Washington, D.C., 1978).
3. B. E. Blair, Ed., Time and Frequency: Theory and Fundamentals, NBS Monograph 140, U. S. Government Printing Office, SD Catalog Number C13.44:140 (Washington D. C., 1974)
4. A. Bates, Applied Physics Laboratory, private communication (1980)
5. M. Bloch, Frequency Electronics, private communication (1983)

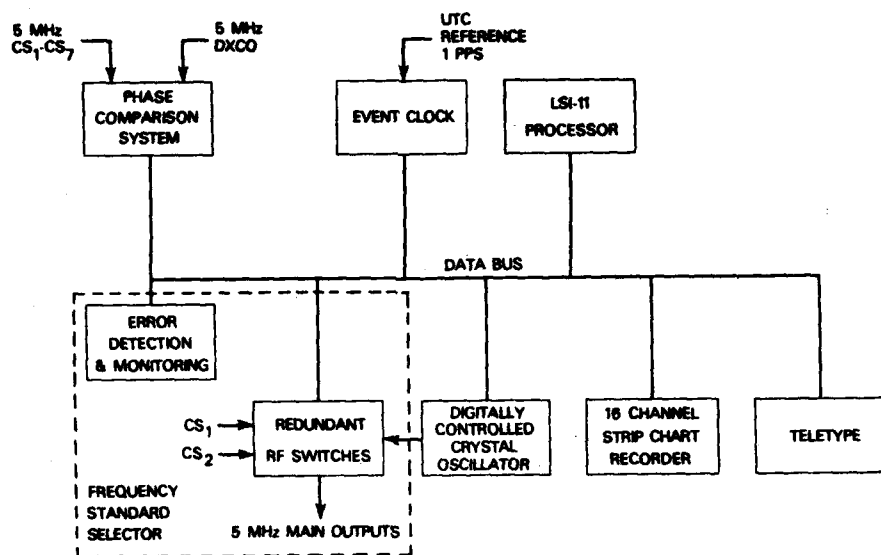


Figure 1 A Simplified Block Diagram of the Frequency Combiner Selector



Figure 2 The Rack Containing the FCS (Right) Next to a Teletype and a PDP-11 Computer

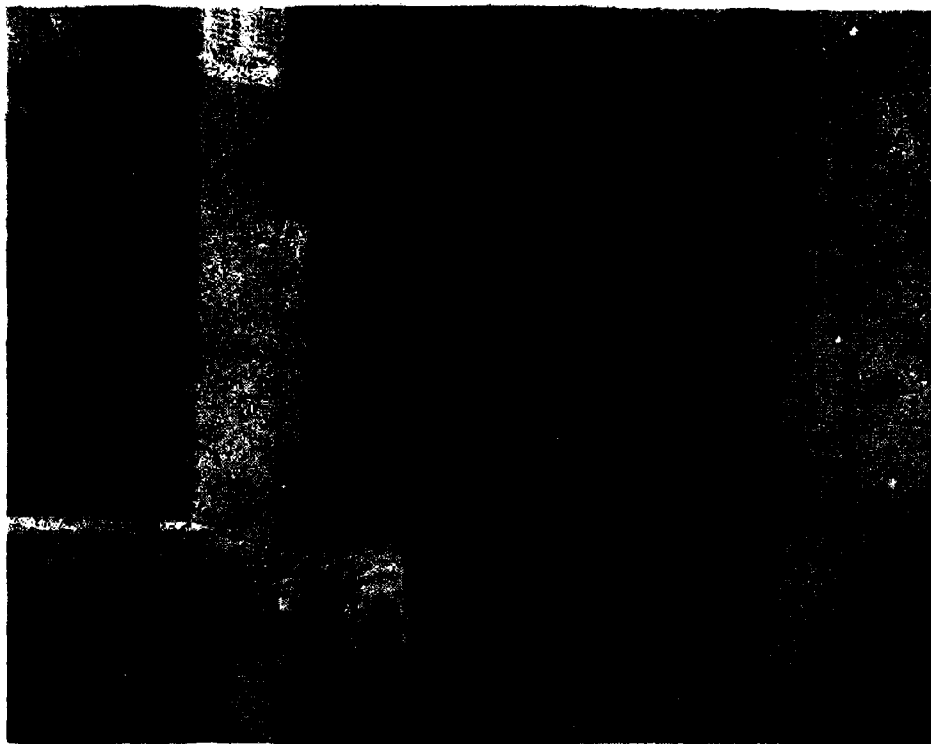


Figure 3 The CAMAC Crate Containing the LSI-11, the 18 Bit Digital to Analog Converter, and the Multichannel Event Clock

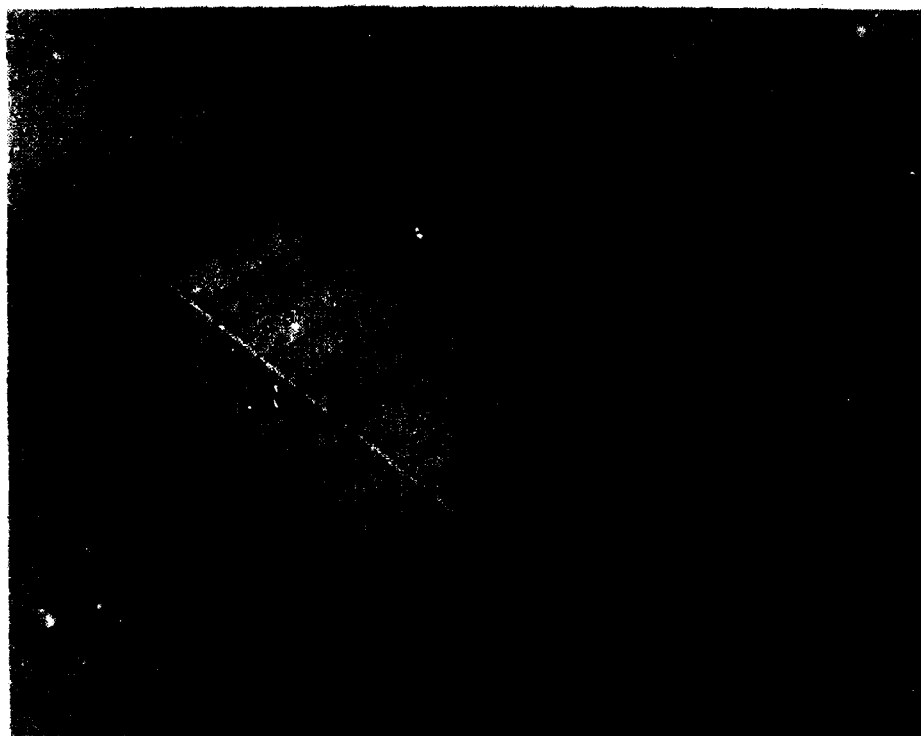


Figure 4 The 8 Channel Phase Comparator

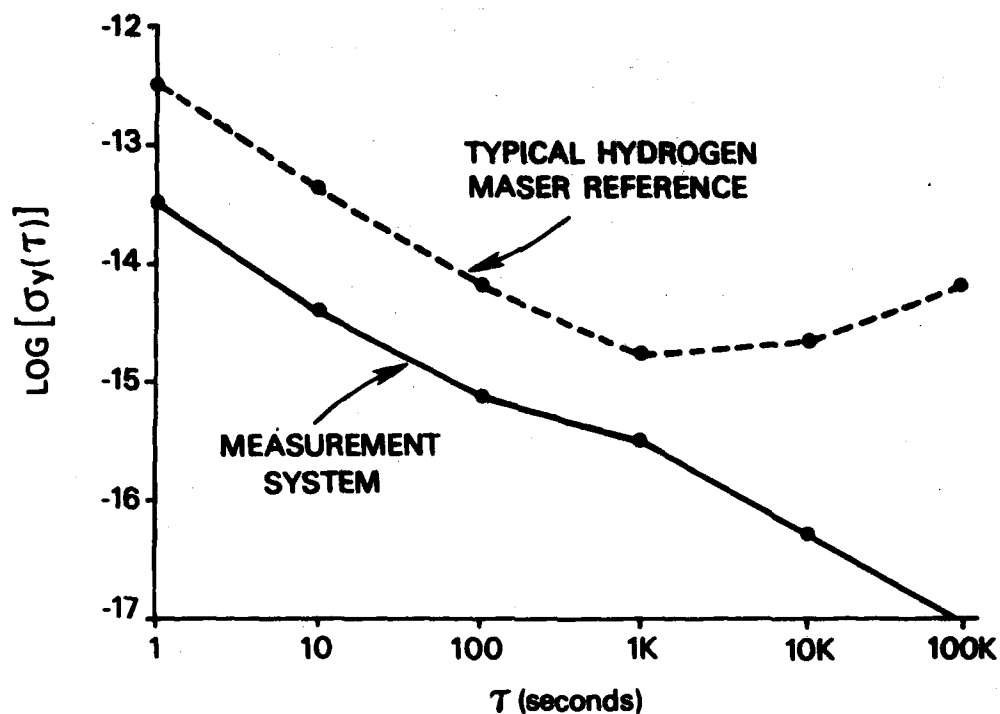


Figure 5 Phase Comparator Frequency Measurement Resolution

DUAL MIXER HETERODYNE TECHNIQUE

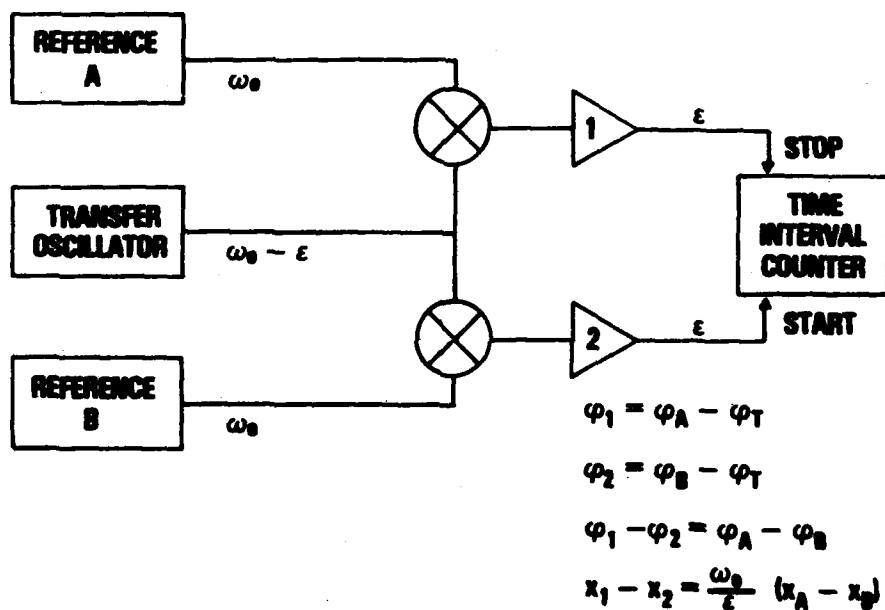
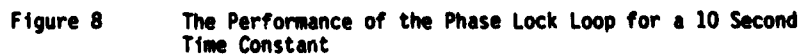
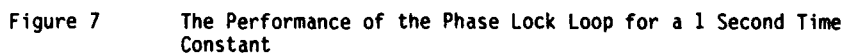


Figure 6 A Simplified Schematic Diagram of the Multiple Mixer Phase Comparison System



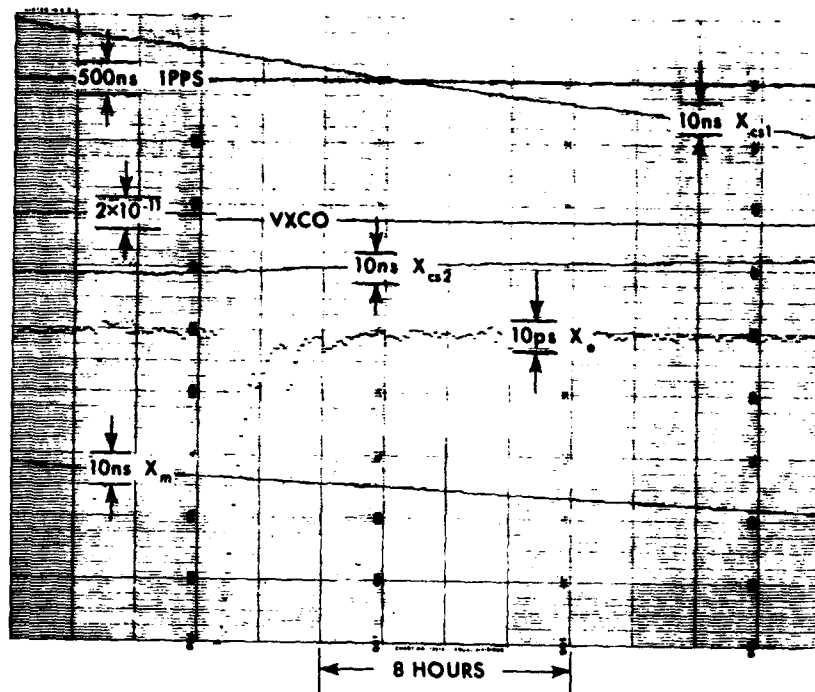


Figure 9 The Performance of the Phase Lock Loop for a 100 Second Time Constant

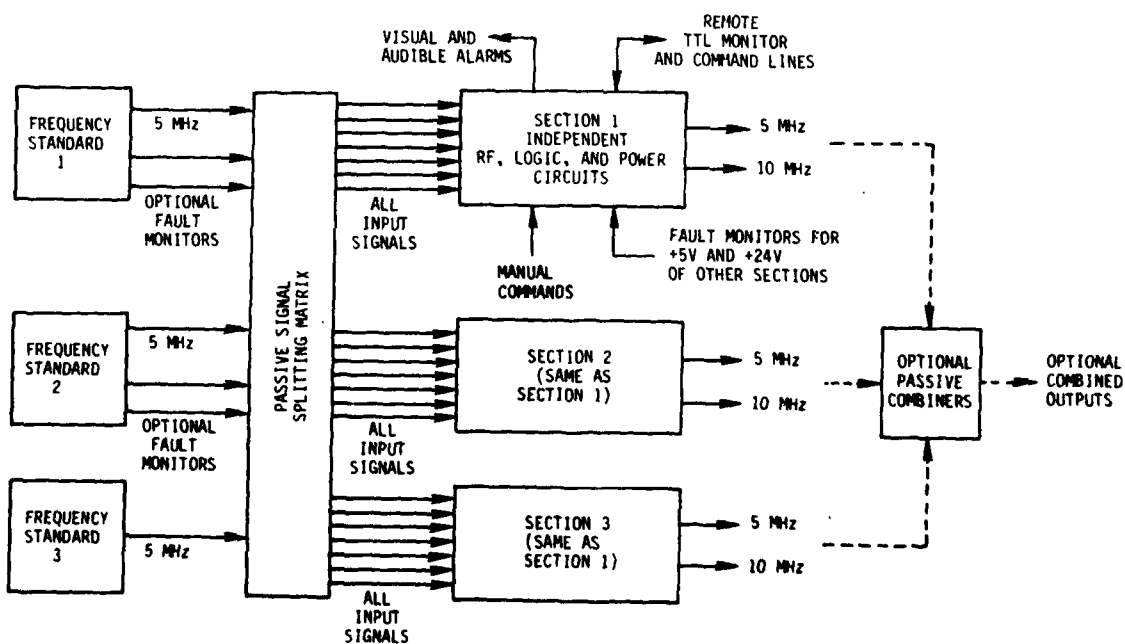


Figure 10 Frequency Standard Selector Block Diagram

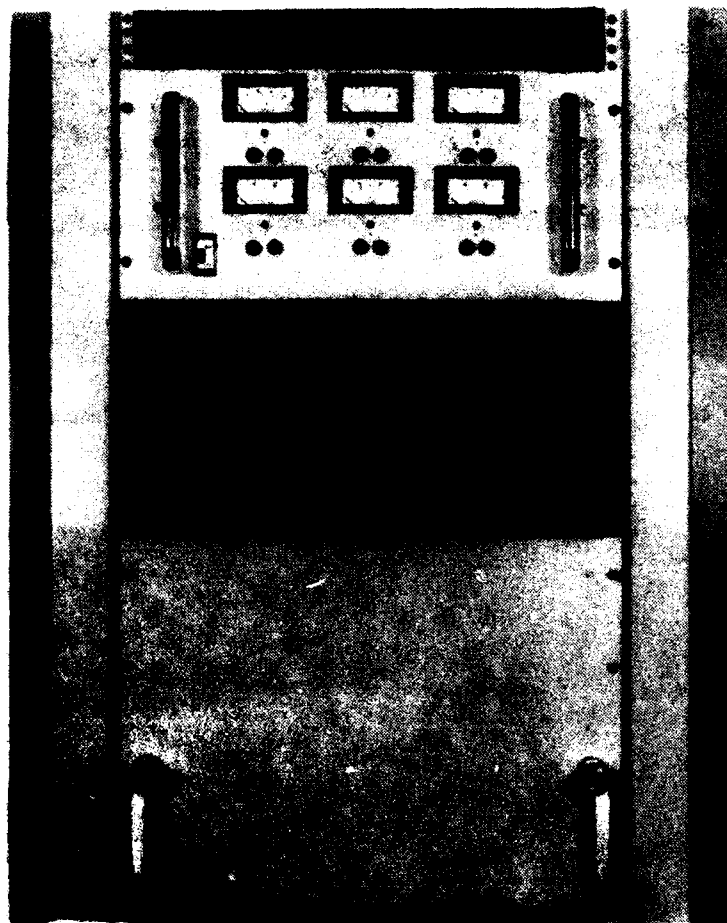


Figure 11. Front View of the Frequency Standard Selector

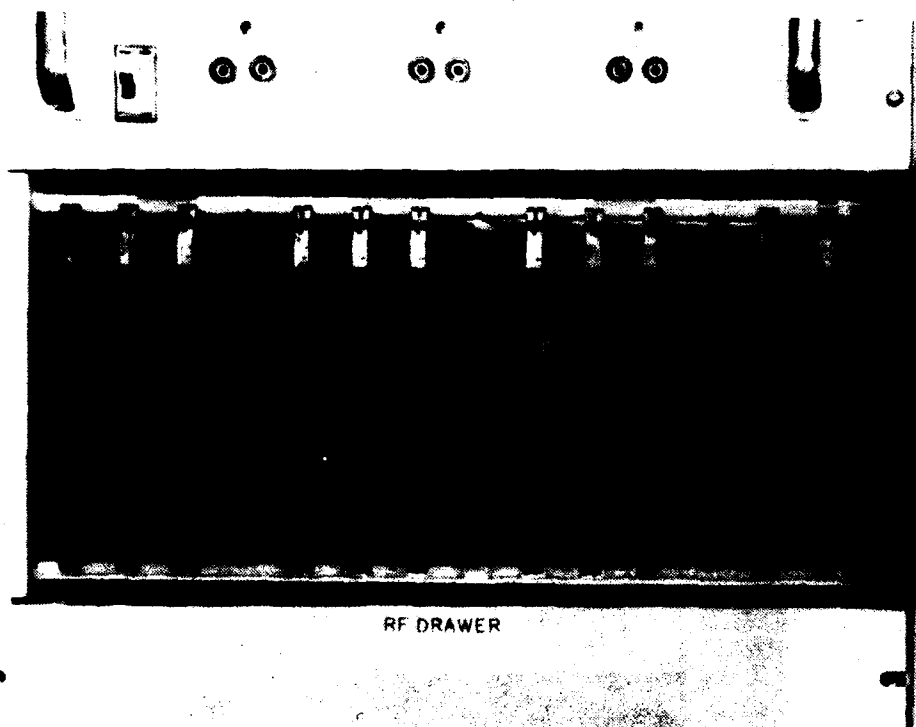


Figure 12 The Triply Redundant Digital Section of the Frequency Standard Selector

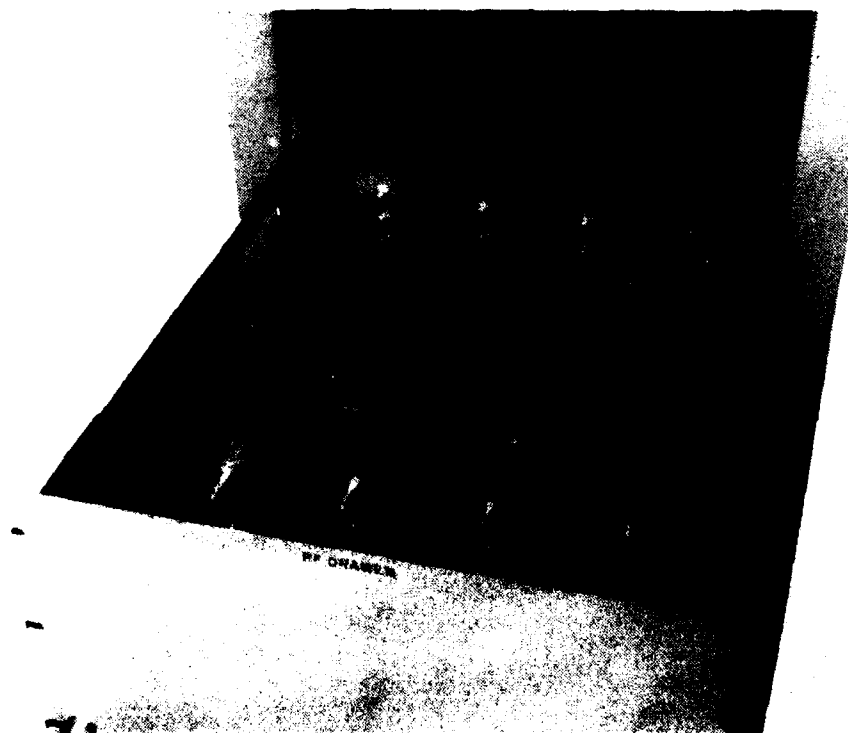


Figure 13 The Triply Redundant RF Section of the Frequency Standard Selector

NON DESTRUCTIVE OBSERVATION OF RANDOM ELECTRICAL TWINNING IN CULTURED QUARTZ

H. Morigoux, J.F. Darces, P. Zecchini
Laboratoire de Cristallographie et Synthèses Minérales
Faculté des Sciences et des Techniques
25030 Besançon Cedex - France

J. Lamboley
Q.E. Alcatel
B.P. 19
92602 - Asnières - France

Summary

It is well known that in a cultured quartz bar we can observe, around the seed, several growth zones which are indicated by Z, +X, -X... These zones have been described either by X-ray topography or by γ -ray irradiation.

By infrared spectroscopy we find that we can distinguish between these zones. Since their infrared spectra are different according to the trapped impurities it is easy to identify each growth zone from its spectrum.

Sometimes, among our irradiated blanks, we note that we can see, inside the normally -X zone, the +X zone's characteristic blackening. The IR spectra confirm the irregular presence of the two zones.

Taking in account the crystal symmetry we can explain these observations. There is a new supplementary two fold axis. This axis commands the electrical twinning. According to the IEEE convention this kind of twinning only changes the resonator's θ and ϕ angle sign.

With the X-ray diffraction method developed in Besançon, we know how to determine these θ and ϕ angles with their sign. Our measurements of these abnormal zones show that the electrical twinning law and not the combined law may be present in all classical cuts used for piezoelectrical purpose. Such a result can be obtained because we are able to measure simultaneously with our goniometer the orientation of two associated crystals.

Introduction

It seems to be accepted that in cultured quartz we cannot find an accident such as a twin crystal. Indeed when we manufacture quartz we place in the pressure vessel only controlled seeds of quartz free from twinning. During the growing process it is assumed that the deposited new quartz crystal presents the same structure that the seed does : i.e. the same handedness and the same electrical axis orientation. So we are confident of the quality of the cultured quartz when we cut a blank. However we sometimes find troubles with some resonators showing unexpected frequency features. Then if we take out the quartz from the resonator and make a new X-ray measurement of its orientation we will find, using a classical procedure, that the orientation is a good one or one with some small variations which cannot explain the observed anomalous motion of the resonator. So we must focus our attention on the quartz itself and characterize the quality of the quartz by others methods. Using gamma-irradiation and infrared spectroscopy we have found the key to this problem and now we are able to explain this discrepancy between the

apparently correct X-ray orientation and the bad properties of the resonator. We are able to detect this defect before the blank is transformed into a resonator.

IR observation of the growth zones

A cultured quartz is obtained by deposition, in a high pressure solution, on a seed having a Z-cut or nearly Z-cut orientation.

During the growth the speed of the deposition is slightly anisotropic. The highest speed is observed in both Z directions and the lowest in the -X direction. So we may describe the bulk crystal as a juxtaposition of several crystals grown nearly independently around the seed, this difference comes from their direction of growth. Each of them is called the Z or \pm X growth zone. These zones had been observed by X-ray topography^{1,2}, but you may see them directly on the quartz³ itself if you irradiate a Y-cut with gamma-rays. A blackening appears more or less intense according to the growth zone (figure 1).

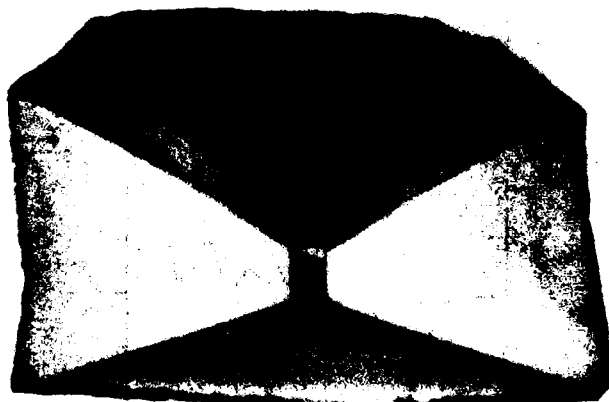


Figure 1 : Y-cut of cultured quartz bar after gamma-irradiation.

The color is unchanged for both Z-zones, but the +X is darker than the -X one. Between the +X and the two Z-zones we observe zones called S-zones ; they are divided into two subzones S₁ and S₂. The growth zones differ also according to their physico-chemical properties and their blackening is linked to the

electrical changes in their trapped impurities. Knowing their exact position we are now able to record their infrared spectra, with the help of a mask.⁴ We observe either an adjacent slide of the irradiated Y-cut blank or the same slide after having heated it. The initial electrical level of the crystal impurities is then restored. Both methods give the same results. All spectra are obtained on a 580B Perkin Elmer infrared spectrometer, at the nitrogen temperature, between 3800 and 3100 cm^{-1} .

In figure 2 we have the infrared spectra of the six zones. The quartz presented here is of medium quality, the IR measured Q is about $1.2 \cdot 10^6$. Each growth zone is identified by a black mark in a small cartouche. The two Z-zones are alike with regard to the two fold axis of the quartz. This axis lies on the surface of the observed Y-cut.

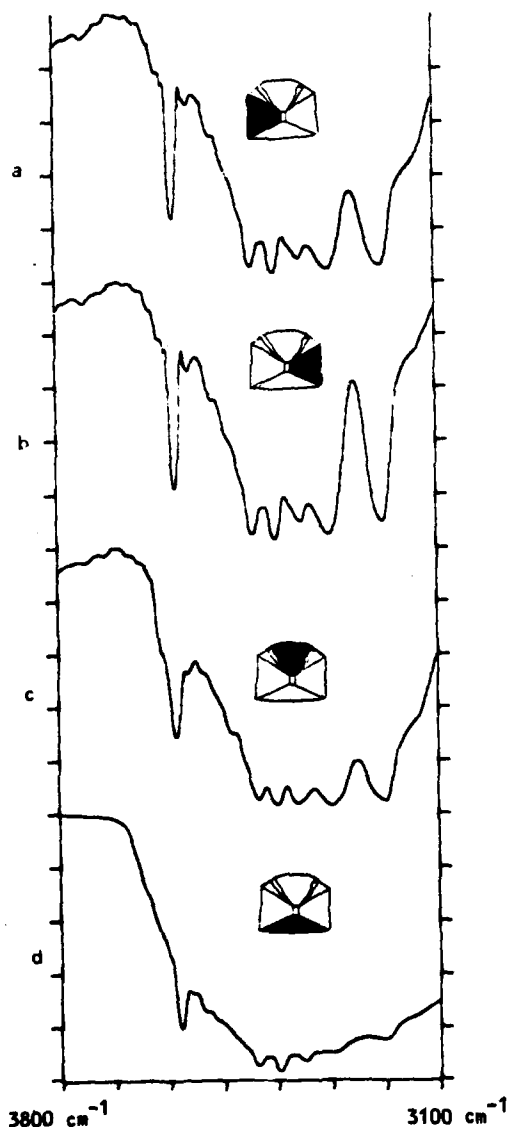


Figure 2 a-b-c-d : Infrared spectra at Nitrogen temperature

a-b : Z growth zone
c-d : X growth zone

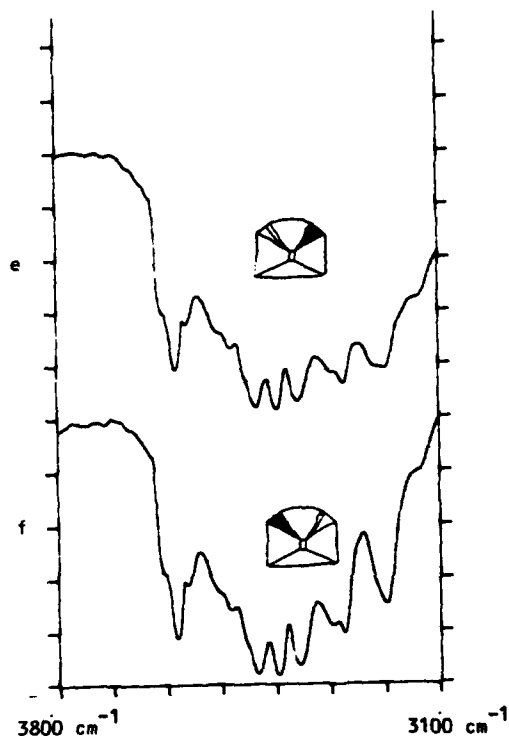


Figure 2 e-f : Infrared spectra at Nitrogen temperature

e-f : S growth zone

We recognize SiO absorption bands at 3200 and 3300 cm^{-1} and the four OH "S" bands.^{5,6} The +X-zone, is quite similar to the Z-zones. The OH-bands are slightly stronger. In the -X-zone, details of the spectrum are hidden inside in a very large band. This wide band is particular to this zone. With the last two S-zones the spectra are quite similar to those observed with the +X-zone : alkaline impurities are stronger.

Observation of an external defect

Once on the -X-surface of a cultured quartz, the nearly flat one, we have observed small hillock as shown figure 3.

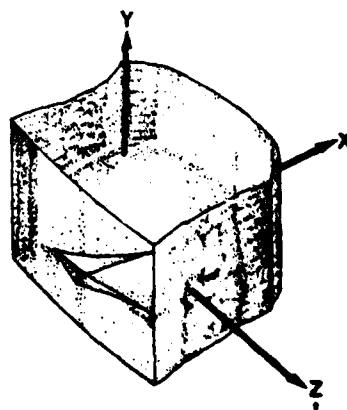


Figure 3 : Hillock on the -X-surface

After irradiation of a Y-cut including this defect we have observed an anomalous blackening in the -X-growth-zone, near the hillock. The infrared spectra figure 4 a-b-c, confirm that in the -X-zone we have a small area of a +X-zone, figure 4.b ; both SiO absorption bands are seen again figure 4.c.

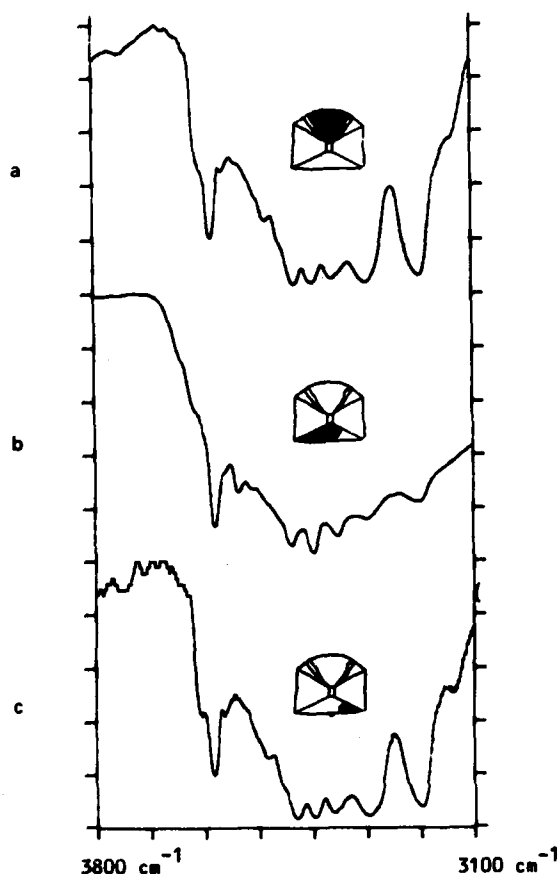


Figure 4 - Y-cut at Nitrogen temperature

a : +X growth zone
b : -X growth zone
c : near the hillock

The explanation is that we have an accidental twinning in the crystal. We do not yet know if this phenomenon obeys the electrical or the compound twinning law.

Quartz twinning

The quartz symmetry is $32 (A_2, 3A_2)$, the fourth part of its lattice symmetry $6/m \text{ mm}$ ($A_6 3A_2' 2 3A_2'' 2$ C). Then there are several ways to place the crystal in its lattice. What we are saying may be simply illustrated with a very common example. If we take matches in their box, there are two ways to place them. The phosphorous head may be placed at one end or at the other end. These two possibilities come from the elements of symmetry match-box lacking matches. With quartz the situation is more complicated because there

are more different elements of symmetry able to turn over the crystal in its lattice. So if we partition the symmetry elements of the lattice not present in the crystal, we have, according to their effects on the turn over, three sets of elements.

- $A_6, 3A_2'' 2$: they are associated with a proper rotation of the reference axis, they induce the electrical twinning law
- $M, 3M''$: they are associated with an improper rotation changing only the handedness of the crystal, they induce the optical twinning law
- $C, 3M'$: they are associated with a change on both the handedness and the direction of the reference axis, they induce the compound twinning law.

X-ray observation of a twinned quartz

In the "quartz lattice" any direction $[hkl]$ is repeated 24 times by the symmetry elements. In the "quartz crystal" they are repeated only 6 times. So there are only 6 equivalent planes having the same X-ray reflective power among the 24. We have four sets of six planes each. Three of these sets are obtained from the fourth by one of the three twinning laws. The principle of the search for twins is based on the observation of the reflective power of planes of the different sets. The discrepancy, if any, between theory and observation determines the type of twinning law. But there is a supplementary phenomenon called "anomalous scattering". In a non centrosymmetrical crystal, such as quartz, reflective power for two planes connected by the center of the lattice are nearly the same. On a piezgoniometer it is not easy to differentiate the variation induced by the anomalous scattering. In a first approximation we can say that X-rays add a symmetry center to the quartz when observed on a piezgoniometer. If the goniometric observation adds a center, it hides the effect of the center involving the compound twinning law. In the same way, during such a measurement optical and electrical twinning law are combined.

Instead of 4 sets we can only see 2 sets of 12 planes each. In the first one there are planes coming from the untwinned crystal and from the twinned part by the compound law. In the second one, if any, there are those coming from the electrical or optical law.

Experiments and X-ray diffraction apparatus

Once the orientation of the blank is measured on the piezgoniometer, we can calculate the angular positions required for the observation of planes coming from the two sets. These planes are selected between those having a large variation in the reflective power when we change sets. The most useful planes come from the (052) and (124) families because we can use them with nearly all the classical cuts. During the observation of a plane the crystal is translated, not rotated, on the reference metallic plane of the piezgoniometer. We know we are observing the same lattice plane because there are no rotation movement on the apparatus. All we have to do is record the variation of the intensity of the reflected beam. The piezgoniometer we use in this work allows us to determine the orientation of any blank. Because we are able to observe any lattice plane, we just have to

adjust the Bragg angle for the X-ray detector and search the reflected beam in combining two rotation movements.⁷ From the angular positions read on the goniometer we can calculate the value of the θ and ϕ angles, with their sign if we know the quartz handedness.⁸ The only special device needed for the twinning research is a translation motion on the sample holder.

Discussion and conclusion

X-ray goniometry and infrared spectrometry are two complementary methods for twinned quartz crystal observation. With the first one, we see the orientation error of the crystal in its lattice. It can work only if the twinned crystal appears on the surface blank. With the second one, we see ponctual defects in the volume of the sample. It can work only in the X-zones because they have very different trapped impurities. Infrared spectra can be obtained at room temperature as shown in figure 5 for a doubly rotated cut of a cultured quartz. In figure 5.a we have an actual +X growth zone. In figure 5.b an actual -X growth zone. In figure 5.c, instead of the -X expected spectrum, we have the two SiO bands of the +X growth zone. The spectra are less accurate than those obtained at the Nitrogen temperature, but the discrepancy between the two X-zones is still noticeable.

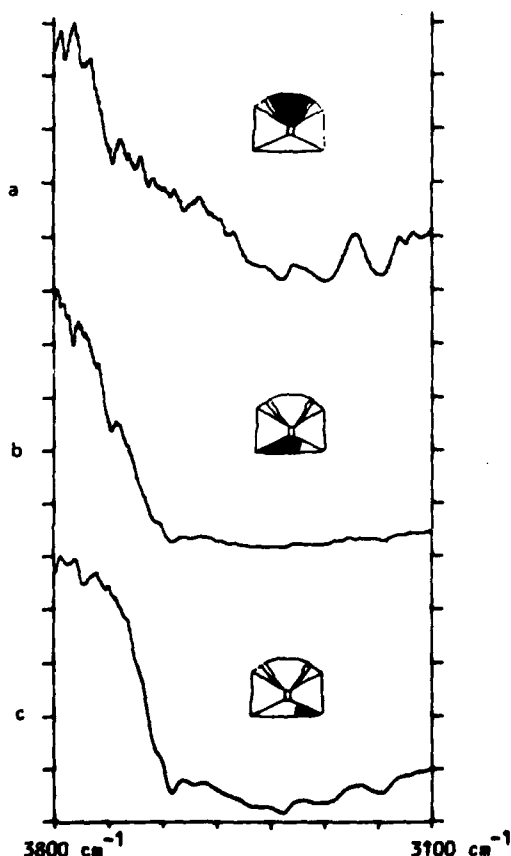


Figure 5 : doubly rotated cut at room temperature

- a : +X growth zone
- b : -X growth zone
- c : Twinned area

The detection of twinning in quartz crystal is a very important problem, because twinning changes the crystal orientation. For instance, with the electrical twinning law we have the $(-\theta, -\phi)$ orientation instead of the (θ, ϕ) expected angles (IEEE Conventions). The twinned cultured quartz crystal we have shown may give soon a lot of troubles when making a resonator.

It seems that the defects observed by infrared spectroscopy are not linked to the seed ; we may call it an accidental twinning. Such a defect has no importance for a resonator because we do not use the -X growth zone. But if we use such a crystal as a seed we will induce twinned area in a new quartz crystal. Such a mechanism may be described in the figure 6. There is a irradiated Y-cut. We can recognize the growth zones in the crystal and in the seed.



Figure 6 : The seed of this crystal is cut in both X and Z growth-zones.

If the seed possesses a twinned area, the new crystal will grow twinned even in the Z-zones. Such a mechanism if not eradicated may spoil all a cultured quartz production.

Acknowledgments

We would like to express our thanks to the D.R.E.T. of the Ministry of Defense for her generous support.

References

- 1 - Homma, S., Iwata, M. (1973)
X-ray topography and EPMA studies of synthetic quartz - J. Cryst. Growth, 19, pp. 125-132.
- 2 - Zarka, A., Liu-Lin, Buisson, M. (1981)
Influence de la localisation sectorielle du germe sur la qualité cristalline de quartz de synthèse. J. Cryst. Growth, 54, pp. 394-398.
- 3 - Charkraborty, D., Lehmann, G. (1977)
Infrared studies of X-ray irradiated and heat treated synthetic quartz single crystals. N. Jb. Miner. Mh. 7, pp. 289-298.

- 4 - Leung, C.S. (1982)
Etude par spectrométrie infrarouge de quartz synthétiques. - Thèse de 3ème cycle, Université de Franche-Comté-Besançon.
- 5 - Kats, A. (1962)
Hydrogen in alpha-quartz. - Philips Research Reports. 17 (3), pp. 133-279.
- 6 - Dodd, D.M., Fraser, D.B. (1965)
The 3000-3900 cm^{-1} absorption bands and inelasticity in crystalline α -quartz. - J. Phys. Chem. Solids 26, pp. 673-686.
- 7 - Darces, J.F., Mèrigoux, H., (1978)
Final X-ray control of the orientation of round or rectangular quartz slides for industrial purposes. - Proc. of the 32nd Annual Symposium on Frequency Control, pp. 304-309.
- 8 - Darces, J.F., Lamboley, J., Mèrigoux, H., (1982)
Dissymetry used for ϕ , θ angle sign determination of a piezoelectric crystal blank by a non destructive method. - Ferroelectrics, vol. 40, pp. 245-248.



TENSILE FRACTURE STRENGTH OF ST CUT QUARTZ

H. Lin Chao & Thomas E. Parker
Raytheon Research Division
131 Spring Street
Lexington, MA 02173

Abstract

The tensile fracture strengths of a total of 83 ST cut, single crystal, quartz discs were measured under a variety of experimental conditions. The major emphasis of this study was to compare the fracture strength of four different types of quartz including natural, and three synthetic grades (optical, premium Q and electronic). Since fracture strength often depends on surface finish, two different polishing techniques were also assessed. One technique used a Syton polish starting from a rough ground surface while the second technique involved a fine mechanical polish followed by a short Syton polish. A biaxial flexure test was used for measuring the tensile fracture strength on the discs which were 1 inch in diameter and nominally 0.1 inch thick.

The average fracture strength of polished samples was 21,400 psi with a standard deviation of $\pm 6,000$ psi. There was very little difference in fracture strength for the four types of quartz and the two different polishing techniques. The differences between the average fracture strength for each category fell well within the standard deviation. Five unpolished discs were also tested and the average fracture strength was found to be $8,900 \pm 800$ psi. Clearly, polishing the surface serves to increase the fracture strength by better than a factor of two.

Two particular techniques for altering the surface quality were also investigated. It has been reported that chemical etching of very thin quartz discs can significantly increase the fracture strength. To test the effect of chemical etching on thick discs, ten chemically etched samples were also broken. The average fracture strength was significantly lower than for polished discs, but many of the etched samples showed numerous etch pits or channels. Some the samples had few or no etch pits or channels and these showed a fracture strength comparable to the polished samples. Another seven samples had 1600 shallow grooves (~ 1000 A deep) ion milled into a polished surface. These samples showed no significant change in average fracture strength. In addition to the parameters mentioned above, other factors such as relative humidity, Q (from infrared measurements), and sample thickness also were found to have no correlation with fracture strength.

Introduction

This paper discusses research made on the basic tensile fracture strength of quartz. It has been over 10 years since a value for

the fracture strength of quartz ($\sim 14,000$ psi) was published¹ and this number should be re-evaluated in light of the progress made in the technology of growing and polishing quartz. The major emphasis in this study was to compare the fracture strength of discs made out of four different types of quartz. These were natural quartz and three synthetic grades (optical, premium Q, and electronic). Since fracture strength often depends on surface finish, two different sample polishing techniques were used. The electro-mechanical Q of many of the quartz discs was also measured using infrared absorption techniques. The test method used for fracturing was the biaxial flexure test. Results from these fracture tests will be presented as well as results which show the effects, on the fracture strength of quartz, of chemical etching and ion beam etched grooves about 1000 A deep.

Sample Preparation

Eighty-three ST cut, Z-zone, 1-inch diameter and 0.1 inch-thick (after polishing) quartz discs were purchased. Both natural and synthetically grown quartz were supplied, including the three synthetic grades of optical, premium Q, and electronic. The fracture strength of a material depends upon the quality of the surface finish, where surface flaws such as scratches can concentrate the applied stress and thereby lead to fracture. Complete removal of flaws from the surface permits the strength to approach the theoretical maximum for the particular material. For comparison purposes, two different sample preparation and polishing techniques were followed.

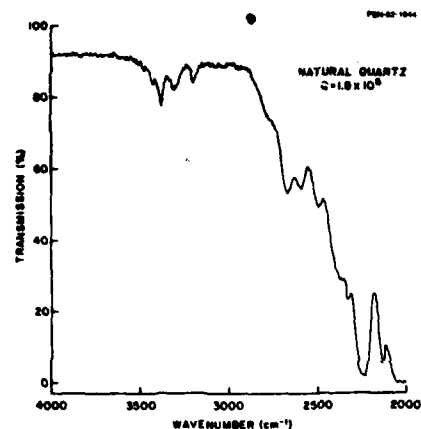
The first polishing technique (referred to as Polish A) was simple, going directly from a 220-diamond saw blade cut finish to Syton polishing for a half hour on each surface. Syton is a chemical-mechanical polish consisting of a colloidal suspension of SiO_2 in NaOH. The second polishing technique (referred to as Polish B) was more complex. Starting with a similar 220-diamond saw finish, each surface was ground down using successively finer grit sizes of 35, 12, 9, 5, and 3 microns of alumina oxide. Next a two-step polishing sequence was employed, first using cerium oxide and then finishing with Syton. The duration of the cerium oxide and Syton polishes varied, depending upon the grade of quartz, from 2 to 5 hours for cerium oxide and 1/2 to 2 hours for Syton. Both Polish A and Polish B samples had slightly bevelled edges. For both types of samples, the mechanical wheels used were randomly varied among the different types of quartz in order to avoid any systematic errors resulting from usage of a particular wheel on a particular type of quartz.

AD P002462

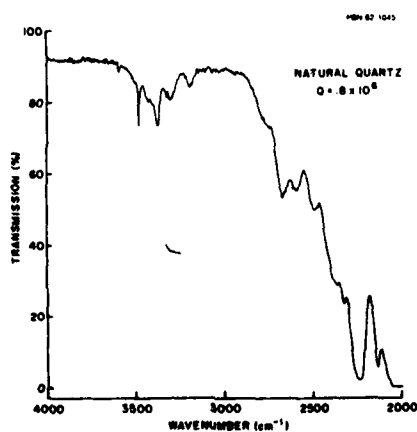
Infrared Absorption Characterization

The literature shows that there is an inverse correlation between dislocation density and electromechanical Q of quartz. High Q crystals tend to have fewer impurities than low Q crystals. There is also a direct correlation between Q and infrared absorption. Hence, infrared measurements were a convenient method for determining the Q and, therefore, the material quality of the quartz discs. All of the synthetic quartz ordered originated from Sawyer Inc., with a quoted minimum Q of 1.3×10^6 , 1.8×10^6 , 2.2×10^6 for optical, electronic, and Premium Q grade, respectively.

A Perkin-Elmer 580 B Infrared Spectrophotometer was used with a selected resolution of 1.4 cm^{-1} and the plotted wavenumber spectrum from 4000 to 2000 cm^{-1} . Examples of plotted infrared spectrum measurements of

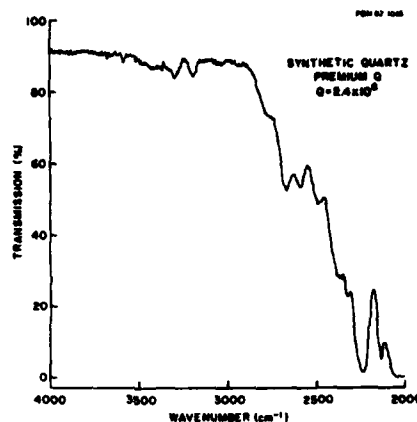


(a) Natural quartz



(b) Natural quartz (note that this spectrum is quite different from above)

Figure 1. Infrared transmission spectra for two samples of natural quartz and one of synthetic quartz: Note that sample-to-sample variation in natural quartz can be quite large, whereas all synthetic samples of a given quartz grade were quite similar.



(c) Synthetic quartz

natural and synthetic quartz are shown in Fig. 1. With natural quartz, great differences were seen in the level of absorption at a given band. This indicates that the impurity content is distinctly variable from sample to sample. This is not surprising considering the different origins, growth conditions, and environmental factors forming each natural quartz stone out of which one to two quartz discs were cut. Because of the large variations, infrared spectrum measurements were made on each natural quartz disc. For synthetic quartz, little variation was seen in the infrared spectrum from disc to disc within a particular grade, hence only a sampling of the disc were measured.

The equation used for converting infrared transmission levels into Q was obtained from a paper by Philips Research Labs.² The basic equations used are as follows:

$$\alpha_{3500} = \frac{1}{t} \left(\log \frac{T_{3800}}{T_{3500}} \right)$$

$$\alpha_{3500}^* = \alpha_{3500} + 0.25$$

Where

α_{3500} = Extinction coefficient at wavenumber 3500 cm^{-1}

α_{3500}^* = Extinction coefficient with the correction factor shown above.

t = Thickness of sample in centimeters.

T_v = Fraction of incident light of wavenumber v transmitted by sample

$$Q = 1.35 \times 10^5 / \alpha_{3500}^*$$

Two other calculation methods for Q were assessed where one used a 3410 cm^{-1} line² and the other used a different equation.³ They all resulted in somewhat different values of Q for a sample. However, the equation above generally gave a Q between the high and low values.

For natural quartz, the Q varied from 0.3 to 4×10^6 . Most values fell into the range from 0.7 to 1.1×10^6 . Synthetic samples nearly always satisfied the minimum Q quoted. A major limitation on the accuracy of the infrared measurements and hence the calculated Q was the thickness of the quartz discs. The literature² recommends at least 7-mm thick samples to provide sufficient infrared absorption for high Q material. Although our discs were only 2.54 mm thick, the low Q of most of the natural quartz samples provided high enough absorption that accurate measurements could be made. For the higher Q material, which included a small fraction of the natural and all of the synthetic quartz, two or three discs were stacked together to provide improved accuracy.

Biaxial Flexure Test and Experimental Conditions

Though there are a number of techniques for measuring tensile fracture strength, the biaxial flexure test⁴ was chosen, since this technique has been used for years at Raytheon to measure the strength of optical and ceramic materials and also a large number of samples can be easily handled with this method. In this test, a 1-inch circular sample is supported below by a somewhat larger ring and is loaded above by a ball (5/8 inch diameter) with a machined flat surface (1/2 inch diameter). For an isotropic material, this creates a large area of uniform tensile stress directly under the area loaded by the ball.⁵ The bottom surface of the specimen is under tension. Once loading and specimen deformation begins, the upper loading occurs along a ring having the diameter of the flat region of the ball; hence, edge preparation is not of major importance, so that sample preparation is relatively easy. Figure 2 shows the load fixture and Fig. 3 shows the Instron universal testing instrument on which the tests were conducted.

The fracture strength of the quartz discs was determined by loading the samples until fracture occurred. The load force at

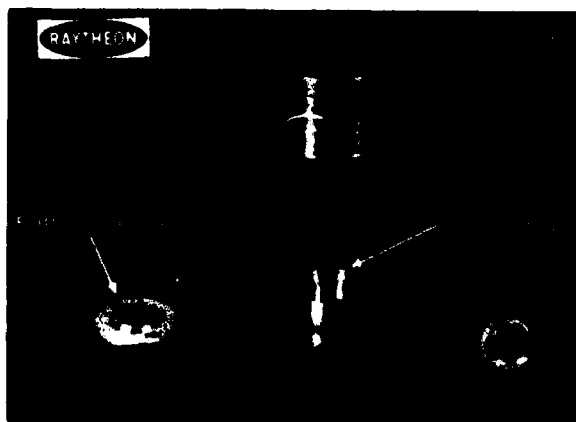


Figure 2. Load fixture for biaxial flexure test.

CE 66277



Figure 3. Instron universal testing instrument used for loading the quartz discs.

fracture was recorded and this was used to calculate⁵ the tensile stress from the equation below.

$$\sigma_r = \frac{3}{2\pi} \left[(1-\gamma) \frac{a^2 - r_o^2}{2b^2} + (1-\gamma) \ln a/r_o \right] \frac{P}{t^2}$$

Where

- γ = Poisson's ratio
- a = Radius of supporting ring
- r_o = Radius of flat on load ball
- b = Radius of quartz disc
- P = Load force (in pounds)
- t = Sample thickness (in inches)

The values used for the various parameters are: $a = 0.420$ inch, $r_o = 0.25$ inch, $b = 0.50$ inch, and $\gamma = 0.16$. The value used for Poisson's ratio was that of fused quartz. This obviously is not the correct value for single-crystal quartz, but the anisotropy of crystalline quartz means there is no single value. The above equation for tensile stress is not strongly dependent on γ , however, so the exact choice of a value for γ is not critical. A 25 percent change in γ causes only a 4 percent change in σ_r .

It was impossible to conduct the tests in a fully controlled environment, but temperature and humidity were recorded during each test to provide for evaluation of any possible correlations between these parameters and the observed fracture strengths. The tests were conducted at room temperature which, through the course of the tests, showed only a small peak-to-peak variation of about 4°F. The relative humidity, however, ranged from 46 percent to 72 percent over the duration of the tests. A special test designed to reduce the effect of humidity on tensile strength was performed and will be discussed in the next section.

During the course of the fracture tests on the first two samples, it was necessary to stop the application of increasing load in order to make adjustments in the recording equipment. In both cases the samples fractured after several tens of seconds at a high but constant load. This phenomenon is known as delayed fracture and results from a slow crack propagation velocity. Since this introduced a time-dependent variable into the fracture strength, great care was taken to ensure that all subsequent tests were made at a fixed load rate. This rate was 1300 lb/min which gave a typical time from the start of the test to fracture in the range of 15 sec to 45 sec.

In order to prevent the introduction of another variable, all of the samples (with the exception of some with strain gauges) were placed on the support ring with their x axis pointing in the same direction. Also, since very high load forces (350 to 1000 lb) were required to fracture the discs, some damage to the support ring was incurred during each test. To prevent this from becoming a factor in the fracture strength, the support ring was dressed after every seventh test. No significant damage was incurred on the load ball throughout the tests.

A final comment on the experimental procedures: one assumption of the biaxial

stress test is that once the sample bends under the load force, the only region of contact between the flatted load ball and the sample surface is at the outer edge of the flatted area. Since quartz is a very hard and brittle material, it was decided to test this assumption by replacing the flatted area with a ring of the same diameter for a few tests. Two samples were broken in this manner and their fracture strengths were completely in line with other samples fractured in the normal manner. Therefore, all remaining discs were fractured with the flatted load ball.

Fracture Results from Main Group

The main group of samples consisted of 50 discs prepared as previously discussed. Of these 50 discs, 15 were natural quartz, 14 optical grade, 14 premium Q, and 7 electronic grade. Polish B was used on 22, while the remaining 28 had Polish A. This group of 50 samples provides the main body of data from which the effect of grade of quartz and type of polish could be evaluated. It also provides a baseline to which some special cases could be compared. These special cases (involving 30 additional discs) will be discussed in the next section.

Table 1 shows the average fracture strength (in psi) and standard deviation for each of the seven combinations of the type of quartz and the polishing technique as parameters that were tested. No electronic grade samples with Polish B were used. The far right column and bottom row show the combined results from the columns to the left or the rows above, respectively. During the fracture tests, the samples were broken in groups of seven (one sample for each parameter) so that any systematic variable could be minimized. The average fracture strength for the entire population was 21,400 psi, but two other characteristics are quite noticeable. One is that the standard deviation is quite large (~ 30 percent), and that any differences in fracture strength among the seven

TABLE 1
FRACTURE STRENGTH OF ST CUT QUARTZ DISCS

	Polish B	Polish A	Polish A & Polish B
Natural	(8) 21,600 ± 7,600	(7) 20,900 ± 5,200	(15) 21,400 ± 6,300
Optical	(7) 24,400 ± 6,400	(7) 23,200 ± 8,100	(14) 23,800 ± 7,100
Premium Q	(7) 19,400 ± 4,200	(7) 20,100 ± 4,400	(14) 19,700 ± 4,200
Electronic	--	(7) 20,400 ± 6,100	(7) 20,400 ± 6,100
Total	(22) 21,800 ± 6,300	(28) 21,100 ± 5,900	(50) 21,400 ± 6,000

(N) = Number of samples. All values in psi.

TOTAL POPULATION (N=50) 21,400 ± 6,000 psi

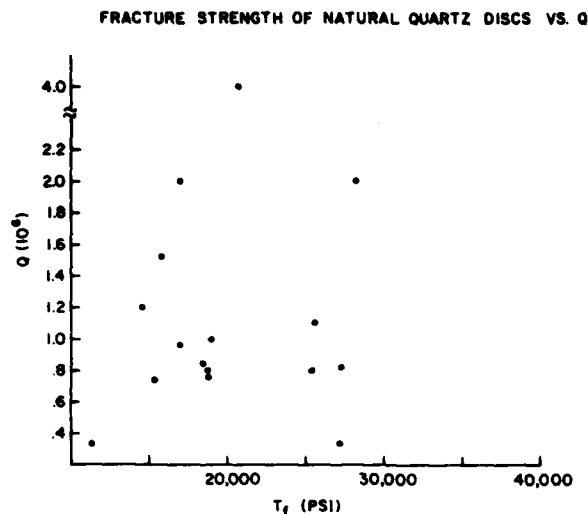


Figure 4. Fracture strength of natural quartz discs relative to their measured Q values.

parameters are small compared to the standard deviations. Virtually no difference existed between Polish A and Polish B samples. More variation among the types of quartz was observed, but even between the strongest (optical) and weakest (Premium Q) the difference was only 4,100 psi. This is only slightly more than $2/3$ of the average standard deviations. Also, no correlation was observed between the measured Q of the natural quartz and the fracture strength. This is shown in Fig. 4.

In view of the large standard deviations, it is of interest to investigate the distribution of fracture strengths. Figure 5 shows the distribution for the entire population of 50 discs. As can be seen, the distribution is not symmetric but has a tail which stretches up to nearly 40,000 psi. Yet, the peak of the distribution lies somewhere between 15,000 and 20,000 psi. The ratio of the strongest to weakest sample is a very large 2.85. Figures 6 and 7 show the fracture strength distribution for the Polish B and Polish A samples separately. Both show the same general shape as the distribution for the entire population. In these figures, the grade of quartz for each data point has been indicated by a letter (N = natural, O = optical, P = premium Q, etc.)

In addition to the fracture strength of each disc, seven of the first samples were broken with strain gauges on them. The strain gauges were located on the bottom side of the discs (the side under tensile stress). Four were oriented so as to measure the strain along the x axis, while the other three were oriented to measure the strain along the direction perpendicular to the x axis. The strain at fracture ranged from 1.2×10^{-3} to 1.6×10^{-3} for seven samples. In the x direction the average ratio of stress to strain was $17.9 \pm 1.0 \times 10^7$ psi.

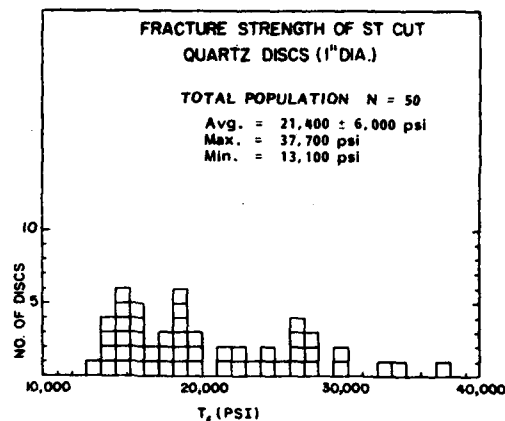


Figure 5. Distribution of fracture strengths for the main group of 50 discs.

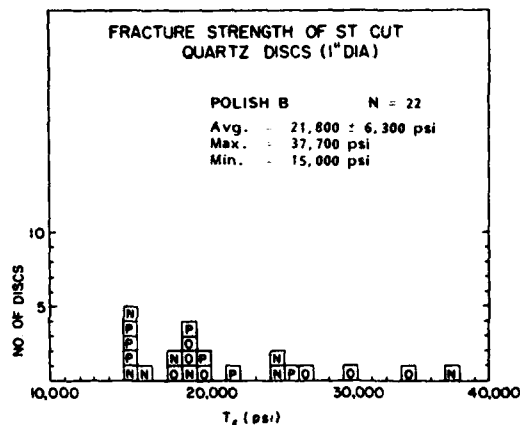


Figure 6. Distribution of fracture strengths for the Polish B discs. The grade of quartz is shown for each sample; N = natural, O = optical, P = premium Q.

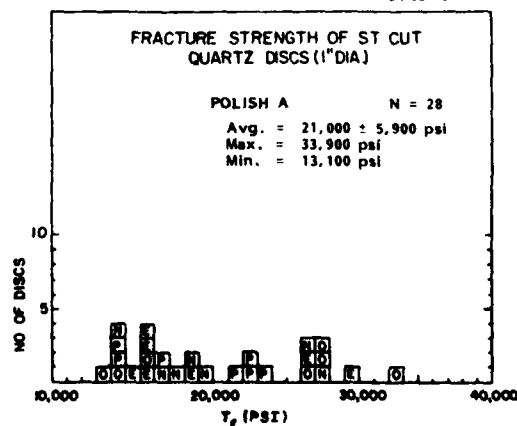


Figure 7. Distribution of fracture strengths for the Polish A discs. The grade of quartz is shown for each sample; N = natural, O = optical, P = premium Q, E = electronic.

In the direction perpendicular to the x axis, the ratio was $17.5 \pm 0.9 \times 10^7$ psi.

Some observations made during the course of the tests led to minor changes in test conditions. One observation was that the average fracture strength of the 7 discs with strain gauges was 25,000 psi while the next 14 discs, without strain gauges, averaged 21,000 psi. This led to some concern that the strain gauges effected the fracture strength. To check this, the next 7 discs were broken with a piece of thin plastic, similar to a strain gauge, fastened to the disc in the same manner as a strain gauge. The average strength of these 7 discs, however, was found to be 21,400. This seemed to rule out the strain gauges as a cause for the high strength of that particular group.

Since the relative humidity was increasing through the course of the fracture tests, it was also decided to evaluate the effect of humidity on fracture strength. Therefore, a special test was arranged for the final 14 of the 50 discs. Seven were broken as usual, but the other seven were baked at 175°C for 4 hours to drive off absorbed water. While hot, one surface was coated with a silicon grease and then the discs were stored in dry nitrogen (at room temperature) until immediately before they were to be fractured. The greased side was placed down so as to be under tension. Though the humidity was high during these tests, both groups showed above-average strengths and little difference was observed between the greased and ungreaed samples. Since neither the dummy strain gauges or the silicone grease had any significant effect on fracture strength, these parameters were not called out in Table 1. In hindsight, it appears that our concerns about systematic effects were caused by statistical fluctuations related to the large standard deviation.

The statistical theory of brittle materials is commonly based on the Weibull probability distribution. The relationship between probability of failure (P_f) and applied stress is:

$$P_f = 1 - \exp \left[- \left(\frac{\sigma}{\sigma_0} \right)^m \right]$$

where σ is the maximum stress at fracture and σ_0 and m are material constants. The Weibull modulus is represented by m . Figure 8 shows a plot of the Weibull distribution for our samples.

The estimation of the Weibull parameter m is made by using the maximum likelihood method in \ln - \ln space. The Weibull distribution expressed in terms of probability of survival is:

$$P_s = \exp \left[- \left(\frac{\sigma}{\sigma_0} \right)^m \right]$$

Taking the logarithm of both side yields:

$$\ln P_s = - \left(\frac{\sigma}{\sigma_0} \right)^m$$

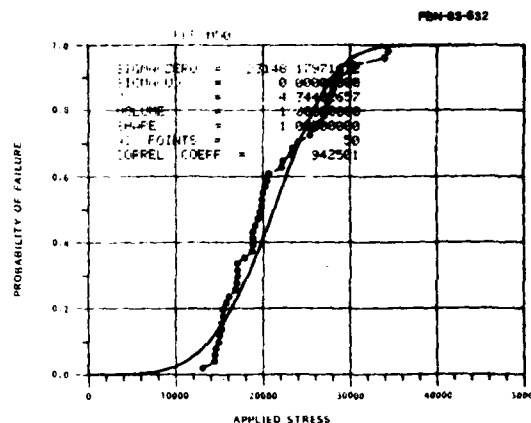


Figure 8. Weibull distribution for the total population of 50 samples tested.

Similarly:

$$\ln(-\ln P_s) = m \ln \sigma - m \ln \sigma_0.$$

Note that $\ln(-\ln P_s)$ is a linear function of $\ln \sigma$ where m is the slope and the intercept is $(-m \ln \sigma_0)$. Figure 9 shows a plot of $\ln(-\ln P_s)$ vs. $\ln \sigma$. Different Weibull modulus can indicate that more than one fracture mechanism is occurring. From Fig. 9 more than one Weibull modulus is apparent and can indicate that one type of fracture mechanism occurs for low fracture strength, and another fracture mechanism occurs for high fracture strengths. This may explain the large standard deviation that was observed.

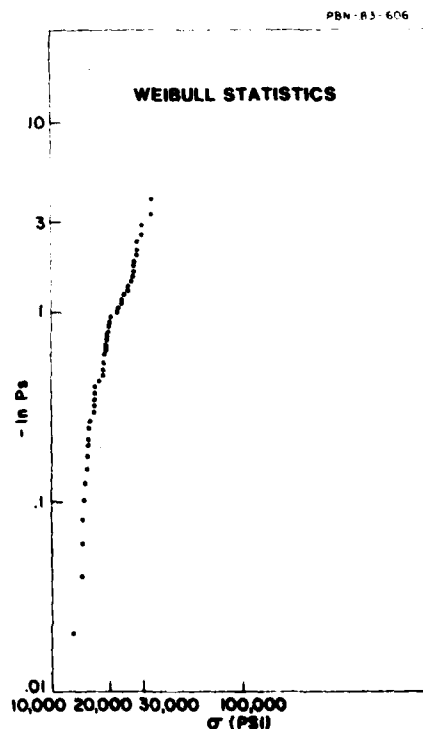


Figure 9. \ln - \ln plot of $-\ln P_s$ vs. σ where the slope is the Weibull modulus.

As a final step in evaluating the main group of discs, seven discs were reassembled in order to study the fracture patterns. When broken, the quartz shattered and the higher-strength discs disintegrated into far too many pieces to attempt reconstruction. Therefore, only the weaker discs could be reassembled. Figure 10 shows the fracture patterns of four of the seven samples along with the corresponding fracture strengths (in psi). A great deal of similarity is seen in the patterns among the four samples. All show a central section with fracture lines running completely around it and usually through it. The diameter of the central section is roughly the same as that of the flat on the load ball. In some cases, some of the fracture lines on different samples appear to be the same crystallographic planes. Outside of the central section the fracture lines tend to run radially, giving a flower petal appearance. Generally it was not possible to tell where the fracturing originated, but in one sample the central section was cracked but still intact. Obviously, for this case, the fracturing did not originate in the central section. This was unexpected since that is precisely where it is supposed to start in a biaxial stress test. This observation, in combination with the very similar fracture patterns in all seven of the reassembled samples, has led to the possibility that the fracturing of all of the samples may have started near the edge of the flat on the ball or perhaps even near the support ring. A possible explanation for this is the affect of anisotropy of a single crystal. The stress analysis for the biaxial stress test was done for isotropic materials, such as ceramics, and would not be accurate for most single crystal materials. A complete analysis for ST cut quartz discs was beyond the scope of this program, but it may show high tensile stress levels on the bottom surface, directly under the edge of the load ball. This further increases the uncertainty in the calculated values of the fracture stress. Therefore it would be best to assume that these values are a lower limit and the frac-

CE 66278

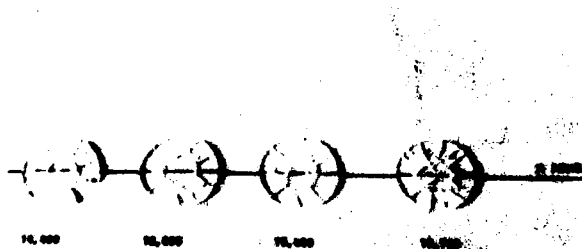


Figure 10. Fracture patterns for four quartz discs along with their measured fracture strengths. Each disc is oriented such that its x axis is parallel to the indicated line.

ture stress may in fact be higher. The present questions about the accuracy of the calculated stress values should in no way affect the results of the comparison among different grades of quartz and types of polish.

Fracture Results from Special Groups

Thin Discs

In this special group, eight quartz discs consisting of two each of X cut, Y cut, AT cut and BT cut plates were used. The discs were 1.25 inch in diameter, 0.05 inch thick, and were polished on one side. It was determined that the discs were at least 10 years old, but the source, grade of quartz, and type of polish were unknown. The average fracture strength of the eight discs was found to be $22,110 \pm 6160$ psi, which is very consistent with the results from the main group. This was reassuring since these discs were considerably different in both geometry and history from the main body of samples.

The two X cut discs had the highest average strength ($\sim 26,700$ psi) of the four different cuts, but a population of only two is insufficient to draw any conclusion about the relative strength of the various cuts.

Unpolished Discs

In order to compare the relative strengths of polished and unpolished discs, five unpolished quartz discs were broken. Of the five, four were natural quartz with a 120- μ m lapped surface finish, and the average tensile strength among the four was 9200 ± 440 psi. The one remaining disc was an optical-grade quartz with a 220-diamond saw blade finish which broke at 7,600 psi. By comparison, the average tensile strength of polished discs was $21,400 \pm 6000$ psi. These experiments indicate, not surprisingly, that unpolished quartz is substantially weaker (by a factor of two to three) than polished quartz, and therefore that polishing is an important part of maximizing the strength of a given configuration.

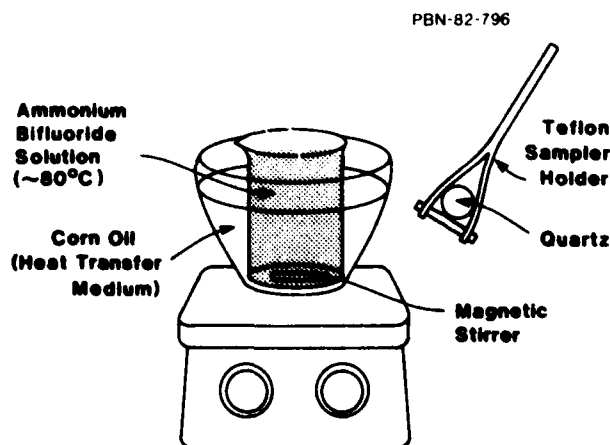


Figure 11. Experimental setup used for chemically etching quartz discs.

Chemically Polished Discs

Chemical etching of natural quartz discs in a saturated solution of ammonium bifluoride can produce chemically polished surfaces. Interest was generated because very thin chemically polished discs are shown to be extremely strong.⁶ The etching experiments were performed in a double beaker arrangement, as shown in Figure 11. An outer glass beaker contains corn oil, which proved to be a better heat-conducting medium than water because of its slower evaporation rate and higher boiling point. The inner teflon beaker contains a saturated solution of ammonium bifluoride with a composition of 65 gms ammonium bifluoride flakes per 100 ml of solution. The ammonium bifluoride solution is heated to about 80°C, and is constantly agitated with a stirring bar. The rate of etching depends primarily on the temperature of the solution where the higher the temperature the faster the removal of material from the substrate. A teflon sample holder is used to hold the quartz disc in the etching solution.

Prior to etching, all discs underwent a thorough cleaning to remove surface contaminants such as grease, which may inhibit etching. The cleaning procedure consists of ultrasonic agitation in a detergent solution followed by thorough rinsing with TCE, acetone, and methanol.

Both natural and synthetic quartz discs, with either Polish A or Polish B surfaces, were chemically etched. The rate of removal from both surfaces varied from 1.5 to 2.2 mils per hour dependent upon temperature and amount of depletion of the etching solution. A total

of 4 to 6 mils was typically etched off to ensure complete removal of surface layers damaged by mechanical lapping and polishing processes.

After completion of the chemical polishing, the surface topography was examined. Two commonly observed surface defects are etch pits and etch channels, as shown in Fig. 12. At the surface end of an etch channel, etch pits are always found; however, etch pits are not always associated with etch channels. The exact mechanism for generation of etch pits and channels is unclear. Etch channels are likely due to dislocations in the crystalline lattice. Bulk and/or surface defects and impurities in the quartz may lead to etch pits. One piece of evidence for a bulk-related mechanism was seen in a sample where etch pits were heavily concentrated in corresponding areas of the surface on both sides of the disc. A surface-related mechanism instead would have resulted in etch pits evenly distributed across only one surface, since the whole disc surface should have undergone the same preparation and polishing conditions.

Large numbers of etch pits and channels were observed in synthetic grade quartz. Hence, chemical polishing using ammonium bifluoride proved to be unsuitable for synthetic quartz. Though the amount of etch pits and channels varied from sample to sample, all chemically etched natural quartz surfaces were of much better quality than chemically etched synthetic quartz surfaces. Of the ten chemically etched surfaces, the average fracture strength was $14,400 \pm 6500$ psi, which proved to be substantially weaker than the regularly Syton polished surfaces of strength $21,400 \pm$

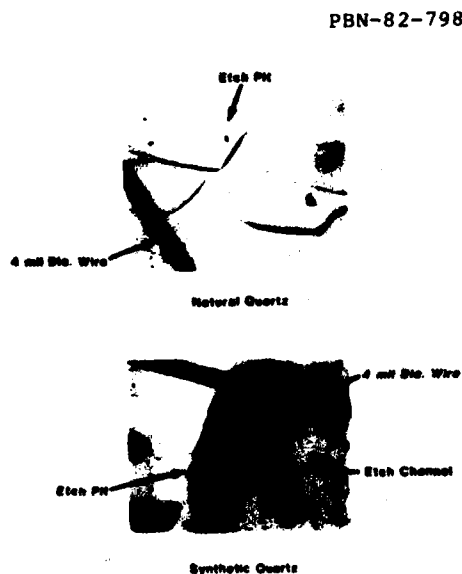


Figure 12. Photographic comparison of surfaces of natural and synthetic quartz disc after chemical etching.

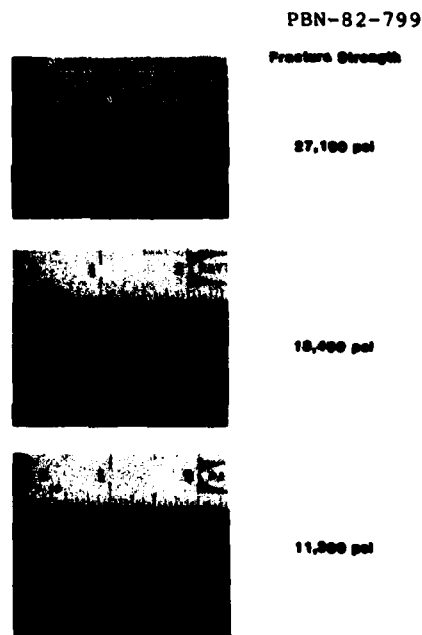


Figure 13. Three chemically polished natural quartz discs, showing the correlation between fracture strength and surface quality after etching.

6000 psi. This weakening of tensile strength due to chemical etching is contrary to results noted in the literature.⁶ One important observation was that, among etched samples, the better quality surfaces with little or no etch pits had a higher fracture strength of $20,700 \pm 7500$ psi. Although this is not stronger than the regularly polished Syton finish surfaces, it is substantially stronger than the average strength of the general population of chemically polished samples where the surfaces were full of etch pits and channels. Hence, etch pits and channels do cause weakening of the fracture strength. Shown in Fig. 13 is this observation that the fewer the etch pits and channels on the surface, the higher the fracture strength. An interesting possibility may be first to chemically etch the surface and then finish with a slight Syton polish to smooth out the etch pits.

Discs with Ion Etched Grooves

Placing ion-milled grooves onto the surface of quartz discs and then testing for tensile strength would show the effect of the grooves on the disc structure and show whether the grooves would substantially weaken the discs.

Seven quartz discs had grooves ion-milled onto the surface using standard photolithographic techniques. The quartz is coated with photoresist. Then the photoresist is exposed through a mask with UV light. A mask with 1600 2.5 microns wide grooves was used. Following the exposure, a development step removes the exposed photoresist and hence bares the surface in the groove regions. An ion-miller is used to etch down 1000 Å into the bare substrate, and the rest of the substrate is protected by the photoresist.

The average fracture strength was $19,300 \pm 2700$ psi, which is comparable to regular unmilled surfaces with strengths of $21,400 \pm 6000$ psi. Hence, at least for discs, ion-milled grooves on the surface did not substantially weaken the tensile strength.

Conclusion

The major emphasis of this study was to compare the fracture strength of four different types of quartz including natural and

three synthetic grades (optical, premium Q and electronic). A biaxial flexure test was used for measuring the tensile strength on quartz discs which were 1 inch in diameter and nominally 0.1 inch thick. The average strength of polished samples was 21,400 psi and this was independent of both quartz grade (natural quartz and three distinct synthetic grades) and polishing details. Unpolished samples, however, fractured at approximately 9,000 psi indicating that polishing is indeed a critical processing step, increasing material strength by a factor of 2 to 3. Chemical etching was also evaluated, but caused etch pits in natural quartz which actually reduced average strength. The ion-beam milling of extremely shallow grooves (1,000 Å deep) on the disc surfaces did not significantly alter material strength. In addition to the parameters mentioned above, other factors such as relative humidity, Q (from infrared measurements), and sample thickness also were found to have no correlation with fracture strength.

References

1. H.E. Karrer and J. Leach, "A Quartz Resonator Pressure Transducer," IEEE Trans. on Industrial Electronics and Control Instrumentation, IECI-16, 44 (1969).
2. J.C. Brice and A.M. Cole, "The Characterization of Synthetic Quartz by Using Infra-red Absorption," Proceedings of the 32nd Annual Symposium on Frequency Control, pp. 1-10, 1978.
3. Technical Brief, Material Specification for Cultured Quartz, No. 7, Sawyer Research Products, Inc., Eastlake, Ohio.
4. J.B. Wachtman, Jr., W. Capps, and J. Mandel, "Biaxial Flexure Tests of Ceramic Substrates," Journal of Materials, 7, 188 (1972).
5. T.R. Wilshaw, "Measurement of Tensile Strength of Ceramics," Journal of the American Ceramics Society, 51, 111 (1968).
6. J.R. Vig, J.W. LeBus and R.L. Filler, "Chemically Polished Quartz," Proceedings of the 31st Annual Symposium on Frequency Control, pp. 131-143, 1977.

EFFECT OF ALKALI IONS ON ELECTRICAL CONDUCTIVITY AND DIELECTRIC LOSS OF QUARTZ CRYSTALS*

J. Toulouse, E.R. Green and A.S. Nowick
Henry Krumb School of Mines
Columbia University, New York 10027

Summary

Because of the importance of alkali ions in relation to frequency stability of quartz resonators, we have investigated the electrical properties of various Li-swept and Na-swept quartz crystals (both natural and synthetic). These properties are of two types: dielectric loss measurements, particularly at cryogenic temperatures, and electrical conductivity measured between 100° and 400°C. In the case of the dielectric measurements, we studied the two well known loss peaks due to Al-Na pairs and established a calibration for the lower temperature peak ($\tan \delta_{\max}$ of 1.5×10^{-5} corresponding to 1 ppm of Al-Na pairs). High quality synthetic crystals and a natural crystal were found to have primarily Li as compensation for Al in the as-grown condition. The absence of an analogous Al-Li peak leads to the conclusion that Li resides on the 2-fold symmetry axis of the AlO_4 tetrahedron to which it is bound. Following irradiation, large loss peaks were found, at 8.4 K for Na-swept and at 6.85 K for Li swept crystals. These peaks may be useful to characterize quartz crystals.

The ionic conductivity is different for Na and Li in the same crystal, and also for synthetic and natural crystals. Conductivity is controlled by a combination of the energy of migration of free M^+ and of association of the Al-M pair.

Introduction

It is well established that frequency instabilities in quartz resonators, especially after irradiation, are related to impurities in the crystals.¹⁻⁴ Probably the most important impurity is Al^{3+} substituting for Si^{4+} , which (due to its lower valence) requires an additional defect for charge compensation. The latter is most commonly an interstitial monovalent ion, M^+ , notably the alkalis, Li^+ and Na^+ , or hydrogen. Because of the Coulombic attraction between the M^+ and the Al^{3+} (which carries an effective charge of -1), the M^+ interstitial is usually located adjacent to the Al^{3+} ion at low to moderate temperatures.

Two types of electrical properties provide useful ways to observe the effects of alkali ions; these are dielectric loss, and conductivity. The pioneering work of Stevels and Volger⁵ showed that dielectric loss (or "dielectric relaxation") observed at cryogenic temperatures showed a number of peaks, two of which could be attributed to Al-Na pairs. Subsequent work in this laboratory^{6,7} showed that a Na^+ ion can occupy two pairs of equivalent sites (a total of 4 sites), which may be considered nn and nnn positions, about a given AlO_4 tetrahedron. An electric field gives rise to dielectric loss by producing preferential reorientation between an equivalent pair of sites, the lower temperature loss peak being due to nn reorientation and the higher temperature one to nnn reorientation. Analogous acoustic loss (or "anelastic relaxation") peaks are

also known.⁸

An important question is whether there are similar loss peaks due to Al-Li pairs. Such a peak was claimed by Stevels and Volger⁵ to be present at 60 K (for a frequency of 32 kHz). There is a need to confirm this claim, since such a peak would provide a valuable method for characterizing the Li^+ content of quartz crystals.

The second property of interest is electrical conductivity. Because of the large band gap of crystalline quartz (> 8 eV) such conductivity is entirely ionic, and is due to alkali ions that have broken away from Al-M pairs as a consequence of the dissociation equilibrium. A detailed study by Jain and Nowick⁹ has shown some interesting differences between natural and synthetic crystals, in particular, a lower activation energy for natural crystals.

Both electrical properties are strongly dependent upon the morphology of the alkali defect, which can be modified by irradiation.¹⁰ Therefore we have also begun to investigate the effects of X-irradiation upon dielectric loss and conductivity.

In order best to study these electrical properties related to alkalis, it is desirable to obtain crystals that have Al compensated by just one alkali at a time. This objective can be met by the process of electrodiffusion (or "sweeping"), i.e. application of an electric field parallel to the z-axis to introduce the desired interstitial ion into the lattice replacing those that were grown in. This method was previously developed by Kats¹¹ and by King¹ and is now a relatively standard technique. In addition to alkalis, H^+ can also be swept into the crystal. By studying the electrical properties of Li-, Na- and H-swept crystals we hope to achieve a better understanding of these properties than had been achieved in the past with measurements on various as-grown crystals.

Experimental

The principal synthetic crystals studied were high quality crystals taken from the Z-growth region: Toyo Suprem Q (bar SQ-A) and Sawyer Premium Q (bar PQ-E). The natural crystal was a clear crystal from Arkansas.

Electrodiffusion experiments were carried out at Oklahoma State University by Dr. J. Martin. (Method described in ref. 12). Dielectric loss ($\tan \delta$) and ionic conductivity were both measured with a General Radio type 1620A Capacitance Bridge assembly over the frequency range 20 Hz-100 kHz. Complex impedance analysis was used to obtain the bulk conduction.⁹ The samples used were plates of surface area 1 cm^2 and thickness 1.0-1.5 mm coated with sputtered silver electrodes. Most samples were cut so that the electric field is applied parallel to the z-axis, but some were

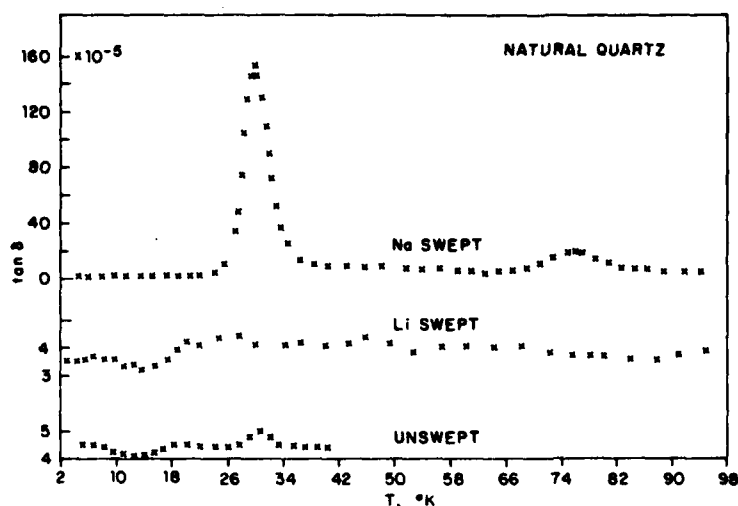


Fig. 1. Dielectric loss of natural quartz crystals that were unswept, Li-swept and Na-swept. Note that the scale for the two lower curves is 20 x larger than that for the upper curve.

cut in a perpendicular orientation. The sample chamber for the conductivity measurements was described previously.⁹ For dielectric loss, the sample was inserted into a Janis Supravertemp cryostat, and measurements were taken between 2.9 and 300 K. The leads from the bridge to the sample electrodes were shielded by a continuous coaxial shield connected to the bridge ground.

X-irradiation was carried out for 4 hours at room temperature using a conventional tungsten-filament tube at 40 kV and 20 mA. The very soft X-rays were filtered out by the layer of sputtered silver used as electrodes.

Dielectric Relaxation Studies

The Al-Na Peaks

In several samples containing Na, we have observed the pair of peaks previously reported by Stevels and Volger⁵ and by Park⁷ as due to Al-Na pairs. At a frequency of 1 kHz these peaks appeared at 30 K and 75 K. An example is given in Fig. 1 which compares dielectric loss ($\tan \delta$) for the natural crystal as received (unswept) and after Li- and Na-sweeping. The two peaks are prominent in the Na-swept sample, very small in the unswept and absent in the Li-swept sample. The 30 K peak is especially interesting as a very sensitive measure of the Al-Na content just as is the 50 K anelastic peak (at 5 MHz).^{8,13} In fact, by comparing the temperature location of the dielectric peak at 1 kHz with that of the anelastic peak at 5 MHz, one obtains $H = 0.052$ eV as the activation enthalpy of the peak. Figure 2 shows a more detailed examination of the peak shape, comparing normalized data for the natural and the Toyo synthetic crystal. It is striking that the shapes are so nearly identical (except for a possible small tail on the high $1/T$ side for the synthetic crystal) in spite of the great difference in origins of these two crystals and a peak height that is 7x larger for the natural crystal. From the peak width of Fig. 2 an apparent activation energy of 0.050 eV is obtained, showing that the peak is only 4% wider than a perfect Debye peak.¹⁴ This relaxation peak is attributed^{6,7} to reorientation between the two equivalent nn configurations of the Al-Na pair. A similar agreement in peak shapes between natural and synthetic is found for the 75 K peak, which is attributed to nnn Al-Na pairs.

The study of the Toyo Supreme SQ-A sample allows a calibration of the 30 K peak. Halliburton et al.^{15,10} developed a method for determining the Al content of a quartz crystal by converting all of it to Al-hole (Al-h^+) centers through an appropriate irradiation procedure; these centers could then be detected and their concentration determined through ESR measurements. In this way, they found, for an SQ-A specimen from the same bar as our sample, a value of 14.4 ppm of Al. If we assume that our Na-swept SQ-A sample contained the same Al content and that the Na sweeping converted all Al to Al-Na defects, then our peak height $\tan \delta_{\text{max}}$ of 22.1×10^{-5} must correspond to 14.4 ppm of Al-Na defects. This result then gives a calibration constant of $1.5 \times 10^{-5}/\text{ppm}$.

The earlier theoretical treatment⁶ permits us to interpret this calibration constant. It was shown that the relaxation strength $\delta \epsilon_{\parallel} / \epsilon_{\parallel}$ of the dielectric constant ϵ_{\parallel} parallel to the z-axis is given by

$$\frac{\delta \epsilon_{\parallel}}{\epsilon_{\parallel}} = 2 \tan \delta_{\text{max}} = \frac{n_d \mu_z^2}{kT} \quad (1)$$

where $\tan \delta_{\text{max}}$ is the measured peak height, n_d the number of dipoles (Al-Na pairs) per unit volume, μ_z the component of the dipole moment parallel to the z-axis and kT has its usual meaning. The above calibration constant then gives a value of $\mu_z \sim 5 \times 10^{-30}$ C-m, or a charge separation $r_3 \sim 0.3$ Å, which is reasonable.

Table 1 presents a summary of the results for the 30 K peak for a variety of samples that we have studied (all measured parallel to the z-axis). In the final column this calibration constant is used to convert peak heights into concentration of Al-Na pairs. It is interesting that these high quality synthetic crystals (SQ-A and PQ-E), grown with Li^+ in the mineralizer, show only a very small amount of Na in the as-grown condition. The total Al content of the PQ-E crystal, for example was ~ 18 ppm;¹⁰ therefore, only 4% of the Al^{3+} is associated with Na in the as grown condition. Similarly there is only 15% Al-Na in the Toyo SQ-A crystal. The remainder is almost certainly Al-Li, since there is no evidence in these crystals for infrared absorption due to Al-OH centers.¹⁶

Table 1. Results for the 30 K peak for various samples

Sample	Peak height $\tan \delta_{\max} \times 10^5$	ppm Na
SQ-A unswept	3.2	2.1
SQ-A Na-swept	22.1	14.4
PQ-E unswept	1	0.7
NQ unswept	1	0.7
NQ Na-swept	151	100

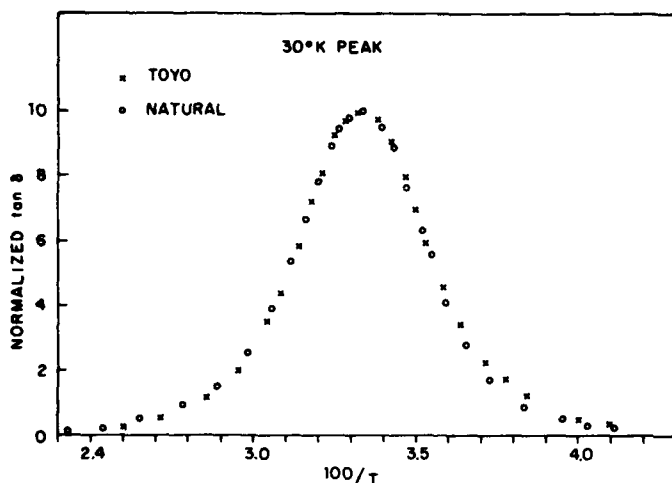


Fig. 2. The 30 K dielectric loss peak on a normalized plot for: \circ the natural crystal, \times the Toyo SQ-A crystal. (The absolute peak height of the natural is 7 x higher than that of the Toyo crystal.)

Search for Al-Li Peaks

In attempting to find one or more dielectric loss peaks due to Al-Li pairs, we have examined several unswept and Li-swept samples both parallel and perpendicular to the z-axis and over the temperature range 2.9 - 290 K. In all of these experiments no identifiable peak was found, in contrast to Stevels and Volger's claim of a peak at 60 K (for 32 kHz). The middle curve of Fig. 1 shows one of the best examples, that of a Li-swept natural crystal. In this case, for which the Al content is 100 ppm (see Table 1) there is no evidence for a peak as large as $\tan \delta_{\max} = 1 \times 10^{-5}$. (Even small fluctuations which might hide very small peaks in Fig. 1 were later shown, by remeasurement, not to be peaks.) Thus, using Eq. (1), if a Al-Li peak exists for measurements parallel to the z-axis, it must correspond to a value of μ_2 at least an order of magnitude smaller than that for Al-Na. Similarly, measurements with electric field in the basal plane, for which the corresponding relaxation-strength equation is⁶

$$\frac{\delta \epsilon_2}{\epsilon_1} = 2 \tan \delta_{\max} = \frac{n_d \mu_2^2}{2kT} \quad (2)$$

where μ_2 is the component of dipole moment in the y-direction of the crystal, again gave negative results, suggesting that μ_2 is also exceedingly small. It is noteworthy that anelastic relaxation, which involves the same relaxational mode⁶ as ϵ_2 , is also absent for Li-swept crystals over a similar temperature range.¹³

It is therefore concluded that the Al-Li pair is most likely oriented along the x-axis, i.e. the direction of the 2-fold (C_2) symmetry axis of the crystal, as shown in Fig. 3. Because Li^+ locates itself on this axis, the two equivalent sites between which reorienta-

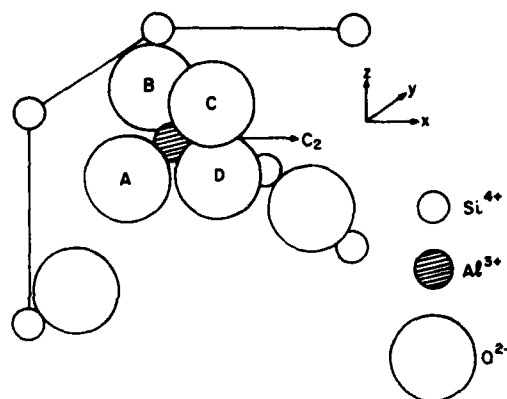


Fig. 3. Diagram showing part of the unit cell of α -quartz containing a substitutional Al^{3+} that forms an AlO_4 distorted tetrahedron. The two-fold symmetry axis, C_2 , which lies parallel to the x-axis, is also shown.

tion occurs in the case of Na^+ , have now collapsed into a single site; thus, the possibility for reorientation, i.e. for dielectric relaxation, no longer exists. [This last remark is just a restatement, in physical terms, of what is contained mathematically in Eqs. (1) and (2).] A possible explanation for the orientation of the Al-Li pair along the C_2 axis may be related to the small size of the Li^+ ion which may permit it to sit between two oxygen ions of the AlO_4 tetrahedron. On the other hand, the Na^+ ion is believed to prefer to form an O-Na type of bond with one of the four oxygen ions of the tetrahedron.⁷ One way to explore the defect configurations further is by computer simulations using the HADES II type program.¹⁷ Such methods have given insight into the minimum-energy configurations of different defect clusters in a variety of ionic crystals and may also be useful in the present case.

Irradiation Peaks

X-irradiation produces a large peak at very low temperatures, with a high rising background below the peak. Figure 4 shows the results for Toyo Supreme SQ-A samples that had been Li-, Na- and H-swept. In the Na-swept case, the 30 K ($100/T = 3.3$) Al-Na peak is considerably reduced and replaced by a huge peak at 8.4 K with a rising background at still lower temperatures. A similar peak is present for the Li-swept case, but it is located at a lower temperature, 6.85 K. The H-swept sample, on the other hand, shows a smaller less well-defined peak after irradiation, which might be interpreted as a composite of the residue of the peaks for the Li- and Na-swept samples.

It is reasonable to ascribe these radiation induced peaks either to the $Al-h^+$ center or to a center that involves the alkali. The fact that the peaks for the different alkalis have different temperatures and that the H-swept sample shows a lower and less well defined peak suggests the possibility of an alkali center. A simple $H-e$ center is not very likely, since such a center with a single electron would have a strong ESR spectrum, yet has not been detected in irradiated samples.¹⁸ A more complex alkali center may be involved; for example, it may contain two electrons and possibly an oxygen-ion vacancy. More work is clearly needed to further explore these interesting irradiation induced peaks as well as the very low-temperature background effect which accompanies them.

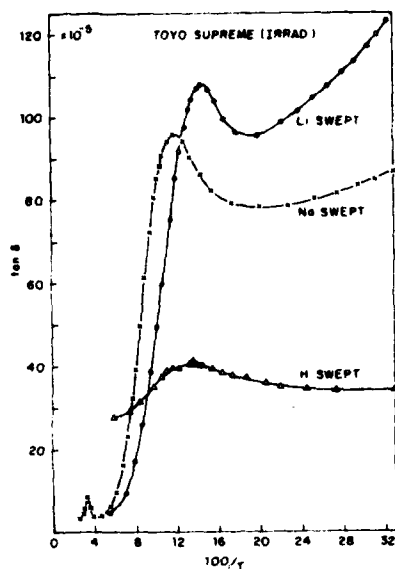


Fig. 4. Dielectric loss of three differently swept Toyo SQ-A samples following X-irradiation.

Conductivity Studies

Conductivity measurements were carried out on unswept and swept Toyo SQ-A and natural crystals. All measurements were made with electric field parallel to z-axis. The results, shown in Fig. 5, are somewhat different for the two types of quartz. For the natural crystal, the Li-swept shows a substantially higher conductivity and a slightly lower activation energy (see Table 2). On the other hand, for the Toyo synthetic crystal, the conductivities are quite close and the Li-swept has the higher activation energy. Also included in Fig. 5 are results for the H-swept Toyo crystal. Here the conductivity has dropped by more than two orders of magnitude, yet the activation energy is comparable to that of the alkali swept crystals. It is therefore reasonable to regard that the conductivity of H-swept sample is not due to H^+ migration, but to the small amount of residual alkali which remains after sweeping.

Table 2 also shows that the activation energies for both natural crystals fall well below those for the synthetics, as Jain had already observed.⁹ These differences were interpreted in terms of the relation.

$$E = E_m + \beta E_A \quad (3)$$

where E_m is the activation energy for migration of a free alkali ion, E_A is the Al-M association energy, and β can vary between $\frac{1}{2}$ and 1 with the value of unity for synthetic crystals, in which the defect concentrations are so small that additional unassociated Al may be present. In any case, the term E_A dominates the expression for E , so that a higher value for Li^+ than for Na^+ does not mean that the larger (Na^+) ion migrates down the open channels of the crystal more rapidly than the small Li^+ ion. Rather, because of the small size of Li^+ , it is possible that E_A for Al-Li is higher than that for Al-Na.

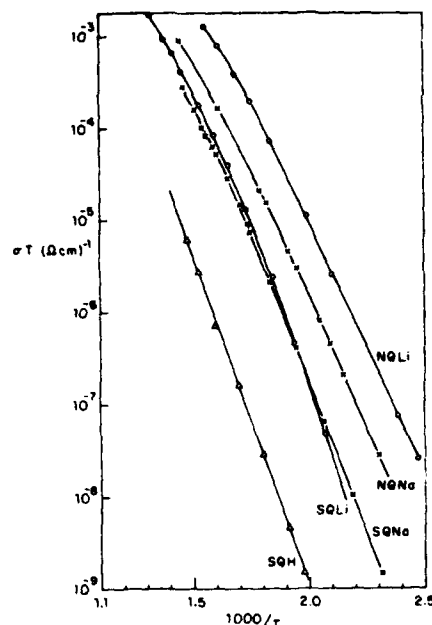


Fig. 5. Conductivity plots ($\log \sigma T$ versus $1/T$, where σ is the conductivity) for Li- and Na-swept natural quartz samples (NQ), and for three differently swept synthetic samples (Toyo SQ-A). The corresponding activation energies are given in Table 2.

Table 2. Activation Energies from Conductivity Measurements.

Crystal	Sweeping	E (eV)
Natural NQ	Li	1.09
Natural NQ	Na	1.11
Toyo SQA	Li	1.44
Toyo SQA	Na	1.31
Toyo SQA	H	1.43

Conclusions

Li^+ and Na^+ behave quite differently in their effects on (i) dielectric loss, (ii) dielectric loss after irradiation and (iii) electrical conductivity. The first of these properties involves the bound Al-M pair; this study has shown that the two alkalis sit on very different sites. The second property may tell us what becomes of the alkali ion during irradiation, information which has thus far not been revealed by other methods. The third property is controlled by free M^+ ions produced by the dissociation of pairs; it involves both the energy of migration and of association. In each category, understanding the difference between Li^+ and Na^+ can help us better to characterize quartz crystals and to understand the origin of frequency instabilities.

Acknowledgements

The authors are grateful to Prof. Joel Martin for carrying out electrodiffusion on the samples studied here, and to Prof. L. Halliburton for helpful suggestions. This work was supported by the U.S. Air Force under contract F 19628-82-K-0013.

References

1. J.C. King, Bell Syst. Tech. J. 38, 573 (1959).
2. B.R. Capone, A. Kahan, R.N. Brown and J.R. Buckmelter, IEEE Trans. on Nucl. Sci. NS-17, 217 (1970).
3. J.C. King and H.H. Sander, Radiation Effects 26, 203 (1975).
4. R. Euler, H.G. Lipson and P.A. Ligor, Proc. 34th Annual Symp. on Freq. Control (ASFC), 72 (1980).
5. J.M. Stevels and J. Volger, Philips Res. Rep. 17, 283 (1962).
6. A.S. Nowick and M.W. Stanley, in Physics of the Solid State, S. Balakrishna, editor, Academic Press, N.Y., 1969, p. 183.
7. D.S. Park and A.S. Nowick, Phys. Stat. Sol. (a) 26, 617 (1974).
8. D.B. Fraser, in Physical Acoustics, Vol. V, W.P. Mason, editor, Academic Press, N.Y., 1968, Chapter 2.
9. H. Jain and A.S. Nowick, J. Appl. Phys. 53, 477 (1982).
10. L.E. Halliburton, N. Koumvakalis, M.E. Markes and J.J. Martin, J. Appl. Phys. 52, 3565 (1981).
11. A. Kats, Philips Res. Repts. 17, 133 (1962).
12. J.J. Martin, L.E. Halliburton, R.B. Bossoli and A.F. Armington, Proc. 36th ASFC, 77 (1982).
13. J.J. Martin, L.E. Halliburton and R.B. Bossoli, Proc. 35th ASFC, 317 (1981).
14. A.S. Nowick and B.S. Berry, Anelastic Relaxation in Crystalline Solids, Academic Press, N.Y., 1972, Chapter 3.
15. M.E. Markes and L.E. Halliburton, J. Appl. Phys. 50, 8172 (1979).
16. J.J. Martin, private communication.
17. W.C. Mackrodt, in Computer Simulation of Solids, C.R.A. Catlow and W.C. Mackrodt, editors, Springer-Verlag, Berlin, 1982, Chapter 12.
18. L.E. Halliburton, private communication.



RADIATION INDUCED TRANSIENT ACOUSTIC LOSS IN QUARTZ CRYSTALS

D. R. Koehler and J. J. Martin⁺
Sandia National Laboratories
Albuquerque, New Mexico 87185

Abstract

When exposed to pulsed ionizing radiation, quartz oscillator crystals can exhibit transient acoustic loss (Q^{-1} , resistance) increases and transient frequency shifts. Although the transient resistance increase is usually observed in unswept quartz, resistance changes and transient frequency shifts have been seen in swept quartz. The transient frequency shifts in swept material are caused by the radiation-induced conversion of the Al-OH center into the Al-hole center and by dynamic thermal effects. For the Al-OH conversion process, the subsequent decay is due to the return of the proton by diffusion along the Z-axis channels.

The transient loss increase which can cause the oscillator to stop must be related to motion of the alkali ions which act as charge compensators for the substitutional aluminum in unswept material. We have directly measured the transient acoustic loss in both synthetic and natural quartz. A set of four 5 MHz 3rd overtone AT-cut blanks were fabricated from an unswept bar of Toyo Supreme Q quartz. One blank was left unswept, one was lithium swept, one was sodium swept, and one was hydrogen swept. The blanks were then mounted in ceramic flat pack holders. The (Brazilian) natural quartz crystals were 5 MHz 5th overtone units. The crystals were mounted in a variable temperature cryostat which allowed the measurements to be made from 110K to 400K. The crystals were irradiated with 70 ns wide x-ray pulses from REBA, a flash x-ray facility at Sandia Laboratories. A typical pulse gave a dose of 6000 Rad (Si). The acoustic loss was sampled approximately every 0.5 second by a log-decrement system. The sampling was started a few seconds before the radiation pulse and continued for approximately 50 seconds. An approximately tenfold increase in acoustic loss was typically observed immediately following the radiation pulse. No transient loss was observed in hydrogen swept crystals.

Following an initial fast transient, the acoustic loss increase decayed as $t^{-1/2}$ over the 20- to 50-second interval for temperatures below 340K. At higher temperatures, the decay went initially as $t^{-1/2}$ then became more rapid going nearly as t^{-1} . The $t^{-1/2}$ dependence is characteristic of one-dimensional diffusion. Plots of the magnitude of the transient loss one second after the x-ray pulse against temperature show a broad background that is nearly the same for the synthetic and natural samples. This broad transient loss is most likely a relaxation loss due to the alkalis drifting along the Z-axis channel. The expected low activation energy for such a motion would be consistent with the broad loss.

Sharper peaks, centered at 360K in the unswept synthetic sample and at 380K in the alkali swept synthetic samples, were observed on top of the broad background. In the natural quartz samples, the transient loss increased rapidly for temperatures above 350K suggesting a peak somewhere above our 400K maximum temperature. These transient peaks may be caused by an unstable defect which momentarily traps the alkali ion.

Introduction

When exposed to ionizing radiation, quartz oscillator crystals often exhibit both transient and steady state frequency and Q shifts.¹⁻⁸ The impurity caused frequency shifts stem from changes in the elastic moduli which are associated with the radiation-induced modification of point defects responsible for acoustic loss peaks at temperatures below the operating temperature of the crystal.⁹ In high purity swept crystals, the frequency changes are caused by dynamic thermal effects in the quartz resonator and supporting structure.^{6,8} The Q or crystal resistance changes are related to radiation-induced acoustic loss peaks near the operating temperature of the crystal.¹

As-grown synthetic quartz contains a variety of point defects.^{10,11} These defects fall into two general categories; (1) substitutional aluminum with its associated interstitial alkali, Al-M⁺, and (2) OH⁻ molecules formed by protons trapped on oxygens near unidentified OH⁻ related defects in the lattice. Ionizing radiation creates mobile electrons and holes in the quartz lattice which interact with these defects. At temperatures above 200K, the alkali is released from the aluminum when a hole reaches an adjacent oxygen and then migrates away along the Z-axis channel until it is trapped at an as yet unidentified site. Protons from the OH⁻ related defects also are released (at any temperature) and become trapped at an oxygen adjacent to the aluminum site. Thus the Al-M⁺ centers are converted into a mixture of Al-OH⁻ and Al-hole centers. The pre-irradiation Al-Na⁺ center is responsible for the large acoustic loss peak near 50K⁹ and the radiation process described above results in the removal of the peak at temperatures above 200K.¹² This process is responsible for much of the steady state frequency shifts. Electrodifusion introduced by King¹³ is used commercially to replace the interstitial alkali ions with protons forming the Al-OH⁻ center. A number of studies have shown that the radiation hardness of oscillators is significantly enhanced by the electrodifusion or sweeping process.^{1,2,3,5,6,7,8,14} Since the Al-M⁺ centers have been converted into Al-OH⁻ centers by sweeping in an air ambient, most of the steady-state frequency shift has been eliminated.

When such swept quartz is exposed to ionizing radiation pulses, transient frequency offsets which decay as $t^{-1/2}$ for several decades of time have been observed. King and Sander^{1,2,3} have observed no resistance changes in swept quartz and have shown that the transient frequency effects are caused by the Al center. A radiation-induced hole releases the proton from the Al-OH⁻ center forming the Al-hole center which has acoustic loss peaks at 25K, 100K and 135K^{2,15} and, thus, changes the elastic moduli. The proton undergoes one-dimensional diffusion along the Z-axis channel until it reaches an Al site where it can recombine. Sosin¹⁶ has pointed out that the $t^{-1/2}$ dependence is characteristic of diffusion in one-dimension.

The transient response behavior of as-grown synthetic and natural quartz is somewhat more complicated since the aluminum is charge compensated by an alkali ion. In addition to transient frequency offsets, crystal resistance increases (Q decreases) are observed. These resistance increases can be large

⁺On leave from Oklahoma State University.

AD P002464

enough to cause oscillation to cease. The radiation releases the H^+ ion from the aluminum site as described above. The increased acoustic loss or resistance observed at the operating temperature of the crystal seems to be caused by the H^+ ion drifting along the Z-axis channel. Hughes¹⁷ has observed transient electrical conduction following a radiation pulse in unswept quartz which lasts for many seconds. He suggests that this conduction is caused by the motion of alkali ions along the Z-axis channels. Nowick and Jain¹⁸ have also recently observed a long-lived radiation-induced conductivity in unswept quartz which could be "frozen-in" by cooling the sample below 0°C. Therefore, a $t^{-1/2}$ dependence of the transient loss might be expected since it is again a one-dimensional problem. King and Sander,¹ however, have reported that a somewhat more complicated decay curve is observed. Hughes' measurements of the transient radiation-induced-conductivity also do not display a simple time dependence.

We report here a study of the radiation induced transient acoustic loss (resistance) in a set of crystals electrodiffused with Li^+ , Na^+ and H^+ ; all crystals were prepared from the same bar of high quality synthetic quartz. Results on natural quartz crystals will also be reported. The study was made as a function of temperature with most measurements made over the 280K to 400K range. A few additional measurements were carried out at cryogenic temperatures.

Experimental Procedure

Most of the measurements were made on a set of four synthetic quartz crystals that are described in detail below and on a pair of Brazilian natural quartz crystals. A set of four 14 mm diameter 5 MHz 3rd overtone AT-cut blanks were fabricated from an unswept bar of Toyo Supreme Q quartz[†]. Studies on this bar at Oklahoma State University show that it contains approximately 10 ppm aluminum and relatively few OH related defects.¹⁹ Using the electrodiffusion process described by Martin, et al.,²⁰ one blank was lithium swept, one sodium swept and one was hydrogen swept. The remaining blank was left unswept. After mounting in ceramic flat pack holders, the crystals were designated as follows: unswept, 5S40; H-swept, 5S38; Li-swept, 5S37; and Na-swept, 5S39. The Brazilian natural quartz crystals were 5.12 MHz 5th overtone units designated 266 and 474. A few additional measurements were made on 5 MHz 5th overtone blanks PQ-ER6 and PQ-ER9 which were hydrogen and lithium swept, respectively. These blanks were fabricated from a Sawyer Premium Q bar designated PQ-E by the Oklahoma State University group.

Pairs of crystals were mounted vertically on the copper cold finger of a variable temperature cryostat. The cryostat allowed the measurements to be made over a 110K to 400K temperature range. The crystals were then irradiated at the desired temperature with the 70 ns-wide x-ray pulses from REBA, a flash x-ray facility at Sandia National Laboratories. The dose for each pulse was measured with a TLD dosimeter taped to the aluminum foil radiation window of the cryostat. A typical pulse gave a dose of 6000 Rad (Si). The actual dose in the crystal resonators will be slightly less than that recorded by the dosimeter because the x-rays passed through the window and parts of the holder.

The acoustic loss, Q^{-1} , data were collected as a function of time for each radiation pulse using a calculator controlled logarithmic decrement system.

[†]Oklahoma State University bar designation SQ-B.

Each crystal was driven at its series resonance frequency for approximately 40 ms and then allowed to freely decay. The crystal signal was then amplified and detected using a superheterodyne receiver. The decay time was measured by gating a timer with the exponentially decaying output from the detector. The calculator switched the system back and forth between the two crystals so that Q^{-1} versus time was measured for both samples on each radiation pulse. Each crystal was sampled approximately every 0.5 second. The sampling was started a few seconds before the radiation pulse and then continued until approximately 50 seconds after the pulse. Since REBA was fired independently of the measurement timing, we have a 0.05 to 0.1 second uncertainty in the t_0 of the radiation pulse for most of the runs. The uncertainty is probably 0.1 to 0.15 second for the run where the Na swept and H₂ swept samples were paired.

Results and Discussion

Figure 1 shows the acoustic loss versus "real" time for the unswept and the hydrogen swept crystals fabricated from Toyo Supreme Q quartz. The radiation pulse took place at the time marked as t_0 on Figure 1. In agreement with King and Sander¹ no transient loss changes were observed in hydrogen swept crystals. The transient acoustic loss for the unswept (5S40), the Li-swept (5S37), the Na-swept (5S39) and the H₂-swept (5S38) crystals was measured as a function of temperature over the 280K to 400K temperature range. A $t^{-1/2}$ dependence is characteristic of second order kinetics for one-dimensional diffusion.¹⁶ King and Sander¹ observed that the transient frequency shift decayed as $t^{-1/2}$ over many orders of magnitude. Since the transient acoustic loss involves the drifting of the alkali ions along a Z-axis channel, a $t^{-1/2}$ dependence would seem reasonable. However, as shown by the 389.8K curve in Figure 2, the situation is more complex as the acoustic loss increase decays there more nearly as t^{-1} . The loss increase per dose versus time is shown for the unswept crystal, 5S40, in Figure 2. The 305.7K curve starts slightly faster then goes as $t^{-1/2}$ while the 389.8K shows a decay faster than $t^{-1/2}$ at the longer times. The time dependence of the transient loss in the synthetic samples can be summarized as follows: at the lower temperatures ($T < 340K$), they follow approximately a $t^{-1/2}$ curve; at the higher temperatures a t^{-1} dependence becomes evident and nearly dominates the decay at the highest temperature.

Figure 3 shows the acoustic loss increase per dose versus time after the radiation pulse for natural quartz crystal 474 at 309.1K and 395.0K. Again, a simple $t^{-1/2}$ is not observed. At 309.1K the initial decay is slower than $t^{-1/2}$ followed by a faster decay. At 395K, the curve follows a $t^{-1/2}$ behavior out to around 10 seconds and then shifts to nearly a t^{-1} dependence. Similar behavior was observed for natural quartz crystal number 266. For the natural quartz samples, the faster (approximately t^{-1}) time dependence seems to come in for temperatures above 340-350K. It does not, however, dominate the entire curve as it shows up only after about 20 seconds of the slower decay.

A few measurements at cryogenic temperatures were made on Li⁺ and H⁺ swept blanks PQ-ER9 and PQ-ER6 respectively mounted in a gap holder. The mechanical vibration that takes place when REBA is fired caused the blanks to bounce on some of the radiation pulses resulting in a loss of data. However, the transient acoustic loss in the Li-swept blank disappeared at temperatures below 200K. Below 200K the alkali ion is not released from the aluminum site.^{10,12} Below 150K, a small long-lived transient loss was observed

in the H-swept blank. This loss was probably caused by the conversion of the Al-OH⁻ center into the Al-hole center and an interstitial H⁰ atom. The Al-hole center is responsible for acoustic loss peaks at 25K, 100K and 135K.^{2,15} The H⁰ atom anneals out within 5 minutes at around 120K.²¹

Figure 4 shows the acoustic loss increase per unit dose taken one second after the pulse for the unswept and Na-swept synthetic resonators as a function of temperature. The results for the Li-swept crystal match those of the Na-swept sample and were omitted from Figure 4 for clarity. The results suggest a peak near 360K for the unswept crystal and near 380-390K for the alkali swept samples superimposed upon a broad transient loss which falls off at temperatures below 300K. The upturn for the peaks near 330-340K seems to match the appearance of the faster time dependence.

Figure 5 shows the acoustic loss increase per unit dose taken one second after the pulse for a natural quartz crystal, and for the unswept synthetic sample, as a function of temperature. Here the results show a sharp upturn near 350K on top of a broad transient loss. Unlike the synthetic crystals which show peaks below 400K the loss data for the natural samples continues to increase for temperatures to the maximum reached. Just as for the synthetic samples, the appearance of the faster decay ($\sim t^{-1}$) correlates with the high temperature increase in the transient loss.

The results shown in Figure 4 for the synthetic samples and in Figure 5 for the natural samples suggest that two mechanisms may be responsible for the radiation-induced transient acoustic loss. An alkali ion, when released from the aluminum site, will diffuse along the Z-axis channel until it is trapped. The migration is thermally activated and should have a fairly low activation energy. Nowick and Jain¹⁸ estimated 0.14 eV for this activation energy and the same activation energy would hold for the anelastic loss associated with the ionic hopping. The low activation energy would result in a very broad anelastic loss peak; this may be responsible for the broad portion of the transient loss data which appears at the lower temperatures. The broad loss at the lower temperatures is similar in both the synthetic and natural samples.

The second mechanism is suggested by the higher temperature peaking of the transient loss spectra. For the synthetic samples, the unswept crystal shows a high temperature peak near 360K while the alkali swept crystals have the peak near 380K or 390K. These peaks may be related to lattice defects which momentarily trap the alkali ions. In the natural samples, a different trap may be responsible for the high temperature loss.

Summary

The radiation-induced transient acoustic loss has been measured over the 280K to 400K temperature range for a set of unswept, Li-swept, Na-swept and H-swept synthetic crystals all fabricated from the same original bar of quartz. Similar measurements were made on a pair of natural quartz crystals. No transient acoustic loss increase was observed in the H-swept sample. At the lower temperatures the radiation-induced loss decayed approximately as $t^{-1/2}$ while at the higher temperatures a faster decay was more evident. The unswept crystal showed a peak in the transient loss around 360K while the alkali swept crystals had a peak near 390K. The transient loss

increased in the natural quartz samples for temperatures up to 400K, suggesting a peak at higher temperatures. Additional IR, ESR and annealing experimentation is suggested to further characterize this defect and such work is now being planned.

Acknowledgements

The authors thank J. C. King and T. J. Young for several helpful discussions. J. J. Martin thanks Sandia National Laboratories for the opportunity to spend his sabbatical leave at Sandia.

References

1. J. C. King and H. H. Sander, *Rad. Eff.* **26**, 203 (1975).
2. J. C. King and H. H. Sander, *IEEE Trans. on Nuc. Sci.* **NS-20**, 117 (1973).
3. J. C. King and H. H. Sander, *IEEE Trans. on Nuc. Sci.* **NS-19**, 23 (1973).
4. R. A. Poll and S. L. Ridgway, *IEEE Trans. on Nuc. Sci.* **NS-13**, 430 (1969).
5. T. M. Planagan and T. F. Wrobel, *IEEE Trans. on Nuc. Sci.* **NS-16**, 130 (1969).
6. D. R. Koehler, T. J. Young and R. A. Adams, *Ultrasonics Symposium Proceedings*, 77CH 1264-1 SU, p.877, (1977).
7. T. J. Young, D. R. Koehler and R. A. Adams, *Proc. 32nd Ann. Freq. Control Symposium*, USAERADCOM, Ft. Monmouth, NJ, p.34, (1978).
8. D. R. Koehler, *Proc. 33rd Ann. Freq. Control Symposium*, USAERADCOM, Ft. Monmouth, NJ, p.118 (1979).
9. D. B. Fraser, In *Physical Acoustics*, W. P. Mason, ed, Vol. V, pp.54-110, Academic Press, NY (1968).
10. L. E. Halliburton, N. Koumoukalis, M. E. Markes and J. J. Martin, *J. Appl. Phys.*, **52**, 3565, (1968).
11. W. A. Sibley, J. J. Martin, M. C. Winteragill and J. D. Brown, *J. Appl. Phys.*, **50**, 5449, (1979).
12. S. P. Doherty, J. J. Martin, H. F. Armington and R. N. Brown, *J. Appl. Phys.*, **51**, 4164, (1980).
13. J. C. King, *Bell Syst. Tech. J.*, **38**, 573, (1979).
14. P. Pellegrini, F. Euler, A. Kahan, T. M. Planagan, and T. F. Wrobel, *IEEE Trans. Nuc. Sci.*, **NS-25**, 1267, (1978).
15. J. J. Martin, L. E. Halliburton and R. B. Bossoli, *Proc. 35th Ann. Freq. Control Symposium*, USAERADCOM, Ft. Monmouth, NJ, p.317, (1981).
16. A. Sosin, *Rad. Eff.*, **26**, 267, (1975).
17. R. C. Hughes, *Rad. Eff.*, **26**, 225, (1975).
18. A. S. Nowick and R. Jain, *Proc. 34th Ann. Freq. Control Symposium*, USAERADCOM, Ft. Monmouth, NJ, p.9, (1980).

19. L. E. Halliburton and J. J. Martin, Oklahoma State University, Private Communications, 1982.
20. J. J. Martin, L. E. Halliburton, R. B. Bossoli and A. F. Armstrong, 36th Ann. Freq. Control Symposium, USAERADCOM, Ft. Monmouth, NJ, in press.
21. M. E. Markes and L. E. Halliburton, J. Appl. Phys., 50, 8172, (1979).

Figure Captions

- Fig. 1 The radiation-induced transient acoustic loss for an unswept synthetic resonator is shown. The radiation pulse took place at t_0 . No transient loss was observed in the hydrogen-swept sample.
- Fig. 2 The transient loss increase per unit dose for the unswept crystal is shown. Here a $t^{-1/2}$ dependence is observed at 305.7K only for the longer times. Again, the high temperature decay goes more nearly as t^{-1} .
- Fig. 3 The transient loss increase per unit dose for a natural quartz crystal is shown versus time after the pulse. Both decay curves go initially as $t^{-1/2}$ and then change to a faster dependence.
- Fig. 4 The increased loss per unit dose one second after the pulse is shown as a function of temperature for the unswept and Na-swept crystals. The results suggest a transient loss peak near 360K for the unswept sample and 380K-390K for the Na-swept sample superimposed upon a very broad transient loss.
- Fig. 5 The increased loss per unit dose one second after the pulse is shown as a function of temperature for a natural quartz crystal and for the unswept synthetic quartz sample. The results for the natural sample may suggest a large transient loss peak somewhere above 400K superimposed upon a broad transient loss.

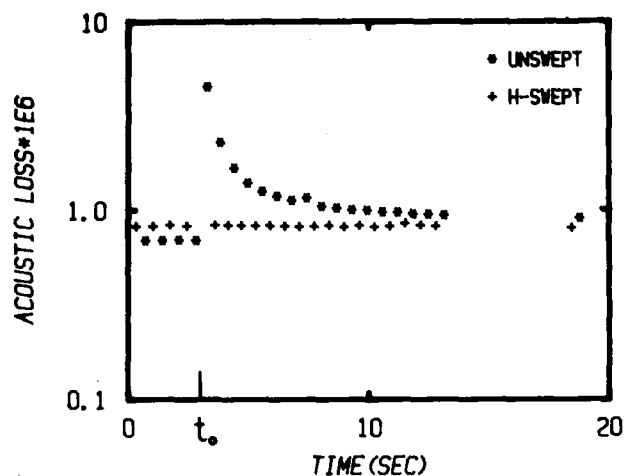


Figure 1. The radiation-induced transient acoustic loss for an unswept synthetic resonator is shown. The radiation pulse took place at t_0 . No transient loss was observed in the hydrogen-swept sample.

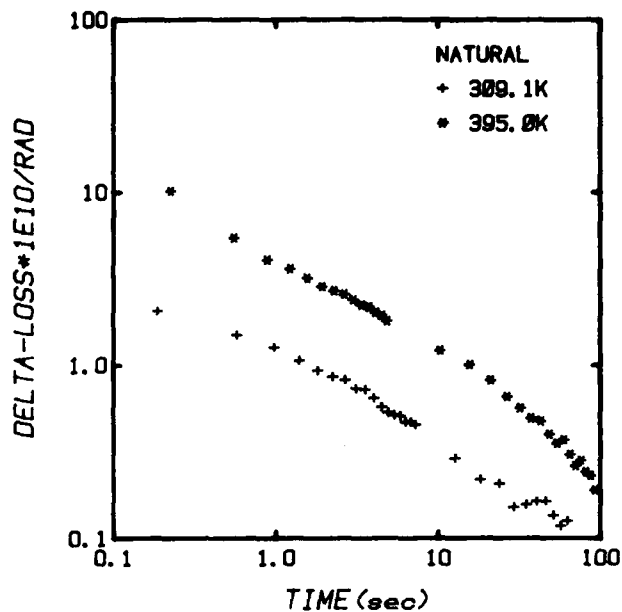


Figure 3. The transient loss increase per unit dose for a natural quartz crystal is shown versus time after the pulse. Both decay curves go initially as $t^{-1/2}$ and then change to a faster dependence.

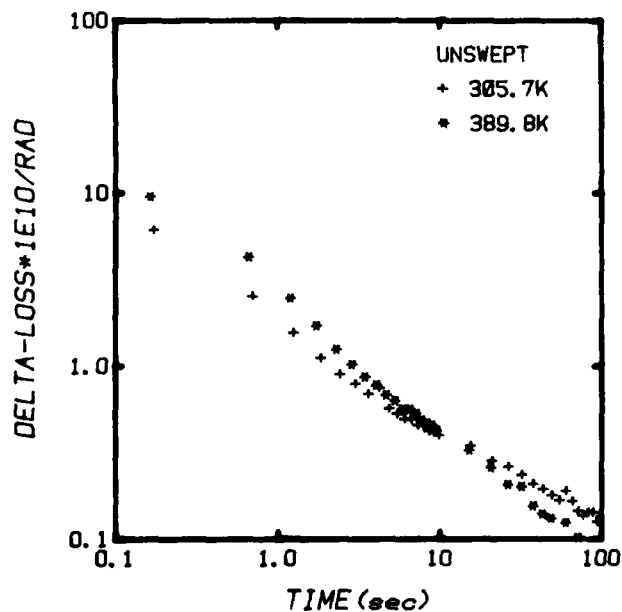


Figure 2. The transient loss increase per unit dose for the unswept crystal is shown. A $t^{-1/2}$ dependence is observed at 305.7K only for the longer times. The high temperature decay goes more nearly as t^{-1} .

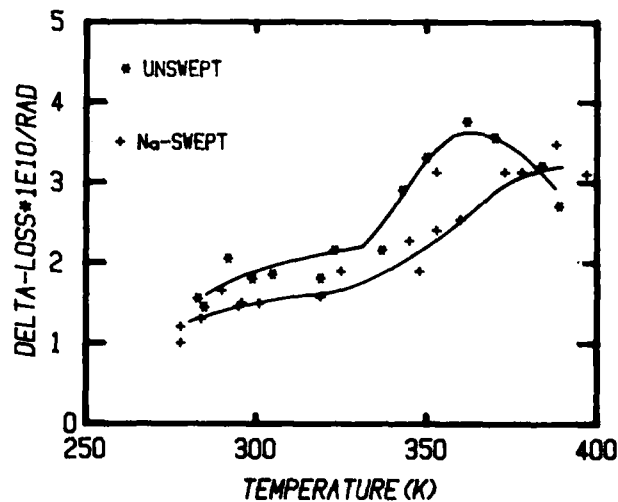


Figure 4. The increased loss per unit dose one second after the pulse is shown as a function of temperature for the unswept and Na-swept crystals. The results suggest a transient loss peak near 360K for the unswept sample and 380K-390K for the Na-swept sample superimposed upon a very broad transient loss.

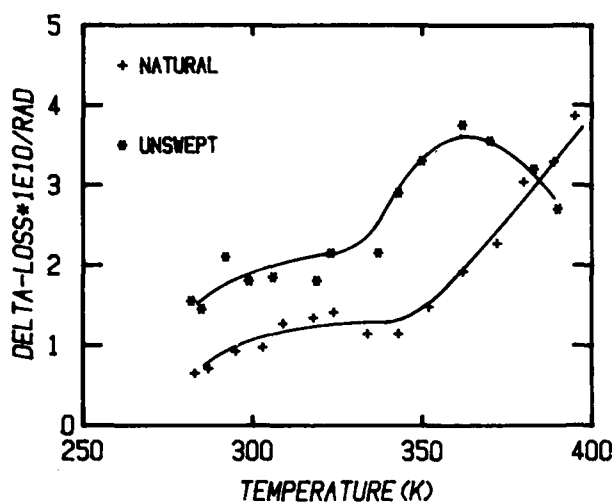


Figure 5. The increased loss per unit dose one second after the pulse is shown as a function of temperature for a natural quartz crystal and the unswept synthetic crystal. The results for the natural sample suggest a large transient loss peak somewhere above 400K superimposed upon a broad transient loss.



THE BULK ACOUSTIC WAVE PROPERTIES OF LITHIUM TETRABORATE

C D J Emin and J F Werner

GEC Research Laboratories
Hirst Research Centre
Wembley UK

ABSTRACT

Lithium tetraborate is a new piezoelectric material which has been shown to have improved properties for SAW applications. In view of this the bulk acoustic wave properties have been predicted for both singly and doubly rotated cuts, using the values of the material constants and their temperature coefficients currently available.

One of the orientations investigated showed potentially interesting properties. An X_2 -cut rotated $56^\circ 40'$ about X_1 displayed a zero room temperature first order temperature coefficient of frequency for the thickness shear mode. Although there is strong coupling to a thickness extensional mode this is well separated out in frequency and the response is effectively single moded.

The piezoelectric coupling constant for the thickness shear mode was predicted to be greater than 6%, comparing favourably with AT quartz (0.8%) and 28° rotated berlinite (3%).

Preliminary measurements on single resonator devices have been conducted. The existence of a parabolic frequency/temperature performance has been confirmed with second order temperature coefficient of $30 \times 10^{-8} \text{ } ^\circ\text{C}^{-2}$. Although the turnover temperature shows a tendency to shift with the value of λ/t (where λ is the electrode length and t the resonator thickness) devices have been designed with a room temperature turnover and a variation of ± 50 ppm between 0 and 40°C . Although this is higher than AT quartz (± 0.5 ppm) the capacitance ratio of these devices has been measured as low as 20 compared with a value for AT quartz in excess of 200.

The maximum attainable bandwidth of a filter using quartz bulk mode resonators is 1%, achieved using complicated synthesis and optimisation techniques. Simple lithium borate bulk wave filters should have fractional bandwidths significantly wider than this and it is expected that by using design techniques similar to those used for quartz filters further improvements could be made.

INTRODUCTION

There is an increasing demand in signal processing systems for compact electronic components using surface acoustic waves (SAW) or bulk acoustic waves (BAW) in piezoelectric substrates. These components include filters, delay lines, encoders, decoders and correlators. The most commonly used substrate material has been α -quartz which offers good mechanical and chemical stability, high intrinsic Q and single rotated cuts with a zero first order temperature coefficient of frequency (TCF) for both SAW and BAW applications.

Unfortunately quartz exhibits weak electromechanical (EM) coupling which is a limiting factor in achieving certain device specifications such as bandwidth and insertion loss. Over certain frequency ranges SAW

transversal filters and BAW resonator filters provide roughly complementary technologies. However there are significant applications in modern radar and communication systems, which require frequency selectivities difficult to achieve with SAW's, combined with bandwidths impracticable for BAW crystal filters. With AT quartz resonators a fractional bandwidth of 1% is feasible but in this region the effects of losses in inductors which are essential to the design are very significant. These problems can be overcome by sophisticated synthesis and optimisation techniques but there is an inevitable practical penalty in terms of sensitivity and ease of construction. There exists, therefore, an urgent need for a material with both high EM coupling and zero first order frequency temperature coefficient.

Materials that have been investigated in recent years include lithium niobate, lithium tantalate and aluminium phosphate (berlinite). Table 1 shows the EM coupling

Material and cut	EM coupling K	k ²	Minimum effective C_0/C_1	Maximum bandwidth %
Lithium niobate $163^\circ X_2$	0.615	0.378	8	12.5
Lithium tantalate $163^\circ X_2$	0.408	0.166	17	6
Lithium tetraborate $56^\circ 40' X_2$	0.281	0.0792	20*	5
Berlinite $28^\circ X_2$	0.15	0.0225	90	1.1
Quartz (AT)	0.089	0.00792	250	0.4

(a) Fundamental mode

Lithium niobate $163^\circ X_2$	0.615	0.042	70	1.4
Lithium tantalate $163^\circ X_2$	0.408	0.0184	170	0.6
Lithium tetraborate $56^\circ 40' X_2$	0.281	0.0088	380*	0.26
Berlinite $28^\circ X_2$	0.15	0.0025	800	0.12
Quartz (Ai)	0.089	0.00088	2000	0.05

(b) Third overtone * Experimentally determined values

Table 1 Comparison of filter performance

AD P 002465

factors and maximum bandwidths of these materials for fundamental and third overtone mode operation. However attempts to build useful filters with these materials have unearthed a number of complications.

Lithium niobate and lithium tantalate exhibit trigonal symmetry with point group 3m, and for both materials there exists, in addition to the major axes, a pseudo-threefold axis rotated 163° from +X₂ about X₁¹ (See Figure 1). A plate prepared at this rotation has a vibrational mode structure similar to AT quartz, but the TCF is of the order of -80 ppm/°C for lithium niobate, and -20 ppm/°C for lithium tantalate. A double rotated (α=75°, β=14°) plate of lithium tantalate has a zero TCF at 25°C, but it has been found that the mode spectrum of resonators operating at fundamental mode for this orientation is distorted by high overtone flexural modes moving through the main response with a TCF of -50 ppm/°C. The TCF for operation at third overtone is -40 ppm/°C.

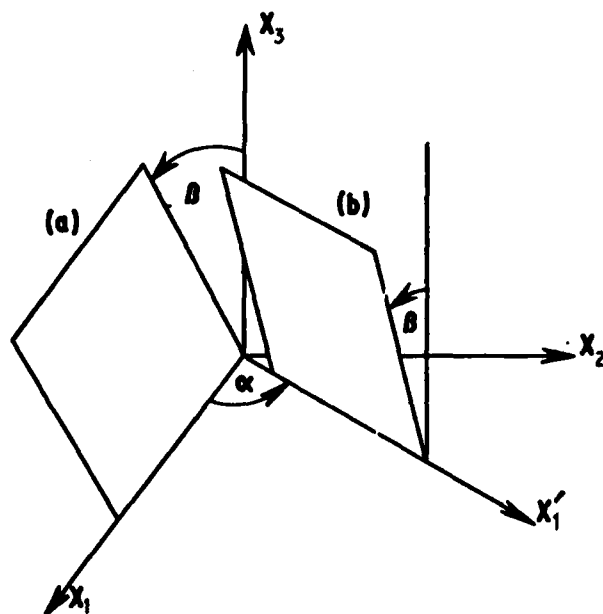


Figure 1 Conventions for specifying plate rotations with respect to crystal axes X₁, X₂, X₃
(a) Single rotation β (b) Double rotation αβ

More recently berlinite has been examined as a possible substrate material^{2,3}. Berlinite is isomorphic with quartz and has a 28° X₂ plate analogous to the AT cut in quartz. Work performed at Hirst Research Centre on berlinite bulk wave resonators has demonstrated a cubic frequency temperature performance with turnovers at room temperature and 100°C. However berlinite offers only a marginal improvement in EM coupling, and berlinite resonators exhibit many undesirable features such as an hysteresis effect on temperature cycling and very low Q values. This was almost certainly due to the fact that the best quality material available had a water content of as much as 1 mg/g in X₂ axis grown material⁴. Growth of single crystal boules by hydrothermal synthesis has proved difficult not least because of its reverse solubility, and investigation of this material has subsequently been discontinued.

Lithium tetraborate (Li₂B₄O₇) is a new material which has been arousing interest for possible SAW applications⁵. It was chosen initially in the hope

that its low density (2451 kg/m³) would result in high acoustic velocities. Single crystal boules have been grown using the Czochralski technique^{6,7} and its symmetry has been determined as tetragonal 4 mm. The elastic, dielectric and piezoelectric coefficients have been measured along with their first order temperature coefficients⁵ and these have been used to calculate BAW properties for single and double rotations of lithium tetraborate plates.

THEORY OF THICKNESS MODE PLATE VIBRATORS

The first exact solution to the problem of vibrations in piezoelectric plates with wave propagation in the thickness direction only was due to Tiersten⁸. He considered the case of an homogenous, anisotropic, infinite plate with infinite massless electrodes subjected to an alternating potential difference. The faces of the plate are assumed to be traction free with X₁ arbitrarily selected as the plate normal. No X₂ or X₃ dependence is assumed. The analysis proceeds in the following stages.

(i) The rotated material constants are obtained for a particular orientation using relations of the form

$$C'_{1jkl} = \sum_{rstu} V_{1r} V_{js} V_{kt} V_{lu} C_{rstu} \quad (1)$$

where V_{ij} is the rotation matrix and C'_{1jkl} corresponds to the rotated elastic constant matrix. In the following text C_{ijkl}, e_{ijk}, ε_{ij} are used to represent the rotated values of the elastic, piezoelectric and dielectric constants.

(ii) The stiffened elastic constants \bar{C} are computed using the relation

$$\bar{C}_{1jkl} = C_{1jkl} + \frac{e_{11j} e_{11k}}{\epsilon_{11}} \quad (2)$$

(iii) Following Tiersten, solutions of the stress equations of motion including piezoelectric stiffening have been assumed to take the form U_j = A_j sin η X₁. This requires

$$(\bar{C}_{1jkl} - C \delta_{jk}) A_k = 0 \quad (3)$$

where δ_{jk} is the Kronecker δ. This matrix equation is solved by Householder reduction⁹ yielding three real positive roots C⁽¹⁾ corresponding to the acoustic velocity of three modes a, b and c. Mode a is therefore a thickness extensional mode whilst b and c are the fast and slow thickness shear modes. Each mode has an associated eigenvector A⁽¹⁾_k corresponding to the polarisation of the particle displacement.

(iv) The EM coupling factor K²₍₁₎ is defined as

$$K^2_{(1)} = \frac{\left(\sum_k A^{(1)}_k e_{11k} \right)^2}{\left(\sum_k A^{(1)2}_k \right) C^{(1)}_{111}} \quad (4)$$

and is plotted in Figure 2 as a function of β for each of the three modes where β corresponds to an orientation from +X₂ about X₁.

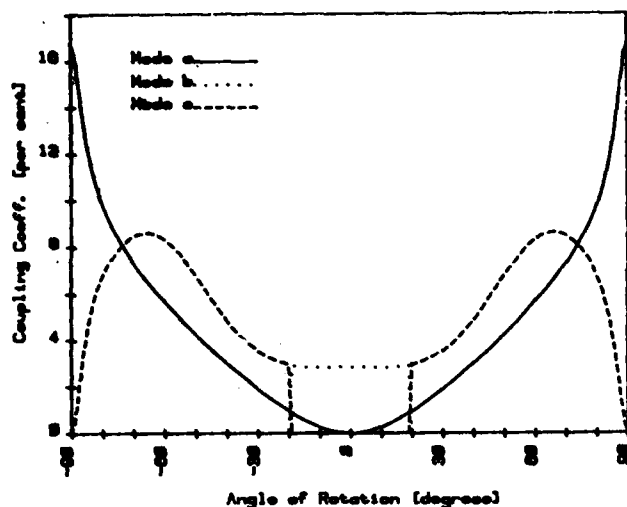


Figure 2 Coupling coefficient as a function of angle of rotation β . Fundamental mode

(v) The natural resonances and antiresonances of the plate are obtained from the poles and zeros of the input admittance¹⁰. These occur at

$$1 - \sum_i k^2(i) \frac{\tan Y(i)}{Y(i)} \rightarrow \infty \text{ for antiresonance (5)}$$

$$1 - \sum_i k^2(i) \frac{\tan Y(i)}{Y(i)} \rightarrow 0 \text{ for resonance (6)}$$

If the plate faces occur at $X_1 = \pm h$ then $Y(i) = h\eta(i)$. Equation (5) is simply satisfied for $Y(i) \rightarrow m\pi/2$ for m odd yielding an infinite set of harmonically related antiresonant frequencies f_a . Equation (6) has three series of coupled roots. For single rotations of $\text{Li}_2\text{B}_4\text{O}_7$ only two modes are excited simultaneously and for simplicity it is assumed that these modes are uncoupled. This is a reasonable assumption for certain values of the ratio of antiresonances for the two excited modes.

(vi) Values of the frequency constants are computed by an iterative solution of the equation

$$Y(i) = \tan \left[\frac{n\pi}{2} \frac{f_r(n)}{f_a(n)} \frac{1}{k^2(i)} \right] \quad (7)$$

(with $n=1$ for the fundamental mode and $n=3$ for the third overtone mode) using the relations

$$Y(i) = \frac{n\pi}{2} \frac{f_r(n)}{f_a(n)} \text{ and } f_a(n) = \frac{n}{4} \frac{V(i)}{h}, \text{ where } V(i) \text{ are}$$

the piezoelectrically stiffened phase velocities $\left(\frac{C(i)}{\rho} \right)$. Frequency constants for the three modes are plotted as a function of β in Figure 3.

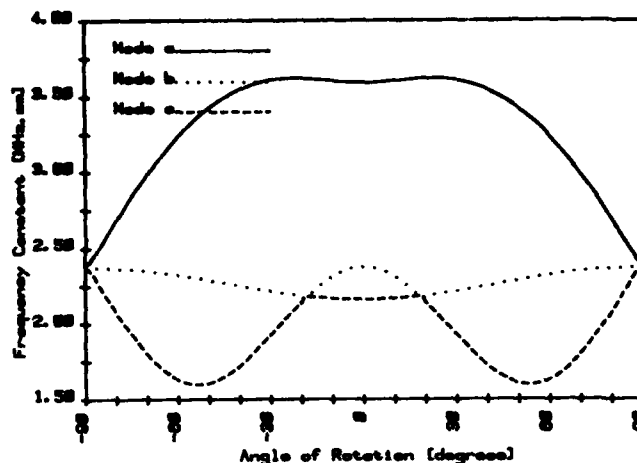


Figure 3 Frequency constant as a function of rotation angle β . Fundamental mode

(vii) Finally the resonant frequencies are adjusted for temperature by using the temperature coefficients of the material constants and their thermal expansion coefficients (see Table 2). The resonant frequencies are computed at room temperature $\pm 0.5^\circ\text{C}$ and a TCF is calculated from the slope obtained from these points.

Coefficient	Magnitude	Temperature coefficient ($10^{-6}^\circ\text{C}^{-1}$)
C_{11}	12.67	-125
C_{12}	0.05	14000
C_{13}	3.0	350
C_{33}	5.39	354
C_{44}	5.50	-23
C_{66}	4.60	-480
e_{15}	0.36	-1300
e_{31}	0.19	-1300
e_{33}	0.89	-1000
ϵ^S_{11}	8.97	-110
ϵ^S_{33}	8.15	-33
α_{11}	13.0	
α_{33}	-1.5	

Table 2 Elastic stiffness (10^{10} Nm^{-2}) piezoelectric (cm^{-2}) and relative dielectric constants and their first order temperature coefficient at 20°C . These were the values used in the calculations. [From Shorrocks et al 1982]. α_{11} and α_{33} are the linear expansion coefficients ($10^{-6} \text{ m}^\circ\text{C}^{-1}$)

These are plotted in Figure 4 for the fundamental mode and Figure 5 for the third overtone mode. It is clear from Figures 2 and 4 that there exists single orientations of $\text{Li}_2\text{B}_4\text{O}_7$ plates for which there are zero TCF's for both the electrically excited modes. These occur at $\beta = \pm 38^\circ 49'$ (thickness extensional) and $\pm 48^\circ 48'$ (thickness shear) for the fundamental mode, $\pm 71^\circ 52'$ and $\pm 68^\circ 17'$ for third overtone.

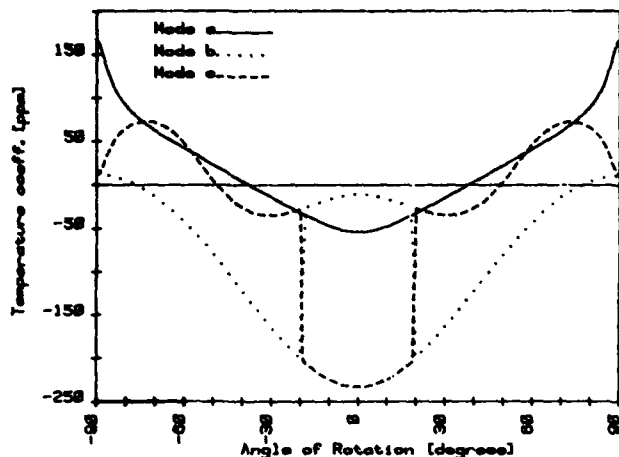


Figure 4 Temperature coefficient of frequency as a function of rotation angle β . Fundamental mode

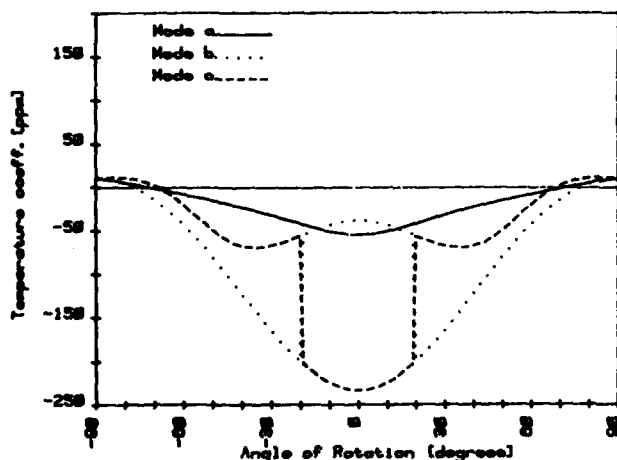


Figure 5 Temperature coefficient of frequency as a function of rotation angle β . Third overtone mode

The experimental data described below was obtained from lithium tetraborate resonators manufactured on plates cut at $\beta = -56^\circ 40'$, an orientation obtained from calculations based on earlier data. It has therefore been necessary to adjust all predictions to this orientation. The $56^\circ 40'$ plate does in fact demonstrate a room temperature turnover for a given value of λ/t where λ is the electrode length and t the resonator thickness. In the case of infinite electrodes where $\lambda/t \rightarrow \infty$ the turnover temperature for this plate tends to an upper limiting value well away from room temperature as anticipated. (see Figure 6).

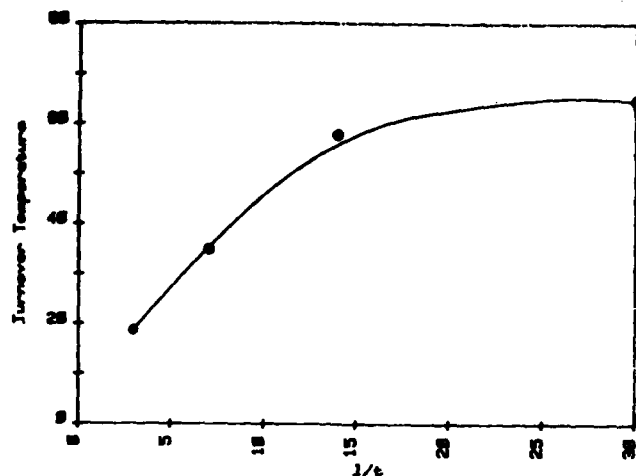


Figure 6 Turnover temperature as a function of electrode size

MATERIAL CHARACTERISATION AND DEVICE PREPARATION

The resonators used in these experiments were prepared with material obtained from an X_1 axis grown crystal boule¹¹. Optical examination revealed an annular core of defective material but away from the central region there were large areas of crystal that were relatively defect free. $56^\circ 40'$ rotated X_2 plates ~ 0.5 mm thick were cut from the boule and then syton polished on both sides. The plates were checked for orientation by X-ray reflection and then the bulk of the material was examined using a Schlieren imaging technique. Circular blanks 8 mm in diameter were trepanned from the best material and then reduced to ~ 0.1 mm thickness using a stretched carrier lapping machine. A final inspection of the blanks by the Haidinger fringe method revealed a thickness uniformity of 1 part in 10^3 .

Blank characterisation proceeded by measuring the blank frequencies at 3rd and 5th overtone in 6 mm diameter air-gap electrodes. Together with a measurement of the thickness using a digital micrometer a value of the frequency constant for the unplated blank was obtained. This was found to be 1700 ± 50 kHz mm.

SINGLE RESONATOR DESIGN

In order to design a single resonator it is necessary to extend the mathematical treatment to allow for the effect of finite electrodes with finite mass. Such a treatment based on a wave propagation analysis for integrated filters has been used previously for the design of AT quartz resonators¹² and an adaptation of this technique has proved successful in predicting the properties of resonators manufactured from higher coupling materials such as 163° rotated X_2 lithium tantalate. As a first approximation this technique has been used for $56^\circ 40'$ rotated X_2 lithium tetraborate. The method involves treating the electroded and unelectroded regions as having the same thickness, t , but separate densities ρ_E and ρ_U ¹³. Solutions are obtained for standing waves in the electroded region and an evanescent wave in the surrounding region; the solutions being matched across the boundary. For square electrodes of length λ these solutions depend on the ratio λ/t , corresponding to the Bachmann conditions for quartz.

The information required is the rotated elastic, piezoelectric and dielectric constants and an estimate of the intrinsic Q of the material. The value used was 10^6 at 1 MHz by measuring the Q of a device at 7th overtone and extrapolating back (Figure 7). Thus values of the fractional mass loading, (defined as $\mu-1$ where $\mu = \mu^2 \rho_u$) and resonator Q have been calculated as functions of plateback for the fundamental and third overtone modes. Here plateback is defined as $(f_u - f_p)/f_u$. f_u is the resonant frequency of the blank and f_p the frequency after plating. These are illustrated in Figures 8 to 11 for a range of values of $1/t$ and were used as the design criteria for energy trapped resonators.

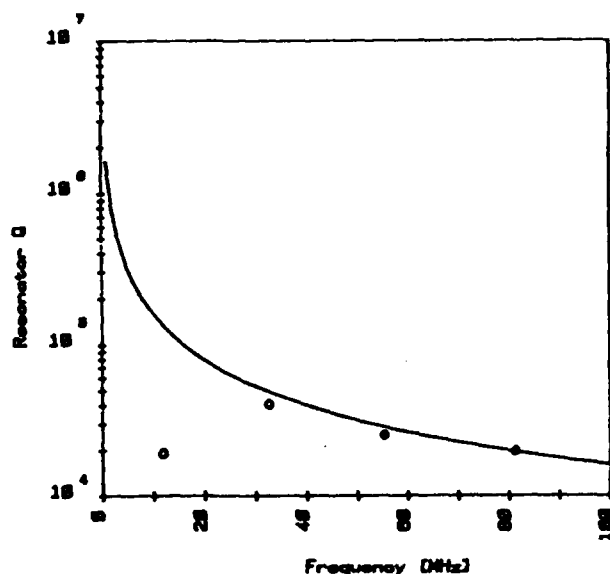


Figure 7 Intrinsic Q as a function of frequency.
 — Extrapolated from resonator Q at third overtone
 o Experimental value for 1, 3 and 5

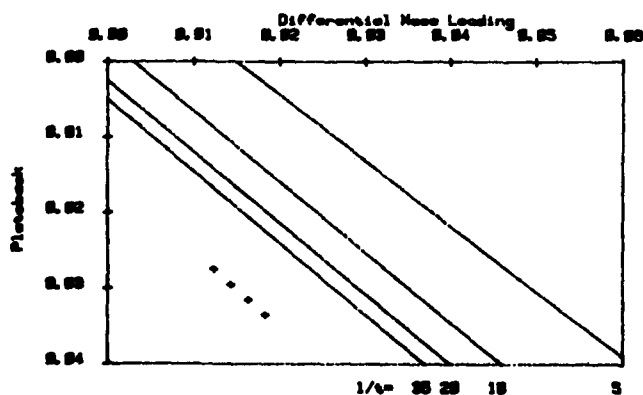


Figure 8 Differential mass loading as a function of plateback for the fundamental mode
 — Theory
 + Experiment ($1/t=14$)

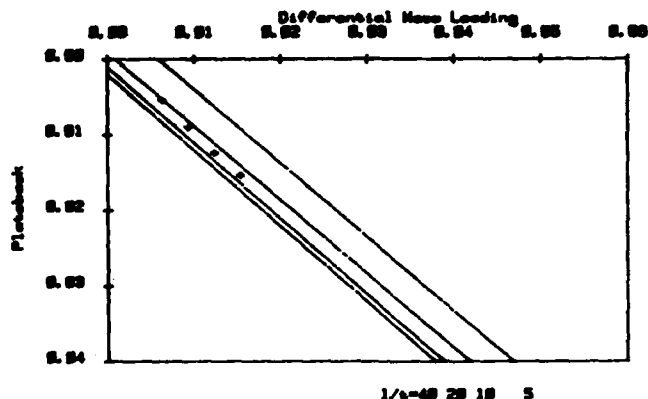


Figure 9 Differential mass loading as a function of plateback for the third overtone mode
 — Theory
 + Experiment ($1/t=14$)

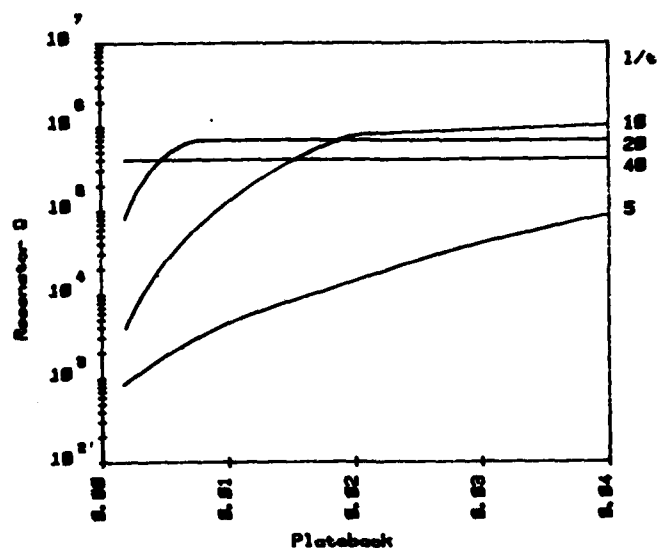


Figure 10 Plot of plateback/resonator Q for the fundamental mode

EXPERIMENTAL RESULTS

Groups of lithium tetraborate resonators have been fabricated with values of $1/t$ from 3 to 30. The crystal equivalent circuit parameters and frequency temperature characteristics of these devices have been examined.

EQUIVALENT CIRCUIT PARAMETERS

The equivalent circuit parameters of the lithium tetraborate resonators have been determined using an automated measurement technique previously described for the measurement of quartz resonators¹⁴. The four S-parameters of the device are measured (transmission measurement), then error corrected using a standard 12

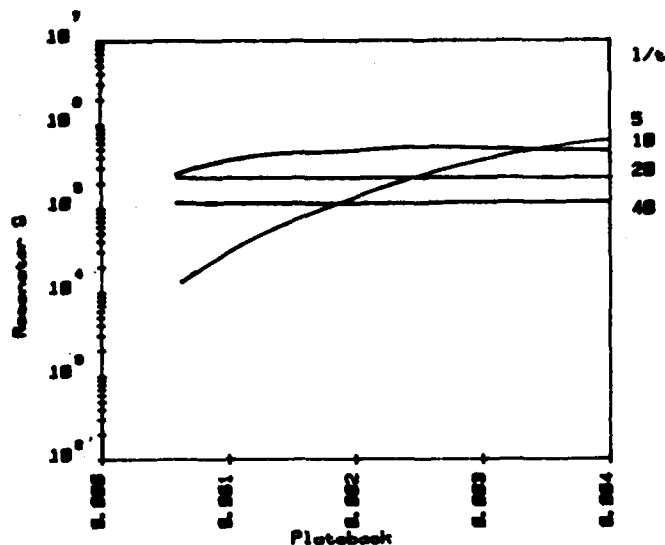


Figure 11 Plot of plateback/resonator Q for the third overtone mode

term error model and finally a non-linear least squares fit is performed using the transmission coefficient, S_{21} . A set of starting values is obtained from an initial survey of the measurements around series resonance which are iteratively refined to obtain optimum values of the crystal parameters R, L, C and C_0 . The measured and calculated responses are then displayed. Figures 12 and 13 show typical outputs for the fundamental and third overtone mode. Accurate, repeatable measurements of crystal parameters can be obtained in this way with traceability to the precision 50 ohm components used in the calibration procedure. A comparison between the values of the measured parameters and those predicted by the single resonator design program is shown in Table 3. The experimental values quoted are averaged over eight resonators. These results are discussed below.

Freq (kHz)	C_0 (pF)	ESR	L_n (nH)	C_n (pF)	Q
11.8588487	1.18	30.0	4.85	3.841E-02	5005

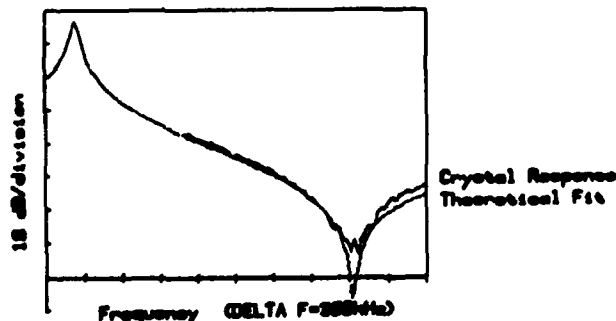


Figure 12 Typical response of a fundamental mode crystal resonator, together with the best value parameter fit

Freq (kHz)	C_0 (pF)	ESR	L_n (nH)	C_n (pF)	Q
23.3821785	1.14	275.50	88.88	2.428E-04	51888

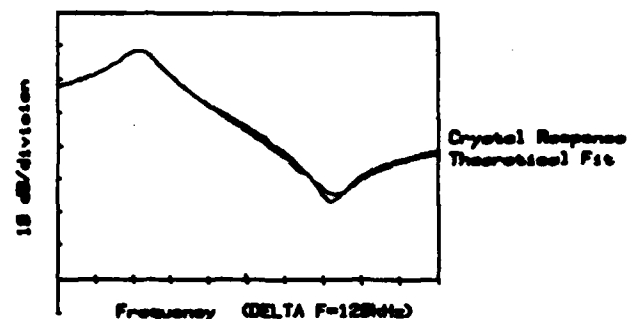


Figure 13 Typical response of a third overtone mode crystal resonator, together with the best value parameter fit

FREQUENCY TEMPERATURE BEHAVIOUR

The variation of resonant frequency with temperature has been investigated in two separate ways.

In order to examine the existence of a room temperature zero TCF, an environmental chamber equipped with both heating elements and CO_2 cooling was used to provide ambient temperature from -20°C to 80°C . Accurate measurements of temperature were made by encapsulating the resonators and then fastening platinum resistance sensors to the container. The frequency was monitored in two ways. Where possible an oscillator circuit was used incorporating a 12 MHz low pass filter to prevent frequency hopping to either the thickness extensional or higher overtone thickness shear responses. The frequency could then be read directly from a counter. Where this was not possible, due to high insertion losses or general lack of stability, frequencies were monitored directly on a spectrum analyser.

The measured data was stored on magnetic tape. Analysis of the data involved making a parabolic fit by the method of least squares. The results are summarised in Figure 14. It is clear from these results that there exists a first order zero TCF but that the turnover temperature is dependent on the value of $1/t$.

The frequency temperature characteristic of a resonator was measured over a wider temperature range, 4 K to 300 K, to look for possible coupling with other modes of vibration which could give rise to lower values of resonator Q. Automated acoustic loss and frequency measurements were made as a function of temperature using a helium/nitrogen flow cryostat system described previously¹⁵. A slightly modified form of the precision crystal parameter measurement technique was utilised. The frequency-temperature characteristic of the resonator is shown in Figure 15 as a percentage change relative to the frequency at 4 K. A large acoustic loss occurs in the 50-70 K temperature regime giving rise to the discontinuity in the frequency-temperature plot. Discontinuities in the range 85-105 K are indicative of coupled modes. A large acoustic loss peak at 160 K gives rise to the inflection in the characteristics in that region. Above about 100 K the main response appears unaffected by spurious modes.

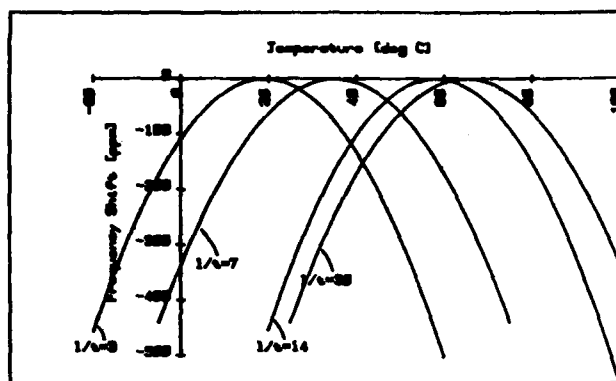


Figure 14 Frequency/temperature characteristics of lithium tetraborate resonators for different electrode sizes. $F=11$ MHz

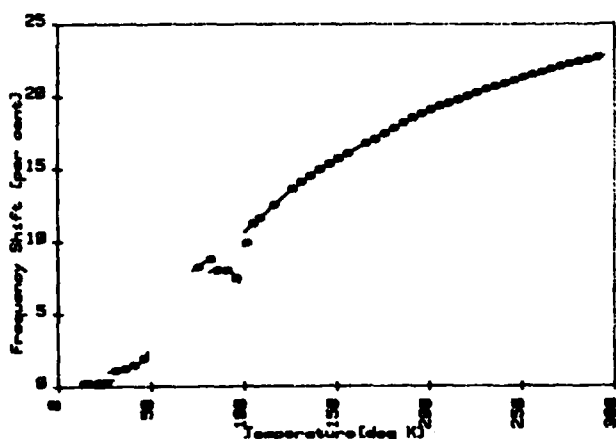


Figure 15 The frequency/temperature characteristics of a lithium tetraborate resonator from 4.2K to room temperature

DISCUSSION

Theoretical calculations have shown a number of orientations of lithium tetraborate offer temperature stability and high electromechanical coupling. A study of the temperature performance of a $56^\circ 40'$ Y rotated plate has revealed some interesting points. For values of $1/t = 30$ the frequency/temperature behaviour is approximately linear with a positive slope of $20 \text{ ppm}/^\circ\text{C}$ as expected. However experiment has shown that as $1/t$ decreases the temperature for zero TCF decreases rapidly reaching 20°C for $1/t = 3$. It is also noted that an increase in turnover temperature with $1/t$ has been achieved by replating the same resonator with different size electrodes. In addition it is independent of electrode orientation with respect to the face axes X_1 and X_3 . Furthermore, using the linear temperature coefficients available a minimum in the frequency/temperature curve is anticipated with a second order coefficient of $+30.10^{-9}^\circ\text{C}^{-2}$ (which is comparable to ST quartz) whereas in practice a maximum has been obtained with a coefficient $-300 \pm 50 \times 10^{-9}^\circ\text{C}^{-2}$. It seems evident that the first order temperature coefficients are only accurate for $25 \pm 5^\circ\text{C}$ and particularly for temperatures in excess of 30° higher order terms are required.

Measurements of the electrical parameters of single resonators have shown a number of interesting properties; the most striking being the large anomalies between predicted and theoretical values for fundamental responses whilst at third overtone far better agreement has been obtained (see Table 3). This would seem to point to cross-coupling between the fundamental mode and some other mode: the only other electrically excited being a thickness extensional mode with a frequency constant more than double that of the fundamental shear mode. Electrical coupling to this mode ($k^2 = 4\%$) cannot explain the increase in coupling required to account for the anomalously high platebacks shown in Figure 8.

The measured Q values for the fundamental mode are an order of magnitude less than anticipated although good agreement has been shown for higher overtone modes.

Measured values														Predicted		
Mode	Frequency	Electrode size	$1/t$	Frequency constant	Plateback	Static capacitance	Rotational capacitance	Capacitance ratio	Rotational inductance	esr	Q	Turnover temperature	2nd TCF	Rotational inductance	esr	Q
	MHz	mm ²		kHz/mm		pF	pF $\cdot 10^{-2}$		mH	ohms		$^\circ\text{C}$	$10^{-9}^\circ\text{C}^{-2}$	mH	ohms	
1	10.56	0.5	3	1720	0.01	1.13	1.86	60	12.2	270	2970	19	-300			
1	11.66	1.0	7	1747	0.017	1.1	3.8	29	4.8	36	9700	35	-290	37.2	34	79200
3	33.3	1.0	7		0.008	1.4			66.6	270	61700			36.0	205	36000
1	11.10	2.0	14	1650	0.030	2.93	14.5	20	1.43	10	9081	58	-314			
1	11.19	4.5	30	1747	0.096	19.0	63.4	30	0.32	10	6600	68	-275			

Table 3 Summary of results obtained on lithium tetraborate resonators

This may be due to mechanical coupling resulting in damping of the fundamental mode or loss of energy through some flexural or thickness twist mode.

It is possible that some approximations in the one dimensional theory and the single resonator theory breakdown in the case of some high coupling materials. Recently a more general theory has solved the problem for coupled acoustic modes in partially contoured crystal resonators¹⁶ and it is hoped to extend this to flat crystal blanks. Even while the fundamental mode behaviour is not fully understood it is still apparent that useful devices can be made on lithium tetraborate at both fundamental and third overtone.

ACKNOWLEDGEMENT

The authors would like to thank Mr D Robertson of RSRE, Malvern for supplying the lithium tetraborate boules.

Part of this work has been carried out with the support of Procurement Executive, Ministry of Defence, sponsored by DCVD.

REFERENCES

- 1 J W Burgess, R J Porter, Procs 27th AFCS, 246, 1973
- 2 Z P Chang, G R Barsch, IEEE Transactions on Sonics and Ultrasonics, SU-23, 2, 127, 1976
- 3 J Detaint et al, Procs of the 34th AFCS, 93, 1980
- 4 M J Bevan, GEC Research Laboratories, Divisional Report, MS16431, 1982
- 5 N M Shorrocks et al, Procs. IEEE Ultrasonics Symposium, 337, 1981
- 6 S R Nagel et al, J. Amer. Ceram. Soc., 60 (3-4), 172, 1977
- 7 J D Garrett et al, J. Cryst. Growth., 41, 225, 1977
- 8 H F Tiersten, J. Acoust. Soc. Amer., 35, 53, 1963
- 9 A Ralston, P Rabinowitz, A first course in numerical analysis, p 511, 2nd Ed. McGraw-Hill, 1978
- 10 T Yamada, N Niizeki, Proc. IEEE, 58, 941, 1970
- 11 Supplied by D Robertson, RSRE, Malvern
- 12 J F Werner, A J Dyer, Procs 30th AFCS, 40, 1976
- 13 W Shockley et al, Procs 17th AFCS, 88, 1963
- 14 R C Peach et al, Procs 36th AFCS, 297, 1982
- 15 S P Doherty et al, Procs 36th AFCS, 66, 1982
- 16 R C Peach, Procs 36th AFCS, 22, 1982

SPUTTERED C-AXIS INCLINED PIEZOELECTRIC FILMS AND SHEAR WAVE RESONATORS

J. S. Wang, K. M. Lakin and A. R. Landin
Ames Laboratory, USDOE
Iowa State University
Ames, Iowa 50011

Abstract

Sputtered piezoelectric films are being studied for potential application to VHF through microwave frequency acoustic wave devices. This paper reports on the growth and characterization of C-axis inclined ZnO and AlN films for shear wave resonators. Orientation relationships for shear wave excitation were calculated for these films which suggest a nearly pure shear wave could be excited with films having C-axis orientation approximately 45° from the surface normal. At other angles, quasi-shear and quasi-longitudinal waves are excited.

The well oriented ZnO or AlN films with C-axis inclined from the film normal were grown in a reactive dc planar magnetron sputtering system having an auxiliary anode structure. These films were evaluated by scanning electron microscopy (SEM) and BAW device measurements. The columnar structure of the C-axis crystallites, inclined from the Si substrate normal, was clearly visible in SEM. Films having C-axis inclination angles up to 45° and thicknesses up to 10 microns have been obtained.

Composite resonators were fabricated by sputtering ZnO or AlN films onto the p^+ Si membrane. Typically, these resonators exhibited Q's of about 5000 in 200 MHz to 500 MHz fundamental resonance range. Of particular interest is the temperature compensation of resonant frequency. Resonators having an absolute temperature coefficient of series resonant frequency of less than 1 ppm/ $^\circ\text{C}$ around room temperature have been made for both ZnO/Si and AlN/Si composite structures. The temperature coefficient of the ZnO and AlN plate resonators was also measured to be -36.2 ppm/ $^\circ\text{C}$ and -25 ppm/ $^\circ\text{C}$ which suggested the temperature coefficient of the p^+ Si membrane to be about $+9$ ppm/ $^\circ\text{C}$ in the shear mode.

Introduction

Piezoelectric films of ZnO and AlN are of continued interest for the excitation of bulk and surface acoustic waves. These films have been widely grown by various techniques including MO-CVD(1,2), plasma-enhanced CVD(3), RF magnetron sputtering(4), DC reactive magnetron sputtering(5) and DC triode sputtering(6). In most of these techniques, the grown films show a strong tendency to grow with their C-axis perpendicular to the substrates and therefore suitable for longitudinal wave excitation when driven by an electric field perpendicular to the film plane. A goal for many years in our group has been to obtain films suitable for shear wave excitation. This task is difficult since shear wave excitation requires that the C-axis of the ZnO or AlN crystallites lie in the substrate plane and be aligned in a specific orientation with respect to the substrate. The C-axis in-plane epitaxial films were found to be grown only on R-plane sapphire using chemical vapor deposition techniques in our early work(7,8). At present, there is no satisfactory technique for growing C-axis in-plane films for general device applications(9,10).

Given the problem associated with growing a C-axis in-plane film, we have shifted our attention to

the oblique oriented films. Orientation relationships for shear wave excitation have been calculated for ZnO and CdS that suggest a nearly pure quasi-shear wave would be excited for films having C-axis orientation approximately 40° from the surface normal and applied direction(11). At other angles, quasi-shear and quasi-longitudinal waves are excited. Based on recently obtained material parameters for ZnO and AlN(12,13), the angular dependence for mode excitation is calculated and plotted in Fig. 1 where the coupling coefficient is given for both longitudinal and shear modes as a function of the angle between the C-axis and the electrical field. As can be seen in these figures, the quasi-shear wave excitation will greatly exceed the quasi-longitudinal wave excitation at an angle 42° for ZnO and at an angle 47° for AlN. Therefore, a good shear wave acoustic device can be made when the films have C-axis orientation approximately 45° from the surface normal.

Although C-axis inclined ZnO films have been deposited by placing the substrate at a 90° angle from the target(14), the film thickness variation is large and therefore this film is not suitable for device applications. Recently, oriented ZnO films having C-axes about 15° from the substrate normal have been sputtered by locating the substrate position far from the target center or tilting the substrate angle with respect to the target(15). This technique was used to make a shear wave transducer. For resonator applications, however, a highly oriented piezoelectric film is needed. This paper describes our well-oriented ZnO and AlN films having C-axis inclination angles up to 45° and thicknesses up to 10 microns. These films were grown in a reactive DC planar magnetron sputtering system. Several high Q shear wave resonators have also been fabricated to evaluate the quality of the films.

C-Axis Inclined ZnO and AlN Films

The ZnO and AlN films were grown in a DC planar magnetron sputtering system having an auxiliary anode structure. The reaction chamber is shown in Fig. 2. A peripheral anode ring is used to collect the electron current and help to prevent arcing to the dielectric film and substrate. The six inch anode ring has a four inch opening and is located below the target and 1 cm above the substrate. The electric field distribution created by the anode ring also has an apparent orienting effect on the growing film. The general sputtering conditions were as follows:

Target: 99.999% pure Zn or Al, 5" diameter
Substrate: Si, SiO₂/Si and Al/Si

Target-substrate spacing: 5 cm
Atmospheric gas: 99.999% pure oxygen or nitrogen
Deposition pressure: 0.5 - 5 millitorr
Substrate temperature: 300 $^\circ\text{C}$ for ZnO, 100 $^\circ\text{C}$ for AlN
Cathode voltage: -300 volts
Cathode current: 400 milliamperes for ZnO,
800 milliamperes for AlN

Anode voltage: 20 - 40 volts dc
Growth rate: 3 $\mu\text{m/hr}$ for ZnO, 1.5 $\mu\text{m/hr}$ for AlN
Cathode and anode voltages were measured relative to the grounded substrate. Typically, substrates were placed 1" from the target center in order to obtain a

more uniform film. To promote plasma stability and to obtain better film quality, the anode voltage was kept in the 20 to 40 volt range. When 30 volts dc anode voltage was applied, a ZnO or AlN film with C-axis inclined approximately 30° from the film normal was obtained. This film was oriented with the crystallographic C-axis preferentially aligned and directed toward the target center. A scanning electron microscope (SEM) photograph of a cross-section of a 6 μm thick AlN film is shown in Fig. 3. Here the columnar structure of the C-axis is clearly visible. In order to achieve films with C-axis 45° from the film normal, the wafer was tilted 15° from base plate normal to obtain an oblique deposition. However, the thickness variation due to oblique deposition degraded the quality of the film and resulted in lower Q shear resonant responses. In general, films having a 30° inclination angle were used for the quasi-shear wave resonators in this work. Their thickness variation was about 0.5% per millimeter as estimated from measurements made in the SEM.

The substrates used here included (100)Si, thermally grown oxide on Si and sputtered Al film on Si. No significant differences of the film structure on these substrates were found. It appears that the substrates used to grow normal oriented ZnO or AlN films can also be used to grow these C-axis inclined films.

Electrical characterization of the material is done in the resonator configuration described in the following sections and detailed in previous publications (16-19).

Shear Wave Resonators

Shear wave composite resonators were fabricated using the sputtered films on Si membranes. The configuration of the composite resonator is shown in Fig. 4. The fundamental parallel resonant frequencies of the resonators are plotted in Fig. 5 as a function of Si thickness for several values of ZnO or AlN thickness. In general, the fundamental resonant frequency is in the range from 200 MHz to 400 MHz for as-fabricated ZnO/Si resonators while AlN/Si resonators have a somewhat higher resonant frequency range from 300 MHz to 500 MHz in this study. For a resonator composed of 2.8 μm of ZnO and 7 μm of Si, a series Q of 3900 was measured. A high parallel Q of 6900 was also obtained for a resonator having 2.0 μm AlN and 6.3 μm Si. Typically the resonator Q was found to be in the 3000 to 7000 range. The phase and absolute value of the impedance around fundamental resonant frequencies were computed and plotted as shown in Fig. 5. The series and parallel resonant frequencies of these resonators are also shown on the Smith Chart in this figure. Some nearby spurs on the high frequency side were also detected.

Edge-only supported ZnO and AlN plates were also fabricated in order to evaluate their shear wave velocities and temperature coefficients. To construct these plate resonators, a ZnO/Si or AlN/Si composite resonator was first fabricated. Then a CF_4 gas plasma etch (20) was used to remove the Si membrane from under the ZnO plate. AlN plates were found to be mechanically more rugged than ZnO plates. ZnO plate bowing was observed if the plate thickness was less than 3 μm . The impedance plots and Smith Charts for a 3.5 μm thick ZnO plate and a 3.8 μm thick AlN plate resonator are shown in Fig. 6. The fundamental shear wave resonant responses were measured at 407.53 MHz and 734.45 MHz for ZnO and AlN respectively. The correspondent shear acoustic wave velocities were calculated to be 2850 m/sec for ZnO and 5585 m/sec for

AlN. The resonator Q was found to be in the range of only 1000 to 2000 for these devices due to a parasitic transducer effect existing at the overlap of the bonding pad onto the substrate p⁺ region.

Temperature Compensated Composite Shear Wave Resonators

The temperature coefficient of the resonators was determined by placing the resonators at the end of a coaxial cable in an environmental chamber and measuring the series resonant frequencies as a function of temperature from -20° to +120°C. Resonators having an absolute temperature coefficient of less than 1 ppm/C° around room temperature have been found for both ZnO/Si and AlN/Si composite structures. The temperature coefficients of the nearly compensated composite resonators are shown in Fig. 7 along with those of AT-cut quartz. The ZnO/Si resonator was composed of 0.85 μm ZnO on 8 μm p⁺ Si membrane while the AlN/Si resonator had 2 μm AlN and 6.3 μm p⁺ Si. The resonant Q's were approximately 4000 for both resonators. Also shown in these figures are the temperature coefficients of quasi-shear wave ZnO and AlN plate resonators. The temperature coefficient for ZnO was shown to be -36.2 ppm/C° and AlN was found to be -25 ppm/C° around room temperature. No parabolic temperature behavior was detected in the measurements.

It is apparent that the ZnO and AlN films effectively compensate the p⁺ Si membrane to give an overall temperature coefficient near zero. Based on linear elastic theory, the p⁺ Si appears to have a positive temperature coefficient. In order to study the temperature compensation mechanism, temperature coefficients were measured near room temperature for different Si to ZnO and Si to AlN thickness ratios. The results are shown in Fig. 8. The temperature behavior exhibited for these devices will agree with linear elastic theory if p⁺ Si is considered as a homogeneously doped membrane and its temperature coefficient is assumed to be +9 ppm/C° around room temperature. The temperature coefficients of these composite structures are plotted in Fig. 8 using the above assumptions. As can be seen, the experimental data are closely fitted to the theoretical calculations.

The experimental results for a positive temperature coefficient in p⁺ silicon corresponds with theories for heavy doping in semiconductors (21,22) and is a fundamental effect. Fig. 10 shows the approximate theoretical prediction for shear wave resonance in boron doped Si as a function of doping concentration. Dashed lines show the approximate boundary regions for the Si membrane used in our study.

Conclusion

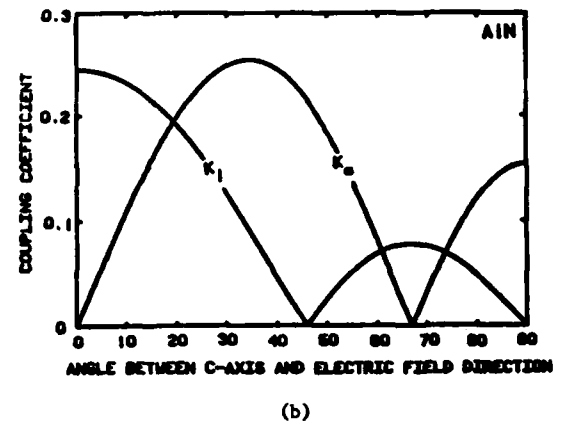
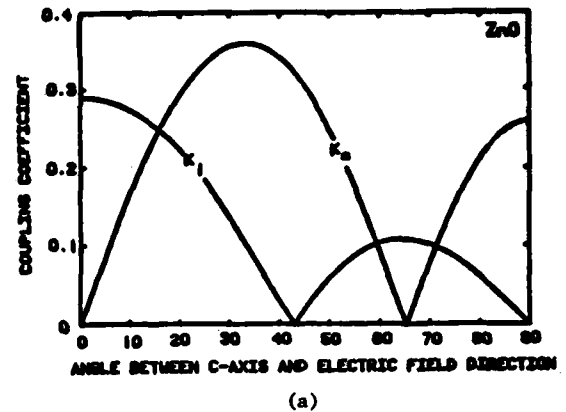
Piezoelectric ZnO and AlN films having C-axes oriented up to 45° from the film normal have been grown on Si substrates in a reactive DC magnetron sputtering system. The quasi-shear acoustic wave velocities were found to be 2850 m/sec for ZnO and 5585 for AlN as measured on films having C-axis 30° inclined from the film normal. The temperature coefficients of these shear wave resonators were also measured to be -36.2 ppm/C° for ZnO and -25 ppm/C° for AlN around room temperature. Temperature compensated composite resonators with the absolute temperature coefficients of the series resonant frequency less than 1 ppm/C° have also been obtained. These devices could be very important for practical resonator applications.

Acknowledgements

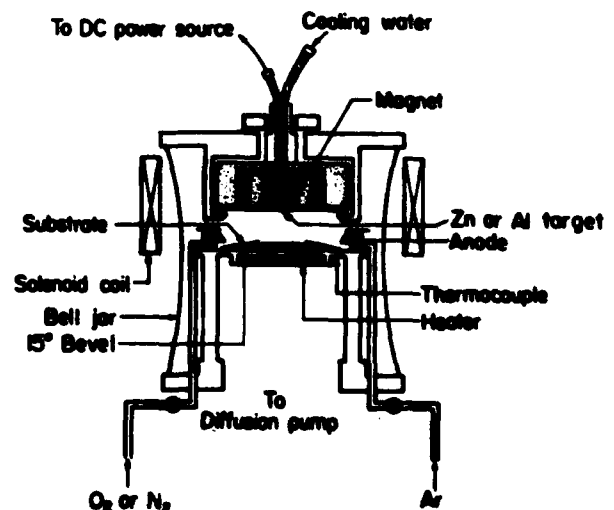
The authors wish to acknowledge G. Kline for sputtering technical assistance and J. Hunt for Si membrane fabrication during the course of this work. This research is supported by the Air Force Office of Scientific Research.

References

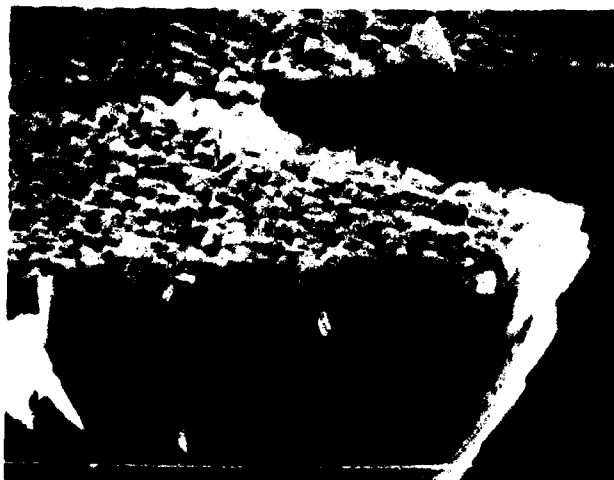
1. S. Ohnishi, Y. Hirokawa, T. Shiosaki and A. Kawabata, Jpn. J. Appl. Phys. 17, 773 (1978).
2. K. Tsubouchi, K. Sugai and Mikoshiba, Proc. IEEE 1980 Ultrasonic Symposium, 446 (1980).
3. T. Shiosaki, M. Shimizu, T. Yamamoto, M. Yagi and A. Kawabata, Proc. IEEE 1981 Ultrasonic Symposium, 498 (1981).
4. T. Shiosaki, Proc. IEEE 1978 Ultrasonic Symposium, 100 (1978).
5. T. Hata, E. Noda, O. Morimoto and T. Hada, Proc. IEEE 1979 Ultrasonic Symposium, 936 (1979).
6. F. S. Hickernell, Proc. IEEE, 64, 631 (1976).
7. S. K. Tiku, C. K. Lau and K. M. Lakin, Appl. Phys. Lett. 36, 318 (1980).
8. J. K. Liu, K. M. Lakin and K. L. Wang, J. Appl. Phys., 46, 3703 (1975).
9. N. F. Foster, J. Vac. Sci. Technol. 6, 111 (1969).
10. H. W. Lehmann and R. Widmer, J. Appl. Phys., 44, 3868 (1973).
11. N. F. Foster, G. A. Coquin, G. A. Rozzanyi and F. A. Vannata, IEEE Trans. on Sonics and Ultrasonics, SU-15, 28 (1968).
12. A. J. Slobodnik, E. D. Conway and R. T. Delmonico, "Microwave Acoustics Handbook", Vol. 1A, 680 (1973).
13. K. Tsubouchi, K. Sugai and N. Mikoshiba, Proc. IEEE 1982 Ultrasonic Symposium, 340 (1982).
14. M. Minakata, N. Chubachi and Y. Kibuchi, Jap. J. Appl. Phys., 12, 474 (1973).
15. S. V. Krishnaswamy, B. R. McAvoy, W. J. Takei and R. A. Moore, Proc. 1982 Ultrasonics Symposium, 476 (1982).
16. K. M. Lakin and J. S. Wang, Proc. IEEE 1980 Ultrasonic Symposium, 834 (1980).
17. K. M. Lakin and J. S. Wang, Appl. Phys. Lett., 38, 125 (1980).
18. J. S. Wang and K. M. Lakin, Proc. IEEE 1981 Ultrasonic Symposium, 502 (1981).
19. J. S. Wang and K. M. Lakin, Appl. Phys. Lett., 40, 308 (1982).
20. J. S. Wang, Y. Y. Chen and K. M. Lakin, Proc. 1982 Ultrasonic Symposium, 346, (1982).
21. R. W. Keyes, Solid State Physics, Vol. 20, pp. 37-90 (1967).
22. P. Csavinsky and N. G. Einspruch, Phys. Rev., Vol. 132, pp. 2434-2440, Dec. 1963.



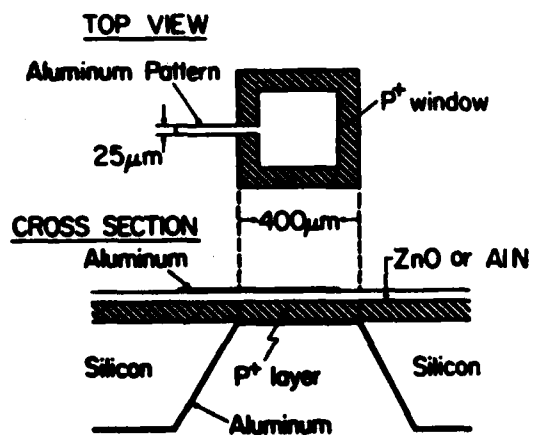
1. Coupling coefficients as a function of angle between the C-axis and the applied electric field (a) for ZnO and (b) for AlN.



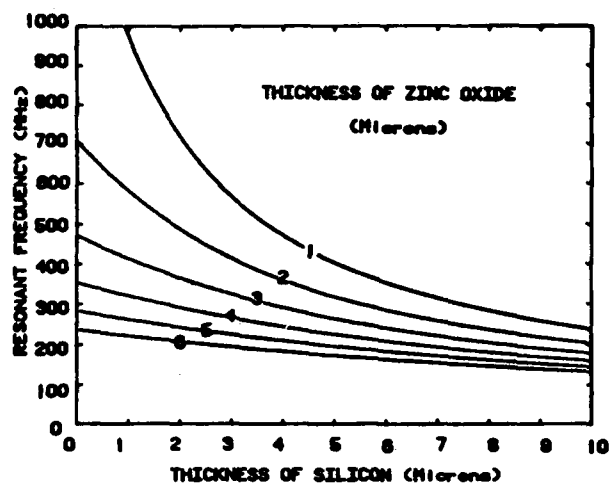
2. Drawing of the dc planar magnetron sputtering system.



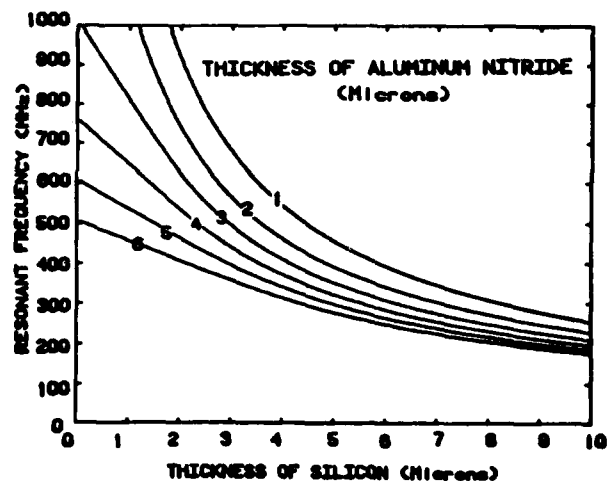
3. SEM photograph of a cross section of a sputtered AlN film with C-axis 30° from the Si substrate normal. The gap in the white line is a one micron reference. (10X magnification)



4. Schematic of a composite resonator. The edge-only sputtered film resonator has the p^+ layer removed from under the ZnO or AlN membrane.

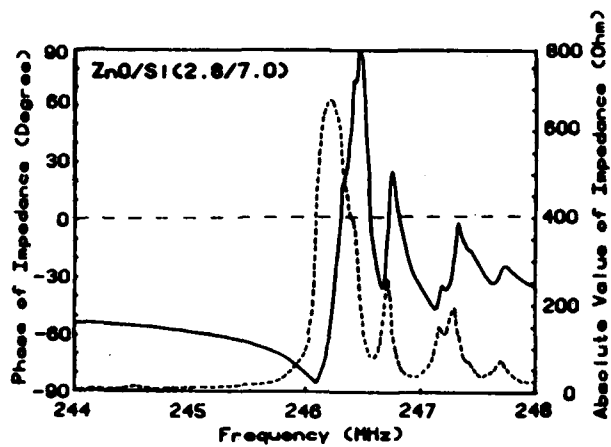
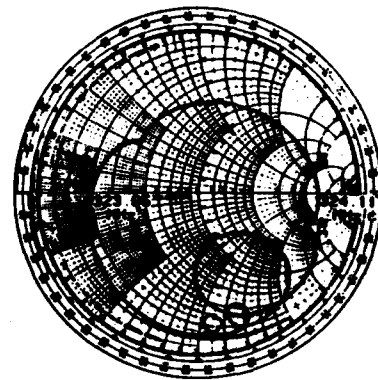
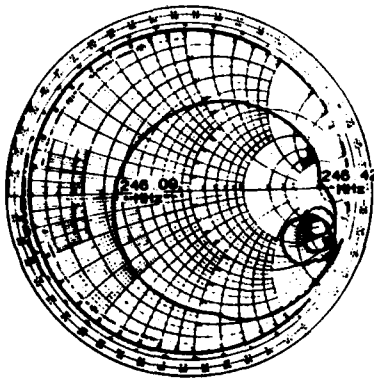


(a)

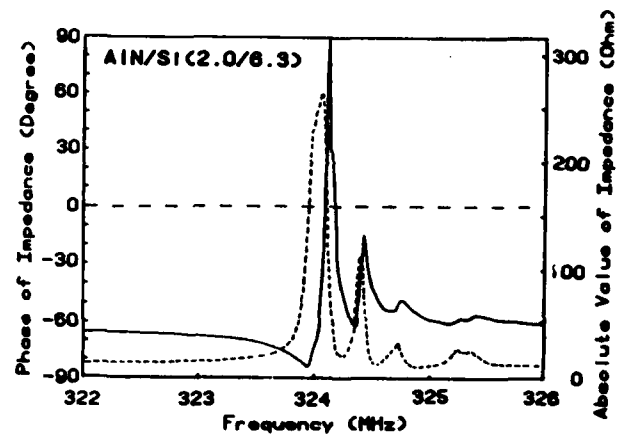


(b)

5. Variation of fundamental resonant frequencies as a function of Si thickness for several values of (a) ZnO and (b) AlN thickness.

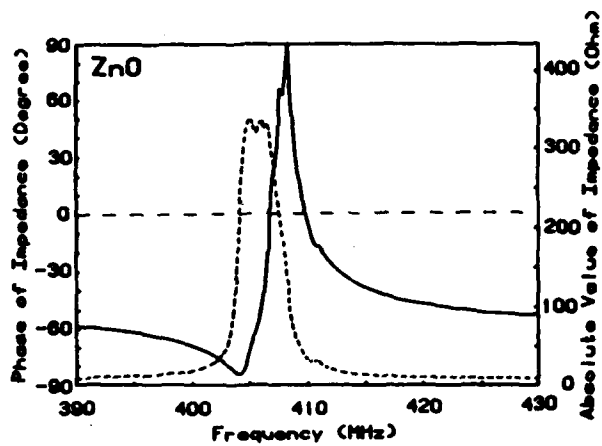
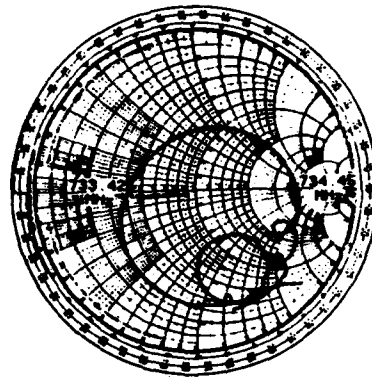
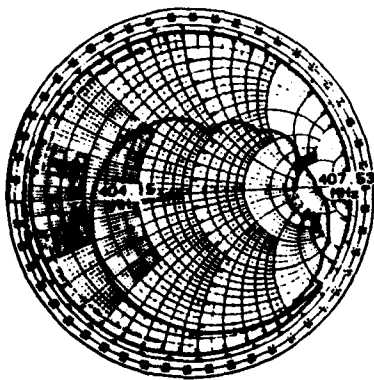


(a)

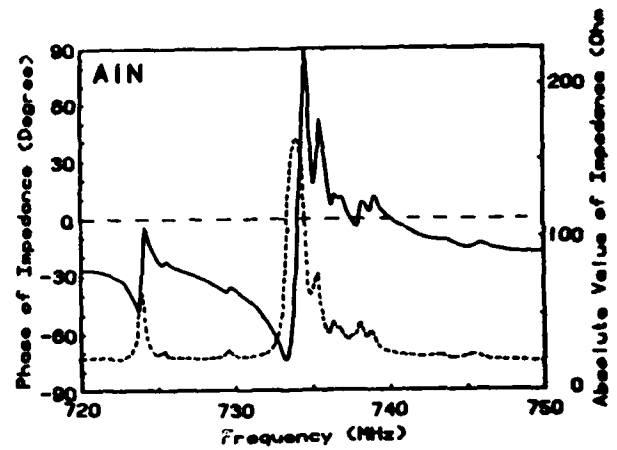


(b)

6. Smith Charts and impedance plots for (a) ZnO/Si and (b) AlN/Si composite resonators around fundamental shear resonant frequencies.

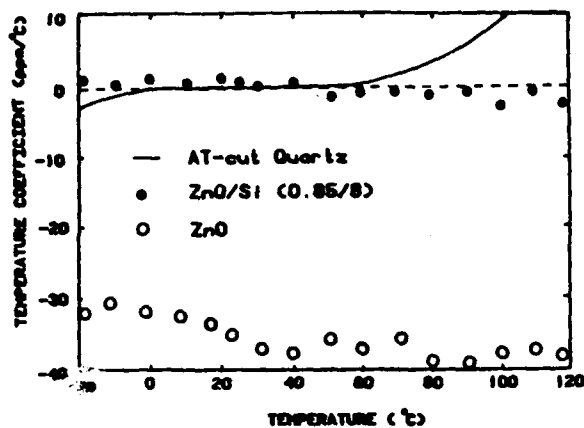


(a)

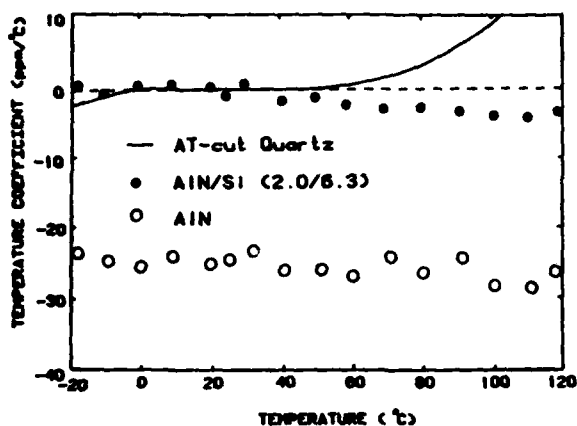


(b)

7. Smith Charts and impedance plots for edge-only supported (a) ZnO and (b) AlN plate resonators around fundamental shear resonant frequencies.

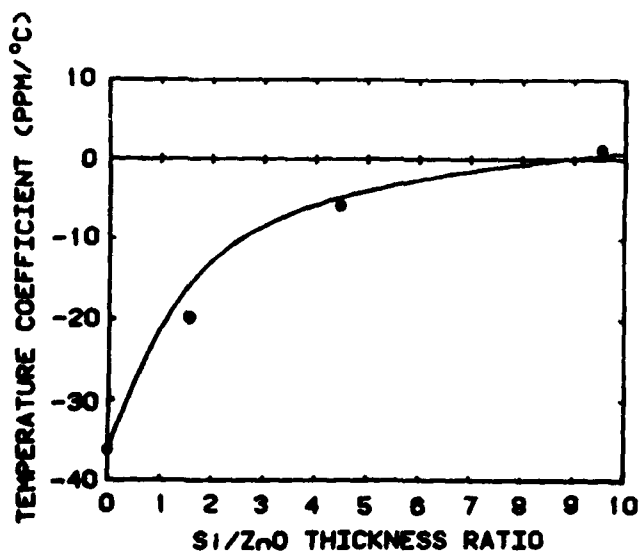


(a)

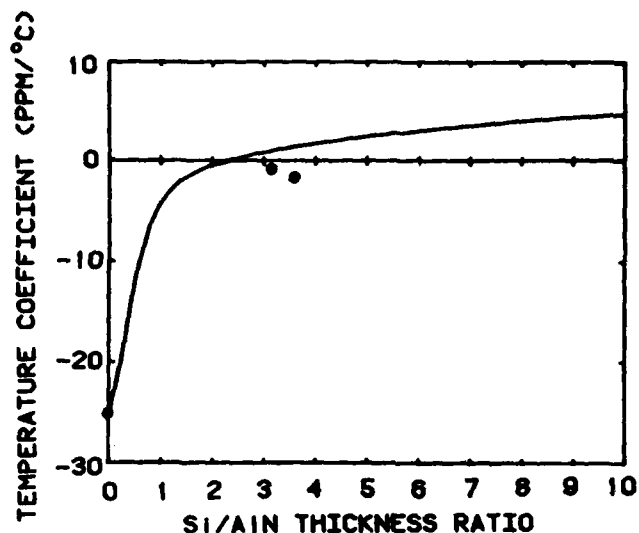


(b)

8. Temperature coefficient of resonant frequency vs temperature for AT quartz conventional resonators, temperature compensated composite resonators and edge-only supported resonators. (a) is for ZnO and (b) for AlN.

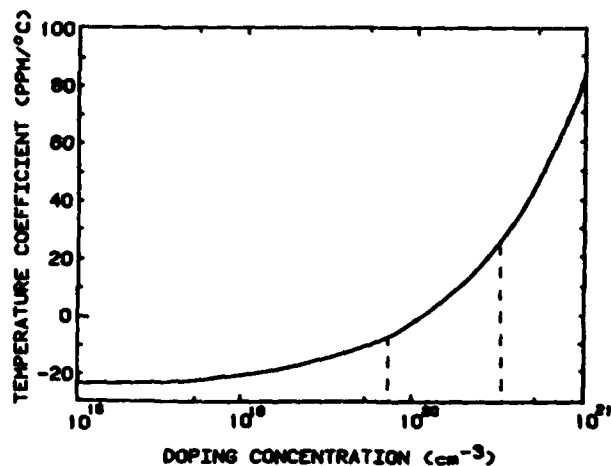


(a)



(b)

9. Temperature coefficient near room temperature vs (a) Si to ZnO and (b) Si to AlN thickness ratios. All data points represent experimental results and the solid lines are theoretical calculations.



10. Temperature coefficient of the resonant frequency of boron doped Si as a function of doping concentration; from band theory.

AD P002467

Recalibration of Q Capability Indications From Infrared Measurements on Cultured Quartz

B. Sawyer

Baldwin Sawyer Crystals, Inc.
Gates Mills, Ohio 44040

SUMMARY

The calibrated formula reported in Sawyer '72¹) that predicts Q capability from measurements of extinction coefficients at 3500 μ m has been re-examined on a broader basis. Quartz samples from four growers have been measured for infrared transmittances at the three wave numbers; 3585, 3500, and 3410, using most of the suggestions in Brice and Cole '78²). In addition, matched pieces of each of the samples were made into Warner's design 5-MHz fifth-overtone resonators by one or more of five manufacturers for Q-measurement use.

Earlier calibrations were generally confirmed for quartz grown in both the carbonate and hydroxide processes. However a small differential in absorption of about 0.013 cm^{-1} between 3500 and the 3800 reference wave number was found to occur in the beam condenser used in the Sawyer '72 work, using a Beckman IR10 prism instrument. The succeeding generation of spectrophotometers utilized more intense IR sources, and gratings, which meant that beam condensers were not needed. For instruments without beam condensers, the Sawyer curve should be revised: moved 0.013 alpha units lower toward the Fraser-Rudd curve (See Figs. 1 & 2 in Sawyer '72¹). This matters little when getting Q indications from large alphas, but results in increasingly overstated Q's as the alphas become small, when the old formula is applied, as it usually is today, to alphas measured without a beam condenser. In fact these measurements indicate that three of the four earlier calibration formulas mentioned by Brice and Cole yield more or less overstated Q's as used today, the exception being the Fraser-Rudd one, which slightly understates.

Of course each of these formulas is as useful as before for ranking samples in Q order, and for testing to calibrated Q tolerance limits. The news here is confirmation that Q's are mostly being overstated at the low-alpha, high-Q ends of the curves. Revised Sawyer and other curves are easily derived if needed, but are not reported here to avoid confusion. Instead a preliminary Q/alpha₃₄₁₀ curve (adjusted slightly from Brice's curve because no polarization correction was used here) is given as follows;

$$10^6/Q = 5.92 \text{ alpha}_{3410}$$

In Sawyer '72 a preference was chosen for a small slit image over a circular one because it permitted high spatial resolution of alpha variations as plotted in a Z-traverse of a Y-section. In the present work a matched slit preference was retained. The slit use and the absence of a rotating sample stage make polarization-averaging as recommended by Brice impractical.

Both formulas in alpha₃₄₁₀, the one in Brice and Cole '78²), and the one above are offered for general purpose use, each repeatable on differing machines. Brice and Cole's will be preferable for averaging over a circular aperture area and polarization averaging: the above one for slit plotting of alpha₃₄₁₀(z) with calibrated results. This formula is still preliminary in that not all Q results have been received, and some modification may be indicated, though little or none is expected.

This recalibration will pose some questions for those who wish to use the recent IEC and EIA standardized Q categories. It seems clear that the category limits in Table I were intended for use on the scales set by the most used formulas in alpha₃₅₀₀ and alpha₃₅₈₅, which tend to give overstated Q's. To be complete when specifying an indicated Q it will be necessary to state which wave number and which Q formula is being used.

For example, a quartz sample which tested at an indicated Q of 3.1 million using the Sawyer '72 formula without a beam condenser should yield an indicated Q of 2.4 million from an alpha₃₄₁₀ measurement and use of the above Sawyer '83 formula. The difference in the indications does not reflect any change in the quartz; only the result of a small differential no longer lost in a beam condenser, whose effect is magnified by taking reciprocals of small numbers.

TABLE I

EFFECTS OF RECALIBRATIONS ON 5MHZ Q CATEGORIES

IEC CODE	$Q \times 10^6$ LIMITS	EIA MINS	ALPHA ₃₅₀₀ BY SAW. '72	RECALIB SAW. '83
	3.8		0.02	2.76
A	3.0	3.0	0.03	2.29
B	2.4		0.04	1.96
	2.2	2.2	0.045	1.83
C	1.8	1.8	0.59	1.53
D	1.0	1.0	0.119	0.91
E	0.5	0.5	0.256	0.48

ACKNOWLEDGEMENTS

The help of many individuals and companies in this effort to accumulate information pertinent to standardization of Q methods is gratefully acknowledged. The writer is indebted for the 5-MHz Q measurements, some still underway, and for many discussions and suggestions. Additionally, support from Sawyer Research Products, and its infrared laboratory are very much appreciated.

REFERENCES

1) B. Sawyer, "Q Capability Indications from Infrared Absorption Measurements for Na₂CO₃-Process Cultured Quartz", IEEE Trans vol.SU-19, No.1, Jan 1972

2) J.C. Brice and A.M. Cole, "The Characterisation of Synthetic Quartz by Using Infrared Absorption", Proc. of the 32nd An. Freq. Cont. Symp. (1978).

PRESSURE-VOLUME-TEMPERATURE BEHAVIOR IN THE SYSTEM H_2O - $NaOH$ - SiO_2 AND ITS RELATIONSHIP TO THE HYDROTHERMAL GROWTH OF QUARTZ*

E. D. Kolb, P. L. Key and R. A. Laudise

Bell Laboratories
Murray Hill, New Jersey 07974

and

E. E. Simpson

Western Electric
North Andover, Massachusetts 01845

SUMMARY

We report here a convenient laboratory method for hydrothermal p-V-T measurements and give pressure data in 1.0M NaOH saturated with quartz as a function of temperature up to 450°C. Measurements were made at percent fills from 65-89% in isothermal laboratory vessels and in autoclaves with temperature differences like those in quartz growth. These results are compared with pressures measured on production sized equipment. The results are used to establish the temperature at which the gas phase disappears under various conditions. Above this temperature $\frac{\partial P}{\partial T}$ is constant and rather large.

One application of these data is to establish safe operating conditions for the autoclaves used for production growth of quartz. Specifically, the steels used for construction of such autoclaves are brittle below a specific temperature which increases slowly with service. The present p-V-T data can be used to assure that high pressures are avoided at temperatures where the autoclave is brittle.

INTRODUCTION

Commercial hydrothermal growth of single crystal α -quartz has been practiced for more than twenty years. It is arguably second only to Si in its importance to electronics. The hydrothermal mineralizer or solvent used is either NaOH or Na_2CO_3 and, particularly in the case of the more generally used NaOH, both the physical chemistry and crystal growth aspects have been quite extensively studied, including phase relations,⁽¹⁾ solubility,⁽²⁾ growth kinetics,⁽³⁾ distribution of impurities (especially OH),⁽⁴⁾ perfection⁽⁵⁾ and electrical properties⁽⁶⁾ of grown quartz. However, the pressure-volume-temperature relations for neither Na_2CO_3 nor NaOH for mineralizer solutions saturated with quartz have been determined. p-V-T data for aqueous solutions of NaOH^(7,8,9) and Na_2CO_3 ⁽⁷⁾ are available, but their relationship to the silica saturated solution used in quartz growth is tenuous. For our purposes the system H_2O - $NaOH$ - SiO_2 is of particular interest since it is most used and probably most studied. We have directed our studies to it and have adopted it for production.⁽¹⁰⁾ In the paper of Kolb et al,⁽¹¹⁾ we reviewed the literature, summarizing p-V-T measurements in hydrothermal mineralizers and described a technique and equipment for rather rapid p-V-T measurements which we applied to the $AlPO_4$ system and which can be easily applied to other saturated hydrothermal systems. As a result of this work and our research and factory experience with quartz, it is our belief that p-V-T measurements in the system H_2O - $NaOH$ - SiO_2 would be particularly useful. A more extended version of this study including experimental details and calibration is available in a recent report⁽¹²⁾ so we summarize here only major results of interest to quartz growers.

*Reprinted (Adapted) from The Bell System Technical Journal. Copyright 1983, AT&T.

RESULTS AND DISCUSSION

p-V-T Behavior of H_2O - $NaOH$ - SiO_2

Figure 1 shows the p-T behavior of 1.0M NaOH saturated with SiO_2 at various fills. 1.0M NaOH was chosen so as to be comparable to conditions used in commercial growth. Figure 1(a) shows the low pressure region while Figure 1(b) covers higher pressures. The line A-B is the coexistence curve. The fills shown on the abscissa are based on the initial cold volumes of solution and autoclave and no correction for temperature or pressure dilatation of the vessel is made to them. In reference (12) these corrections are discussed and applied as are measurements in H_2O -NaOH. Along the curve A-B, liquid and vapor saturated with quartz coexist. As can be seen, at a particular fill once the autoclave has filled with liquid, the pressure departs from the coexistence curve and is linear with temperature; $(\frac{\partial P}{\partial T})_{\%f}$, the slope of the p-T curves at constant %f is constant and greater at higher % fills. Indeed the data of Figure 1(a) may be used to show the temperature at which the system departs from the coexistence curve as shown in Figure 2.

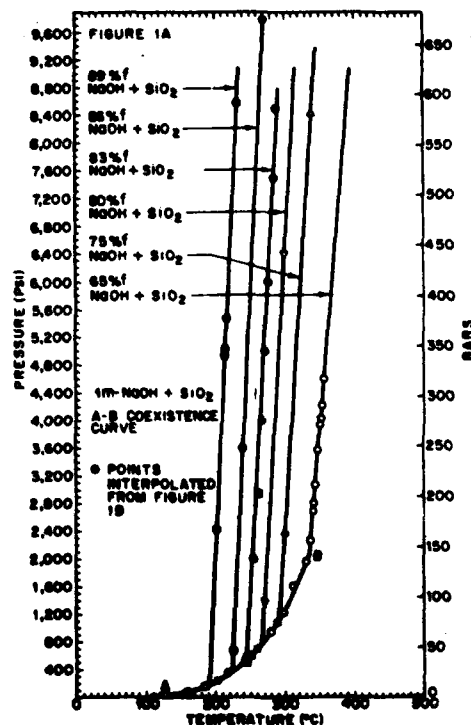


Figure 1 - (a), (b) p-T behavior of 1.0M NaOH saturated with quartz at several % fills

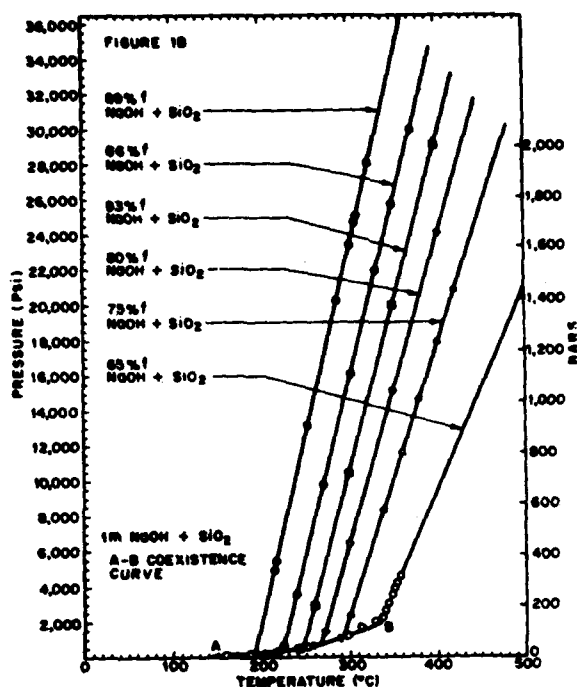


Figure 1 - (a), (b) p-T behavior of 1.0M NaOH saturated with quartz at several % fills

Relationship to Autoclave Embrittlement

Figure 3 shows the brittle-ductile transition of some autoclave steels as a function of exposure time at quartz production conditions (350°C crystallization temperature, 400°C nutrient temperature with the usual ~18 hr warm up and ~24 hr cool down and typical run time of ~30 days). Further identification of the steels is given in reference (13). The data of Figure 3 were obtained by Charpy impact measurements (13)

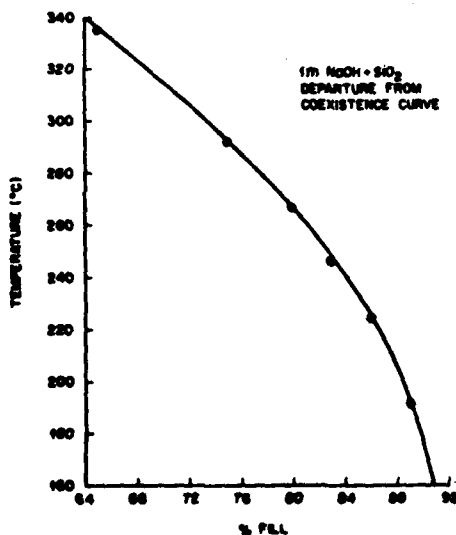


Figure 2 - Temperature at which autoclave fills for 1.0M NaOH saturated with quartz as a function of initial fill.

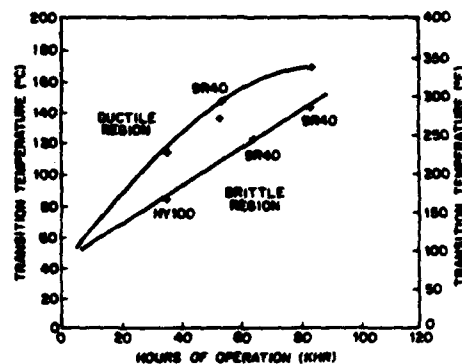


Figure 3 - Brittle-ductile transition of autoclave steels as a function of time at 350-400°C.

The autoclave operating conditions, primarily fill, should be chosen such that high pressure is not developed until the temperature is above the brittle range. On a practical basis, this requires that conditions be chosen using data in Figures 2 and 3 so that pressures do not depart from the coexistence curve until the autoclave temperature is in the ductile region. For example, from Figure 3, an autoclave with 80,000 hours of exposure (about 10 years at a typical 90% duty cycle) will be brittle for temperatures below 140-175°C. From Figure 2, this suggests that fill should be limited to 90%. Thus p-V-T data are very useful in establishing safe operating conditions and should be obtained for other mineralizers and conditions used in commercial hydrothermal processes.

Effect of Temperature Differentials

For practical crystal growth, saturation occurs in a hotter lower region of the autoclave containing relatively finely divided quartz nutrient and growth takes place in an upper, supersaturated, cooler region containing seeds. To obtain information about the effect of temperature differentials on measured pressures, a series of experiments in non-isothermal vessels was conducted. All experiments were done in 1.0M NaOH saturated with quartz at various fills. Figure 4 summarizes these results. We designate temperatures as follows:

T_1 = Upper cooler temperature region of autoclave

T_2 = Mid point temperature of autoclave

T_3 = Lower hotter temperature region of autoclave

$$\Delta T_1 = T_3 - T_1 \quad (T_2 \sim \text{constant})$$

$$\Delta T_2 = T_2 - T_1 \quad (T_3 \sim \text{constant})$$

As can be seen pressure is approximately linear with ΔT_1 and ΔT_2 . Figure 4 can be used to estimate the correction necessary for converting isothermal data to ΔT data. It is interesting to point out that neither hotter lower region, upper cooler region, nor average temperature determine pressure, but a fair fit can be obtained using an average weighted in favor of the upper cooler temperature.

For example, consider Table 1, made from the data of Figure 1 and Figure 4, for 83 %f, 1M NaOH saturated with SiO_2 . Column (a) is the upper cooler temperature region of the autoclave (T_1), Column (b) the measured pressure, Column (c) the temperature difference ΔT_1 , Column (d) the temperature, T_4 , which gives the same pressure in an isothermal autoclave, Column (e)

$\Delta T_3 = T_4 - T_1$ and Column (f) $\Delta T_3/\Delta T_1$. If the average temperature determined the pressure in a non-isothermal autoclave, then $\Delta T_3/\Delta T_1 = 0.5$. As can be seen from Column (f) the average is weighted toward the upper cooler temperature region and the weighting factor is reasonably constant.

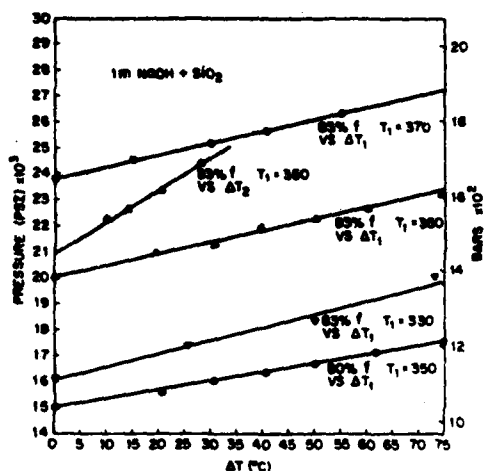


Figure 4 - p-T behavior of 1.0M NaOH saturated with quartz as a function of temperature as defined:

$$\Delta T_1 = T_{\text{Upper cooler region}} - T_{\text{Lower hotter region}}$$

and

$$\Delta T_2 = T_{\text{Mid}} - T_{\text{Upper cooler region}}$$

The pressure is of course uniquely determined by a knowledge of local density and temperature anywhere in the vessel. For example, in a vessel whose average X_f is 83%, if the ΔT is 50°C with a top temperature of 350°C, the observed pressure is 22,180 psi (1530 bar). Using isothermal data for X_f and taking density as $X_f/100$, this suggests a local density of 0.85 g/cc for the region at 350°C and 0.79 g/cc for the region at 400°C. These densities are clearly subject to some errors since they are based on initial X_f fills and do not include contributions due to dissolved solutes. Procedures of this sort might be used to calculate density gradients and hence be used to calculate the driving force for convective circulation in the system.

TABLE 1

Upper Cooler Temperature Region of Autoclave (T_1) (°C)	Pressure P (psi)	ΔT_1 (°C)	T_4 to Give P in Isothermal Autoclaves (°C)	ΔT_3 (°C)	$\frac{\Delta T_3}{\Delta T_1}$
350	23,250	75.0	368	18	.240
350	23,000	67.5	367	17	.252
350	22,000	45.0	361	11	.244
350	21,000	22.5	356	6	.267
350	20,000	0.0	350	0	

Comparison of Laboratory and Factory Measurements

p-T measurements were made on a production crystal growth vessel filled with nutrient, 5% baffle and seeds and using the heater placement and power inputs normal for commercial growth. The volume of the vessel was determined by filling with water to the corrosion mark indicating where the seal ring had rested in the previous run. Volume corrections were made for the leads in the cover. Capillary lead tubing and pressure Bourdon gauge were filled with water using procedures similar to those of the laboratory. The vessel was filled to a nominal fill of 83% with 1M NaOH and warmed from room temperature to operating conditions: 350°C (crystallization temperature), 400°C (nutrient region), ΔT 50°C in a period of 18 hours. During the warm up temperatures and pressures were recorded and are shown in Figure 5.

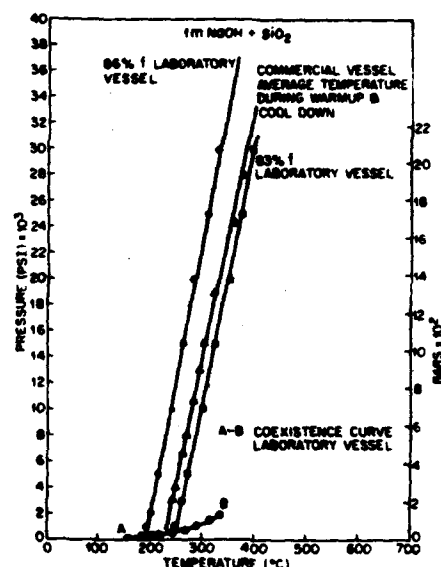


Figure 5 - Pressure-temperature of 1.0M NaOH + quartz in commercial autoclaves during warm up and cool down.

The results are compared with isothermal laboratory data for 83 and 86 X_f . Careful examination of the ring seating region at the conclusion of the run and calculations of the volume from the autoclave dimensions lead us to conclude that the true X_f fill in the vessel was $84 \pm 0.5\%$. On this assumption we can see that the pressures obtained are in reasonable agreement with isothermal laboratory data.

CONCLUSIONS

A reasonably rapid and convenient method for the measurement of the pressures of saturated hydrothermal solutions of the sort used in crystal growth has been developed and applied to the determination of p-V-T relations in 1.0M NaOH and 1.0 NaOH saturated with quartz at degrees of fill from 65-89%, pressures up to ~30,000 psi (2070 bar) and temperatures up to ~450°C.

After about ~10 years of service the brittle-ductile transition of autoclave steel increases to ~170°C so that substantial pressures should be avoided below that temperature. The p-V-T data are used to show the relationship between initial fill and temperature at which the pressure departs from the

coexistence curve and $\left(\frac{\partial p}{\partial T}\right)_f$ becomes large. These data may be used as a guide for the choice of safe operating conditions.

Pressure measurements were made on autoclaves with temperature differentials and the data were related to pressure measurements on isothermal vessels. A weighted average temperature (weighted in favor of the upper cooler region) for the temperature differential reproduced the isothermal measurements. On the assumption that local density and temperature determine pressure, p-V-T data can be used to estimate density differences associated with typical temperature differences (ΔT). A ΔT of 50°C at 350°C growth temperature for 83 Xf 1.0M NaOH saturated with quartz produces a density differential of ~ 0.06 g/cc.

Finally pressure measurements were made on large scale factory autoclaves under conditions of commercial growth and these measurements are compared to laboratory results.

ACKNOWLEDGEMENTS

We would like to thank A. J. Caporaso for suggestions and assistance and especially for help with the design and fabrication of Pt equipment.

REFERENCES

- (1) G. W. Morey and J. M. Hesselgesser, Amer. J. Sci., Bowen Volume, 343 (1952), see also, "Phase Diagrams for Ceramists", Ed. by E. M. Levin, H. F. McMurdie and P. P. Hall, Publ. by American Ceramic Society, Columbus, Ohio, 1956, p. 259.
- (2) R. A. Laudise and A. A. Ballman, "The Solubility of Quartz Under Hydrothermal Conditions," J. Phys. Chem., 65, 1396-1400 (1961).
- (3) R. A. Laudise, "Kinetics of Hydrothermal Quartz Crystallization," J. Am. Chem. Soc., 81, 562 (1959).
- (4) N. Lias, E. Grudenski, E. Kolb and R. A. Laudise, "The Growth of High Acoustic Quartz at High Growth Rates," J. of Crystal Growth, 18, 1-6 (1973).
- (5) 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. of Crystal Growth, 43, 676-686 (1978).
- (6) Proceedings of the 35th Annual Symposium on Frequency Control, 1981, Electronic Industries Assoc., 2061 Eye St., Washington, D.C., see this and earlier Symposia volumes for discussions of quartz properties.
- (7) J. Liberts, "P, T, f-Werte von 4 n K₂CO₃-, Na₂CO₃-und NaOH-Lösungen unter Hydrothermalen Bedingungen," Chem. Eng. Tech., 41, (1969) 1231.
- (8) L. A. Samoilovich, "The Pressure of Hydrothermal Solutions," Trudy VIII, Soveshchaniya, Experimental. Tekhnich. Mineral. i Petrografi, Moscow, (1971) 195.
- (9) S. Kanahara, N. Yamaaki and K. Matsuoka, "Effect of Various Solute Concentration on Hydrothermal P-T Curves. II. NaOH," Rept. Res. Lab. Hydrothermal Chemistry, Kochi, Univ. Japan, 2, (8-15) (1978) 30.
- (10) R. A. Laudise and R. A. Sullivan, "Pilot Plant Production Synthetic Quartz," Chemical Engineering Progress, Vol. 55, No. 5, 55-59 (May 1959).
- (11) E. D. Kolb and R. A. Laudise, "Pressure-Volume-Temperature Behavior in the System H₂O-H₃PO₄-AlPO₄ and Its Relationship to the Hydrothermal Growth of AlPO₄," Journal of Crystal Growth, Vol. 56, No. 1, pp. 83-92 (1982).
- (12) E. D. Kolb, P. L. Key, R. A. Laudise and E. E. Simpson, Bell System Technical Journal, Vol. 62, No. 3, Part 1 (March 1983).
- (13) P. L. Key, E. D. Kolb, R. A. Laudise and E. E. Bresnahan, "Embrittlement of Steels Used in Autoclaves for Hydrothermal Crystallization," J. of Crystal Growth, 21, 164-166 (1974).

STANDARD CHARACTERIZATION METHODS FOR THE DETERMINATION OF THE QUALITY OF HYDROTHERMALLY GROWN QUARTZ

J. F. Balascio and N. C. Lias

Motorola, Inc.
P.O. Box 279
Carlisle, Pennsylvania 17013

Summary

Standard characterization techniques have been developed for the determination of the Q and the etch channel density of hydrothermally grown single crystal alpha quartz. Y-cut slices, for infrared Q evaluation, and AT-cut slices, for etch channel determinations, were respectively lapped to a three micron finish at a final thickness of 6.35 ± 0.03 mm.

In order to properly characterize the Q of the as-grown quartz, the sample should be linearly translated through the beam so that a complete profile along the Z axis is obtained. To eliminate differences due to averaging techniques, the maximum alpha value calculated from the scan should be employed to safely characterize the Q of the quartz. Utilizing standard samples, the repeatability of any alpha value was found to be less than or equal to ± 0.004 .

To determine the etch channel density of as-grown alpha quartz, a standard etching procedure and sample measurement procedure has been formulated. Repeatable channel densities have been obtained utilizing a 24.0 molar NH_4HF_2 solution at $T = 75^\circ\text{C}$ to etch from 4 to 8 samples for two hours. After etching, a $0.25 \text{ cm} \times 0.25 \text{ cm}$ grid is scribed on each AT slice, in the pure Z region, and the channels are counted in each grid area with the aid of a 30x microscope. The average etch channel density (cm^{-2}) is then calculated for the sample.

Introduction

Recent applications for quartz resonators and oscillators require tighter tolerances for performance under hostile environments. These tighter tolerances translate into higher reliability in component characteristics along the chain of manufacture. Uniform and predictable component properties are obtained through proper characterization procedures and traceability throughout the manufacturing process. The first link in the manufacturing chain is the growth and characterization of the alpha quartz. Two properties of interest to consumers of alpha quartz are its Q value and its etch channel density. To aid those involved in quartz manufacture, standard procedures for the characterization of hydrothermally grown alpha quartz with respect to Q, as determined by the infrared technique, and ρ , etch channel density, are presented in this paper.

Within a score of years a number of different techniques and equations have been presented for the characterization of quartz with respect to Q by the infrared technique. (1-6) Some of the techniques involved scanning the sample along the optic axis at fixed wavelengths while others employed a stationary sample with the wavelength changing. Little information was available with respect to sample preparation and the measurement variables present during the determinations. Since the determination of the Q of quartz by the infrared technique can only serve as a reasonable approximation of its true value, a complete and standard

method of sample preparation and evaluation has been developed.

Another property of interest is the etch channel density of alpha quartz. It is important to be able to confidently process alpha quartz into resonators which will exhibit the desired performance requirements. Data gathered over a period of years show that resonators fabricated from low etch channel density quartz have superior mechanical strength than those fabricated from higher channel density quartz. (7) Also, if low etch channel density alpha quartz can be routinely produced, then chemical etching or polishing techniques can replace certain mechanical lapping procedures which would greatly reduce the number of manipulative transfers. This would allow the greater employment of photolithographic techniques in blank manufacture. It would also allow for the production of a greater variety of blank geometries which would be difficult to produce by mechanical means.

To determine the etch channel density of as-grown alpha quartz, a standard etching procedure has been formulated. Included in this procedure are standardizations of the concentration of the etching solution, volume of etching solution, etch temperature and time, number of samples per etching load and, the method of channel density measurement on the etched samples. Typical etch channel density variations within a crystal, as well as, within a growth run have also been determined.

ExperimentalSample Preparation Technique

The preparation of the samples for both infrared and etch channel determinations was identical. Figure 1 summarizes the essential features of the preparation technique employed. The alpha quartz crystal or crystal section is mounted on flat glass with a quick setting adhesive. The minus X surface is in contact with the glass. The crystal is then clamped in position so that a slice, approximately 6.6mm in thickness, is cut from the crystal or section with a diamond blade rotating at approximately 3000 rpm. A Y-cut slice is removed for the infrared measurement and AT-cut slice (35.25°), from the same crystal or section, is removed for the etch channel measurement. Both slices are debonded from the flat glass plate, cleaned and then nested in carriers for lapping. The slices are lapped first with 25 micron abrasive and then finished with three micron abrasive to a final thickness of 6.35 ± 0.03 mm.

Instrumentation

A Beckman ACTA MIV spectrophotometer was employed for the determination of Q by the infrared technique. The wavelength accuracy and resolution of this instrument is better than $\pm 2.5 \text{ nm}$ and $\pm 1.2 \text{ nm}$ respectively. The lapped slice is placed in the instrument compartment so that the focal point of the beam is located at the entering face of the slice. A $1.3 \text{ mm} \times 0.33 \text{ mm}$ rectangular aperture is located between the infrared

source and the slice being scanned. The 1.5mm width is parallel to the Z axis and 6.35mm length is parallel to the X axis. The slice is then linearly translated through the beam so that a complete profile along the Z axis is obtained. The instrument is operated in the absorbance mode with the full scale equal to 0.5A. A scanning speed is used so that one millimeter of sample is displayed on five millimeters of chart paper. Table 1 summarizes the specifications described above.

For the measurement of the etch channel density, a quarter centimeter by quarter centimeter grid is scribed on the AT slice in the pure Z region of the sample and the slice is then viewed through a stereomicroscope with a final magnification of 30x. Under these conditions, one segment of the grid fills the field of view. Each segment is counted and recorded and the average etch channel density (cm^{-2}) is calculated.

Determination of Q by the Infrared Technique

The equation that was employed for the standardization of the infrared measurement of Q was

$$\frac{10^6}{Q} = -0.45(+0.15) \alpha^2 + 7.47(+0.30) \alpha + 0.114(+0.02)$$

The equation is quadratic and relates the absorption coefficient, α , to the Q of a 5 MHz 5th O.T. AT resonator. The numbers in parentheses represent, essentially, the standard deviation about the average value for each coefficient. Figure 2 is a plot of a section of the curve generated from the equation above. In Figure 2, three points were selected to determine the effect of the uncertainty in the values of the coefficients on the resultant Q values calculated. The points selected were 1.5×10^6 , 2.0×10^6 and 3.0×10^6 with absorption coefficients of 0.075, 0.052 and 0.0295 cm^{-1} respectively. For these three points, the variations in the values of the calculated Q due to the uncertainty or the range of the coefficients are, respectively, $\pm 0.10 \times 10^6$, $\pm 0.15 \times 10^6$ and $\pm 0.28 \times 10^6$. In addition to this variation, is the variation in Q dependent upon the change in growth rate as the crystal is grown. That is, depending upon the uniformity of growth in the hydrothermal run, a significant change in Q can occur during the growth of the quartz crystal. Figure 3 illustrates this type of change. This figure is a traced copy of an infrared scan of one of our standard Y-cut slices used to verify the performance of the instrument and scanning arrangement from one test period to another. A test sample scan is always performed before and after the unknown quartz samples are tested. For this scan the instrument was set in the absorbance mode and the test sample was scanned at the two wavenumbers shown in the figure. The little table included in this figure lists the pertinent data logged for each test scan. Listed in the part number section of this table is the thickness of the Y-cut sample scanned in inches (0.58cm).

The large disruption in the center of the scan represents the location of the seed employed to grow the quartz crystal. Both edges of the seed are well defined. This is an indication of the degree of resolution one can obtain. Two points have been defined on this scan: the locations of the maximum and minimum absorption coefficients. These two values are used to determine whether or not an adjustment is to be made to the scanning arrangement. That is, the

variations of the values of the α_{max} and α_{min} are used to determine when the instrument should be adjusted and/or the lamp changed. The Q values calculated from the maximum and minimum alpha values are also shown in this figure. Over a period of years, the reproducibility of the standardization procedure has been documented. Some of these data are shown in Table 2 for the standard sample discussed in Figure 3. Listed in this table are the maximum and minimum alpha values measured from the infrared scans and their calculated Q values. For each column is listed the average value with its standard deviation and the range of the measured or calculated value shown. The repeatability of the α_{max} within one standard deviation is ± 0.0034 and for α_{min} , ± 0.0016 . For shorter time frames (2-3 months), the deviation was found to be ± 0.004 and ± 0.002 respectively.

The data collected over a period of years have shown that there are a number of sources of determinant and indeterminate errors. Limitations on the degree of accuracy of the coefficients of the equation preclude an accurate average Q correlation to the absorption coefficient. In addition to the limitations of the equation, are those variations in Q caused by the changes in growth rate during a hydrothermal run. Dependent upon the thermal program developed by the producer of the alpha quartz for a particular grown product there could be a wide variation in Q in the useable area of the as-grown crystal. Therefore, depending upon the averaging technique employed for the determination of the absorption coefficient different Q values could be obtained from the same infrared scan. Finally, there are also the indeterminate errors associated with the physical implementation of the scanning procedure itself. If we equate the standard deviation of the α_{max} to the acceptable range one would allow the absorption coefficient to span, a range in Q of approximately 0.2×10^6 results. To compensate for these potential sources of error, it has been recommended that the quartz be classified on the basis of the value of the maximum α at 3500 cm^{-1} . (8) This value is obtained from the analysis of a scan of the infrared test slice along the Z axis. The α_{max} is the point of maximum absorption found during the scan of the sample. This procedure had been submitted to the EIA for consideration and was adopted by them last year.

Determination of Etch Channel Density in Alpha Quartz

Interest in low etch channel density alpha quartz has surfaced for a number of reasons, some of which have been stated in the introduction section of this paper. There are many factors that affect the etch channel density found in the as-grown quartz crystal. However, before the effect of the factors on channel density could be ascertained, a standard etching procedure had to be developed in order to quantify, as accurately as possible, the etch channel density variation within a grown crystal, as well as, within a growth run.

The etchant solution chosen was ammonium bifluoride. This solution was chosen because an experimental base was already developed independently. (7) The parameters already fixed for this procedure were the temperature of the etching solution (75°C) and the time of etch (2 hours). The variables in this procedure that still had to be determined were the concentration of the etching solution, the volume of

the solution to be used, the number of samples to be etched per load and the expected variation of etch channel density within a quartz crystal, as well as, within a growth run.

To determine the concentration of ammonium bifluoride which would result in reproducible etch channel densities, multiple "AT" slices were cut and then lapped from three quartz crystals from three different growth runs. These crystals were selected from growth runs in which we anticipated a high etch channel density. Two slices from each single crystal were etched for two hours at 75°C in 0.5 liters of ammonium bifluoride of various concentrations (Figure 4). The etchant solution concentrations were varied from 6.0 molal to 25.0 molal. These data indicated that a relatively constant etch channel density was obtained at 24.0m NH_4HF_2 . It was also noted that a large change in channel density occurred in all three samples between the etchings performed in the 9.0m and 12.0m ammonium bifluoride. Since the solutions were always stirred during the etching process, it seems unlikely that a change from a diffusion controlled process to a reaction controlled process had occurred.

The next variable investigated was the anticipated variation in etch channel density within an as-grown quartz crystal, as well as, within a growth run. In order to remain consistent with the previous parameter investigated, crystals were chosen from a growth run in which was anticipated a high etch channel density. A crystal was selected from both the top and bottom tiers of the growth run. The top test crystal yielded nine "AT" slices which were lapped with one "AT" slice from the bottom test crystal. All ten slices were then etched in 1.0 liter of ammonium bifluoride for two hours at 75°C. The etched slices were then cleaned with acetone and a 0.25cm x 0.25cm grid was scribed on each slice in the pure Z region. Each slice was then viewed through a stereozoom microscope at a final magnification of 30x and each square area was counted. Table 3 contains the data collected on these slices. Seven of the nine determinations from the top test crystal were within one standard deviation of the mean value of these slices. The etch channel density measured on the "AT" slice from the bottom test crystal was easily within the limits of the top test mean. The bottom test slice was removed from one end of the crystal and could be compared with either slice #1 or #9 of the top test crystal. In this case even better agreement with respect to the channel density among these slices is obtained. Figure 5 is a graphical representation of the data in Table 3. Presented in this manner, another interesting feature of these data is more easily seen. There appears to be a pairing of ρ values from the opposite ends of the top test crystal. That is, comparing ρ between slices #1 and #9; slices #2 and #8, etc., a smaller range of etch channel density was found between these slices compared to the range between consecutive slices. The only exception to this was slice #6 which exhibited a very high etch channel density. More data are being gathered with respect to this observation over a range of different channel densities. The initial indication from these data is that a possible twofold symmetry exists with respect to etch channel density in an as-grown crystal. Our main interest here, however, was to develop a degree of confidence that a particular slice from an as-grown alpha quartz crystal would characterize the quality of most of the crystals grown in a particular growth run. The data we have gathered to date have indicated that our original test crystal selection is sufficient for vessel characterization with respect to etch channel

density.

The last variable in the etching process that had to be determined was the number of "AT" slices that could be etched in a given volume of etching solution at one time and obtain reproducible etch channel densities. Again multiple "AT" slices were cut from three different crystals grown in three different growth runs. These slices were then lapped according to the standard lapping procedure previously described. The slices were then etched in 0.5 liters of 24.0 molal ammonium bifluoride for the specified time and temperature. In this case, a different range of ρ values were employed. The number of slices was varied from four to fifteen per load. To reach the prescribed number of samples per load, dummy slices were added to the etching solutions. These dummy slices were equivalent Y-cut slices that had been through an identical lapping stage of the test procedure. Table 4 contains these data. In all of the sample loadings listed in this table, only one slice per stone or crystal was etched in each loading. In this case, the standard deviation listed for the slices measured represents the variation of the channel density within one slice rather than the standard deviation about a mean value for a number of slices from the same crystal as was shown in Table 3. The data in Table 4 indicate that up to eight slices can be etched at one time in the standardized solution without significantly affecting the ρ value for a particular crystal. Less consistent etch channel densities were measured at the fifteen sample loading compared to the results obtained on the respective stones at the other sample loadings. A summary of the total etching procedure developed from these investigations is shown in Table 5. This procedure had also been submitted to EIA and was adopted last year. (9)

Early in the course of this investigation lumbered bars were fabricated from crystals of known etch channel densities. These lumbered bars were sent to our component blank fabrication facility in Franklin Park, Illinois. Round 250 mil "AT" blanks were fabricated from these lumbered bars and these blanks were then etched under standard conditions. These results are shown in Table 6. The specific source of each blank was not maintained with respect to the bar from which it was fabricated and a random sample of these blanks were selected for the etching experiment. The blank data listed in this table were arranged with respect to the determined etch channel density. Certain conclusions still could be made about the values reported and their distribution. The average value of ρ for the five lumbered bars was calculated to be 206cm^{-2} and the mean value for the blanks fabricated from these lumbered bars was $35 \pm 70\text{cm}^{-2}$. This result is in good agreement with some other blank data reported elsewhere by our group. (10) That is, approximately a sixfold reduction in etch channel density seems to occur between the growth of the quartz and the final blank fabrication. For some of the blanks listed in this table there was a degree of uncertainty as to whether or not an etch channel was even present. This uncertainty was because the thinness of the blanks made it difficult to ascertain whether or not an etch channel was present. These blanks were listed as containing three etch channels per square centimeter. The difference between the mean and the median values for these blanks is also approximately sixfold. The probability is that the majority of the blanks that were etched were fabricated from the three low etch channel density bars and therefore, the median value would reflect this distribution. In the unlikely event that all of the blanks counted were

from the low ρ lumbered bars, the reduction in the average etch channel density would still be approximately 50%. In an attempt to discover the origins of this reduction, a large number of etched "AT" slices were analyzed on a $\text{cm}^2 \times \text{cm}^2$ basis to determine if the variation of ρ within an as-grown slice is responsible for the lower channel density seen in the blanks. Figure 6 contains a sketch of a typical low etch channel density pure Z slices analyzed. The dashed parallelogram in the center of the slice represents the original seed used to grow the quartz crystal. The square areas on both sides of the seed represent cm^2 sections in which the etch channel density is counted. A circular section on each side of the seed represents the approximate location where a round blank would be removed from the quartz crystal. Usually, the etch channel density is measured across the entire pure Z region, including the seed. The two columns shown as adjacent to the seed, actually extended into the center of the slice and both contained approximately one half of the seed area. For the purpose of clarity, these columns were moved outside the defined seed area. Listed for each row and column, on both sides of the seed, are the average etch channel density for that particular group of four cm^2 areas. The average channel density for the entire slice was calculated to be 106cm^{-2} and the ρ for the blank areas was 84cm^{-2} . This represents approximately a 20% reduction, however, does not account for the degree of reduction seen between the blanks and the bars previously discussed. Further investigations are being pursued in order to isolate the cause of this reduction. Lastly, it should be noted that the etch channel density tends to decrease from the seed-crystal interface to the outer as-grown Z surfaces in the pure Z region of the crystal. The change in channel density is substantial and, again, reinforces the importance of good seed-crystal interface during growth.

Acknowledgements

The authors would like to acknowledge the contributions of the following individuals to this investigation: Ted Lind and Ken Grata of Motorola's Franklin Park, IL facility for discussions on the infrared Q measurement technique and the fabrication of blanks; Ed Baracchini and Tom Walker of Motorola's Carlisle, PA facility for contributions in all aspects of sample preparation and measurement.

References

1. Dodd, D. M. and Fraser, D. B., J. Chem. Phys. Solids, Vol. 26, 673 (1965).
2. Sawyer, B., IEEE Trans. Sonics and Ultrasonics, Vol. SU-19, 41 (1972).
3. Lias, N. C.; Grudenski, E. E.; Kolb, E. D. and Laudise, R. A., J. Crystal Growth, Vol. 18, 1 (1973).
4. Toyo Tech. Paper: 77-03, May (1977).
5. Brice, J. C. and Cole, A. M., Proc. 32nd Annual Freq. Sympos., 1 (1978).
6. Balascio, J. F. and Lias, N. C., Proc. 34th Annual Freq. Sympos., 65 (1980).
7. Vig, J. R.; LaBus, J. W. and Filler, R. L., Proc. 31st Annual Freq. Sympos., 131 (1977).
8. Motorola, Inc., Materials or Methods Specification, No. 12M80986A74, August (1981).
9. Motorola, Inc., Materials or Methods Specification, No. 12M80986A75, August (1981).
10. Balascio, J. F. and Lias, N. C., "Factors Affecting Etch Channel Density in Hydrothermal α -SiO₂", presented at Fifth International/American Confer.

on Crystal Growth, Coronado, CA, July (1981).

Figure 1. SAMPLE PREPARATION METHOD

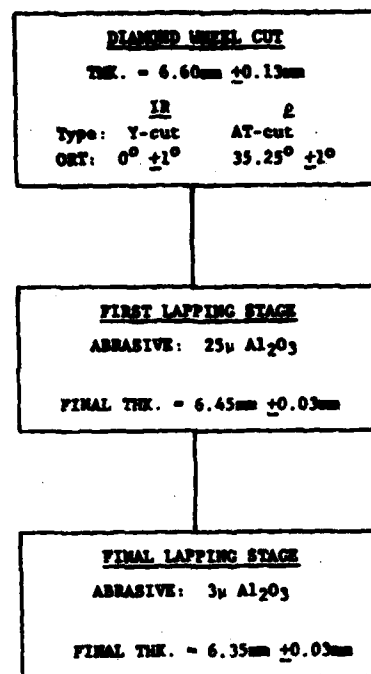


TABLE 1

STANDARD INSTRUMENTATION FACTORS FOR IR Q DETERMINATIONS

INSTRUMENT: BECKMAN ACTA MIV SPECTROPHOTOMETER
RESOLUTION: $\geq \pm 2.5\text{nm}$ ACCURACY: $\geq \pm 1.2\text{nm}$
MODE AND SCALE: ABSORBANCE (0-0.5A)
OPERATING WAVELENGTHS: 2631.6nm (3800 cm^{-1}) &
2657.1nm (3500 cm^{-1})
APERTURE: 1.5mm (112) \times 6.35mm (112)
SCANNING SPEED: 1mm OF SAMPLE/5mm OF CHART

Figure 2. PLOTTED SECTION OF Q CORRELATION EQUATION

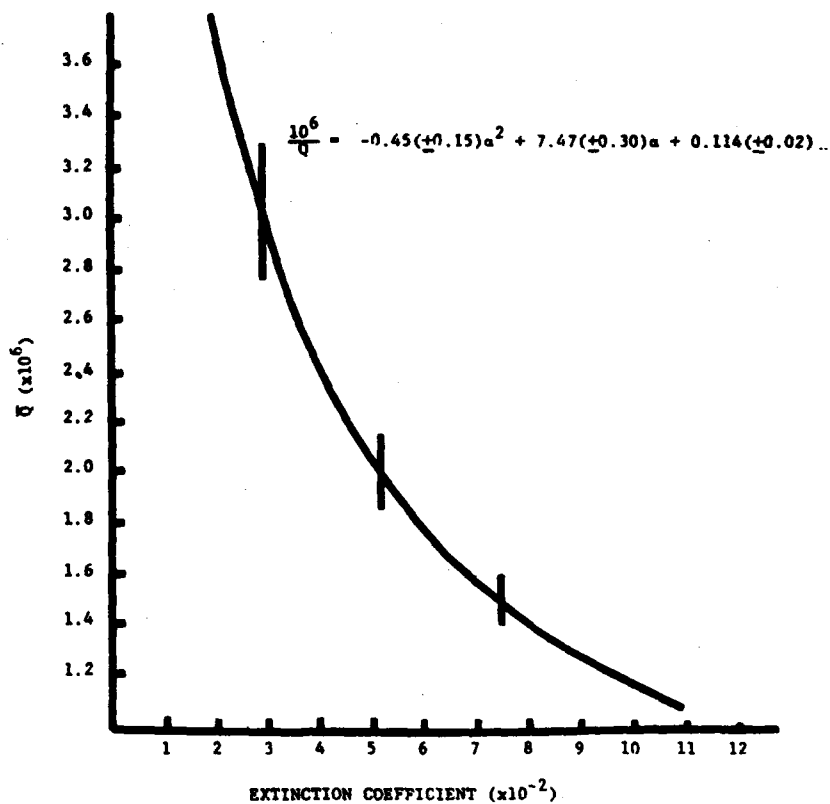
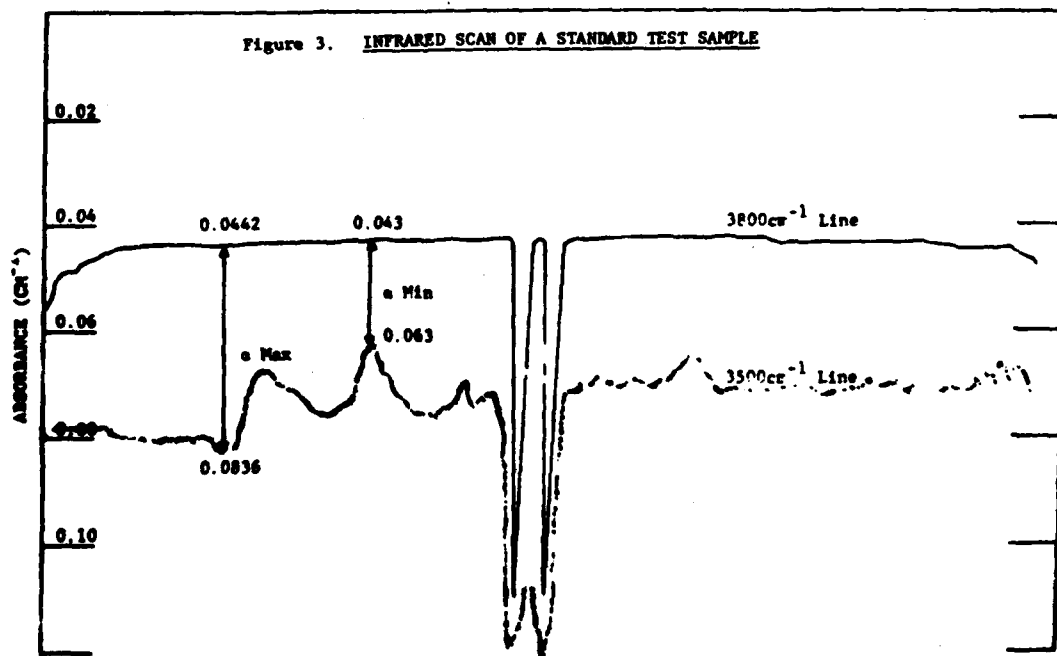


Figure 3. INFRARED SCAN OF A STANDARD TEST SAMPLE



IR Q ANALYSIS

Name _____	Date <u>11-17-80</u>
• P/M <u>0.244</u>	Q Max <u>2.721x10⁶</u>
• P/M <u>0.077</u>	Q Min <u>1.613x10⁶</u>
Run <u>Standard</u>	P/W <u>T=0.229"</u>

TABLE 2

REPRODUCIBILITY OF σ_{\max} AND σ_{\min} .
VALUES OF STANDARD SAMPLE

Date	σ_{\max}	σ_{\min} ($\times 10^6$)	σ_{\min}	σ_{\max} ($\times 10^6$)
11/80	.0677	1.619	.0344	2.700
12/80	.0711	1.555	.0378	2.527
2/81	.0722	1.536	.0368	2.573
3/81	.0654	1.663	.0337	2.738
5/81	.0633	1.709	.0344	2.700
6/81	.0606	1.770	.0372	2.556
8/81	.0661	1.651	.0325	2.807
9/81	.0648	1.677	.0360	2.615
11/81	.0654	1.665	.0341	2.716
12/81	.0698	1.579	.0365	2.590
6/82	.0710	1.557	.0332	2.881
12/82	.0630	1.716	.0345	2.694
3/83	.0663	1.726	.0352	2.657
4/83	.0650	1.673	.0360	2.615
X	.0666	1.650	.0352	2.669
S	.0034	0.071	.0016	0.099
W	.0116	0.234	.0053	0.354

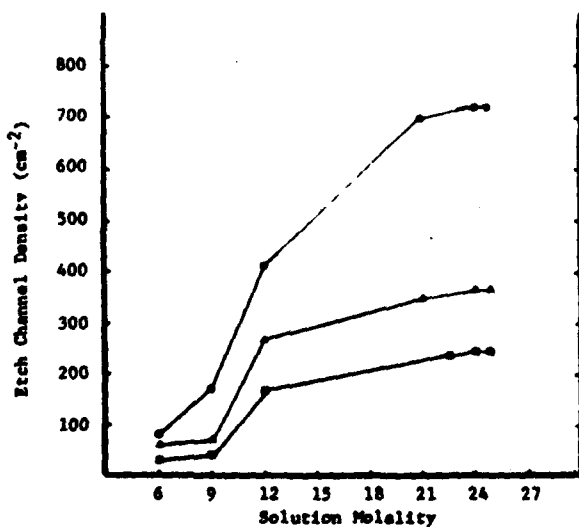


Figure 4. CHANGE IN ρ AS A FUNCTION
OF SOLUTION MOLALITY FOR
THREE DIFFERENT CRYSTALS

TABLE 3

VARIAION OF ρ WITHIN A STONE AND BETWEEN
STONES GROWN IN TOP AND BOTTOM TIERS IN A RUN

Position	ρ (cm^{-2})
Top Test #1	715
Top Test #2	610
Top Test #3	649
Top Test #4	723
Top Test #5	701
Top Test #6	867
Top Test #7	620
Top Test #8	565
Top Test #9	681
Bottom Test #1	704

X (Top Test) = 681cm^{-2}
S = 87cm^{-2}

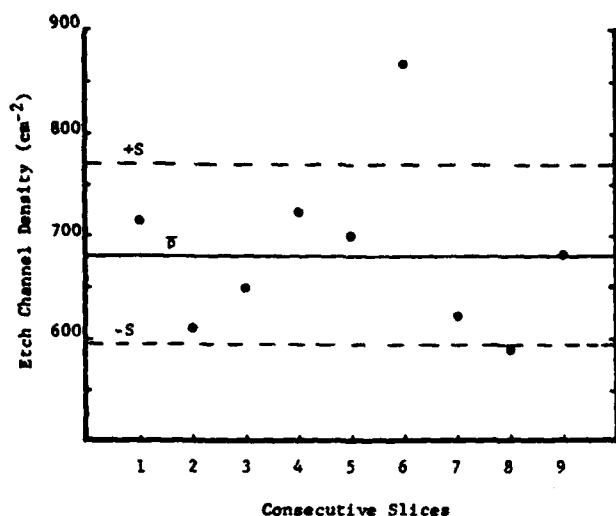


Figure 5. TYPICAL CHANNEL DENSITY VARIATION
WITHIN ONE AS-GROWN CRYSTAL

TABLE 4

COMPARISON OF ETCH CHANNEL DENSITY AS A
FUNCTION OF THE NUMBER OF 'AT' SAMPLES ETCHED

Number	4	8	15
Stone A	270 \pm 76cm ⁻²	280 \pm 24cm ⁻²	281 \pm 39cm ⁻²
Stone B	233 \pm 29cm ⁻²	242 \pm 28cm ⁻²	275 \pm 24cm ⁻²
Stone C	163 \pm 59cm ⁻²	153 \pm 9cm ⁻²	129 \pm 10cm ⁻²

TABLE 5

STANDARDIZATION OF ETCHING TECHNIQUE

SOLUTION: 24.0 Molal NH₄HF₂

ETCH TIME: 2 Hours

ETCH TEMPERATURE: 75°C \pm 5°C

SAMPLES PER LOAD: 4 to 8

SOLUTION VOLUME: 0.5 Liters

TABLE 6

ETCH CHANNEL DENSITIES ON BLANKS
FABRICATED FROM KNOWN CHANNEL DENSITIES

As-Grown Etch Channel Densities					
979-16: 493 \pm 49cm ⁻² (1 bar)					
1095-16: 333 \pm 45cm ⁻² (1 bar)					
1123-16: 68 \pm 20cm ⁻² (3 bars)					
Blank #	ρ (cm ⁻²)	Blank #	ρ (cm ⁻²)	Blank #	ρ (cm ⁻²)
1	322	13	22	25	3
2	249	14	16	26	3
3	155	15	13	27	3
4	85	16	9	28	3
5	66	17	9	29	0
6	57	18	6	30	0
7	54	19	6	31	0
8	44	20	3	32	0
9	32	21	3	33	0
10	28	22	3	34	0
11	25	23	3	35	0
12	25	24	3	36	0
Average: 35 \pm 70cm ⁻²					
Median: 6cm ⁻²					
Range: 322cm ⁻²					

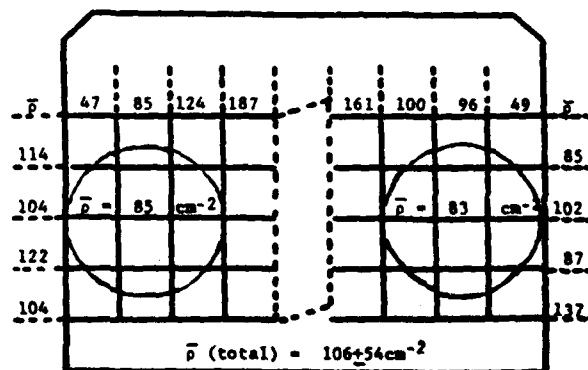


Figure 6. TYPICAL VARIATION OF ETCH
CHANNEL IN A 'AT' SLICE

ELECTRODIFFUSION OF CHARGE-COMPENSATING IONS IN ALPHA-QUARTZ

J. J. Martin, R. B. Bossoli, L. E. Halliburton

Brinda Subramaniam and J. D. West

Department of Physics
Oklahoma State University
Stillwater, Oklahoma 74078Summary

Electrodiffusion (sweeping) is a post-growth treatment which allows the selective exchange of charge compensating interstitial ions in quartz. This technique is employed commercially to enhance the radiation hardness of the material used for quartz oscillator crystals. Nearly all quartz contains Al^{3+} ions substituting for Si^{4+} . Adjacent interstitial alkali ions or protons provide the necessary charge compensation for these Al^{3+} ions. Additional unidentified sites also trap protons to form the OH^- -related defects which are responsible for several IR absorption bands. When thermally released from their trapping sites, the interstitials can migrate along the relatively large Z axis channels. Therefore, if the sample is heated with an electric field applied along the Z axis, the ions can be swept out and replaced either by protons from the surrounding atmosphere or by the desired alkali from a salt electrode. This process also modifies some extended defects, as the etch tunnel density is reduced in hydrogen swept material.

In order to better understand the electrodiffusion process, we are systematically investigating various aspects of sweeping. Current-versus-time curves usually show an initial decay followed by a steady current. We believe that this steady current represents the migration of only one species of ion through the crystal. Once this steady state is reached the temperature is slowly lowered and the ionic conductivity is measured versus temperature to obtain the activation energy. Alkalies show activation energies between .76 and 1.24 eV; while hydrogen shows an activation energy of 1.5 to 1.9 eV. Low-temperature IR absorption spectra, ESR spectra, and acoustic loss measurements were used to characterize the effects on the point defects of alkali and hydrogen electrodiffusion. Four-point bend fracture tests were carried out on a series of AT-cut samples as a function of etching and hydrogen sweeping. Etched samples from hydrogen-swept bars showed a greater fracture strength (i.e., were more difficult to fracture) than etched samples from similar unswept bars.

Introduction

Interstitial alkalis and protons can be trapped by substitutional aluminum ions¹ and other point defects in alpha-quartz. When liberated from such traps,

either thermally or by radiation, the interstitial ions drift along the relatively large Z axis channels. This interstitial diffusion is responsible for the radiation induced transient acoustic loss², the increasing acoustic loss observed at high temperatures,^{3,4,5} and the annealing of radiation-induced defects. An electric field applied parallel to the Z axis can be used at high temperatures to electrodiffuse (i.e., sweep) interstitial ions out of a sample and replace them with other ions.

King⁶ was among the first to develop electrodiffusion as a method for changing the concentration of specific interstitial ions within a crystal. As part of his extensive infrared study, Katz⁷ used the process to sweep specific alkalis and protons into and out of quartz. Fraser⁸ has described the basic process for the selective electrodiffusion of alkalis and Kreft⁹ has shown that holes can be swept into quartz if the process is carried out in vacuum above the phase transition. Alkali sweeping and the use of inert gas atmospheres have been investigated by Brown, O'Conner, and Armington.¹⁰ A number of studies have shown that the radiation hardness of oscillators is significantly enhanced if electrodiffusion has replaced the alkalis in the quartz with protons.^{11,12,13}

Arnold¹⁴ has reported that etching quartz produces very deep tunnels while Nielsen and Foster¹⁵ have observed that these tunnels tend to lie along the general growth direction. The tunnels are apparently caused by dislocation networks which trap impurities.¹⁶ Vig et al.,¹⁷ in a recent investigation of chemically polished and mechanically polished resonator blanks, found that swept material showed fewer of these etch tunnels and that chemically polished blanks were superior in strength.

Electrodiffusion is closely related to ionic conductivity. The main difference is the exchange of ions that takes place in the electrodiffusion process, i.e., specific ions (protons or alkalis) are brought into the sample to replace interstitial alkalis that had been trapped at defects during the crystal growth process. Electrical conductivity in quartz is ionic and one dimensional with monovalent ions moving along the Z axis channels. The number of mobile ions is governed by the number of traps, such as substitutional aluminum. If only one kind of trap, say aluminum, and one species of interstitials are present, the number per unit volume of mobile ions should be given by

$$n = (C_T/2)^{1/2} N_0 \exp(-E_A/2kT) \quad (1)$$

where C_T is the mole fraction of traps, N_0 is the number of SiO_2 's per unit volume and E_A is the association

energy between the ion and the trap.¹⁸ The ion mobility¹⁹ is given by

$$\mu = \frac{ed^2}{kT} \nu \exp(-E_m/kT) \quad (2)$$

where d is the jump distance, ν the jump frequency and E_m is the activation energy for interstitial migration. Thus, the conductivity $\sigma = ne\mu$ can be expressed as

$$\sigma T = A \exp(-E/kT) \quad (3)$$

where

$$E = E_m + \frac{1}{2} E_A \quad (4)$$

and

$$A = \frac{N_O ed^2 \nu (C_T/2)^{1/2}}{kT} \quad (5)$$

According to this analysis, a plot of $\log \sigma T$ versus T^{-1} should yield a straight line. Jain and Nowick²⁰ have reported such behavior and have found activation energies, E , near 1.4 and 1.0 eV for synthetic and natural quartz, respectively. The electrodiffusion process is expected to be more complicated since ion exchange is involved; however, Martin et al.²¹ found a similar behavior for the ionic currents once the sweeping process has neared completion.

A systematic study of the electrodiffusion process in quartz is underway at Oklahoma State University. In addition to the effects of electrodiffusion on the point defects, we are also investigating the effects of sweeping on extended defects and fracture in quartz.

Experimental Procedure

The samples used in this study were cut from unswept pure Z-growth lumbered bars of synthetic quartz. These bars included material from Thermodynamics, Sawyer Research Products (Electronic, Premium Q, and Super Premium Q grades), and an aluminum-doped synthetic bar. The various samples were identified according to the notation established by Markes and Halliburton.²² Z plates 3 mm thick and 15 mm x 17 mm in the X and Y directions were used in the electrodiffusion experiments. Such samples are a convenient size for the sweeping apparatus and for subsequent infrared evaluation. Polished AT-cut rectangular plates and plano-convex AT-cut resonator blanks were swept for the mechanical and acoustic loss measurements, respectively.

Prior to sweeping, the samples were given an optical polish. Our sweeping process, which was carried out in the system previously described,²¹ begins by vapor depositing gold electrodes on the two sample faces. If Li^+ or Na^+ is to be swept, the appropriate salt is first evaporated on one side, then a gold electrode is deposited over the salt. The electrodiffusion is carried out in the desired atmosphere at a temperature between 480°C and 500°C. Alkali sweeps (i.e., taking alkalis into the crystal) are done in a vacuum while an H_2 atmosphere is used for sweeping protons into the crystal. The applied electric field is approximately 15 V/cm for alkali sweeps and 2000 V/cm for hydrogen. A digital multimeter and a laboratory computer are used to monitor the sweeping current as a function of time. Hydrogen sweeps are carried out until the current reaches a steady state; this usually takes less than 24 hours. Alkali sweeps

continue until a sufficient number of ions have been transported; this usually takes less than two hours. As a sweep nears completion, we believe only one species of ion is being transported. Then, while the furnace temperature is slowly programmed down, the sample current is recorded as a function of temperature. This latter data gives the effective ionic conductivity as a function of temperature for the specific ion being swept.

Eight AT-cut polished rectangular plates approximately 15 mm x 20 mm x 2 mm thick were fabricated from each of the bars EG-C, PQ-C, TD-A, and PQ-D. Four plates from each of the bars EG-C, PQ-C, and TD-A were hydrogen swept by the process described above. (Bar PQ-D was Sawyer Swept Premium Q grade material.) Two unswept and two swept plates from each of the bars were then etched in a concentrated ammonium bifluoride solution to reveal their etch tunnels. Except for PQ-D, each bar gave pairs of AT-cut plates representing each of the following four categories: unswept, unswept-etched, H-swept, and H-swept-etched. Samples for the X-parallel fracture tests were 1.5 mm wide with their long axis parallel to X and were cut from one plate of each pair. Z'-parallel fracture test samples, also 1.5 mm wide, were cut perpendicular to the X axis of the other member of each pair. The fracture stress was measured using a four point bend jig mounted in an Instron testing machine.²³ The four-point bend test subjects the sample (at least in the elastic region) to a pure bending moment between the inner loading points. The stress varies from tension along the long axis of the sample on the top to compression on the bottom. The results from at least seven samples from each plate were averaged to obtain the values reported.

Results and Discussion

The electrodiffusion current for hydrogen sweeping usually shows a significant initial decay followed by a steady current as shown in Fig. 1 for sample PQ-G. The results for a Li^+ sweep on a similar sample are also illustrated in Fig. 1. Both the H^+ and Li^+ sweeps show nearly the same final current even though the applied field for the H^+ run is approximately 2000 V/cm while 15 V/cm was used for the Li^+ sweep. This result and similar results for Na^+ show that H^+ is much less mobile than alkali ions. We believe that the large initial current and subsequent decrease that takes place in a H^+ sweep is caused by the removal of the very mobile alkali ions from the sample.

Specific ions, Li^+ from the LiCl source electrode or H^+ from the hydrogen atmosphere, are brought into the crystal at the positive electrode. These specific ions then migrate along the Z axis channel and replace the interstitial alkalis originally trapped at the aluminum and similar sites. Thus, as the process continues a condition is reached where only one species of ion is migrating along the Z axis channel. The number of migrating ions is determined by the number of traps present in the sample. Consequently, the conductivity should be described by Eqs. 1 through 5. Figure 2 shows a plot of $\log \sigma T$ versus $1000/T$ for Li^+ , Na^+ , and H^+ sweeps of previously unswept samples taken from bar HA-A. Bar HA-A, which was doped with aluminum during growth, contains approximately 50 ppm Al, but it contains very few OH^- related growth defects.²¹ The apparent ionic conductivity of H^+ is seen to be much less than that of either Na^+ or Li^+ . The upper half of Fig. 3 compares the apparent ionic conductivity of Li^+ in the HA-A sample, in Premium Q sample PQ-I (≈ 10 ppm Al),

and in Super Premium Q sample SP-A (<1 ppm Al). The bottom half of Fig. 3 compares the conductivity of H^+ in sample HA-A and PQ-I. For both Li^+ and H^+ runs, the conductivity clearly increases with increasing Al content.

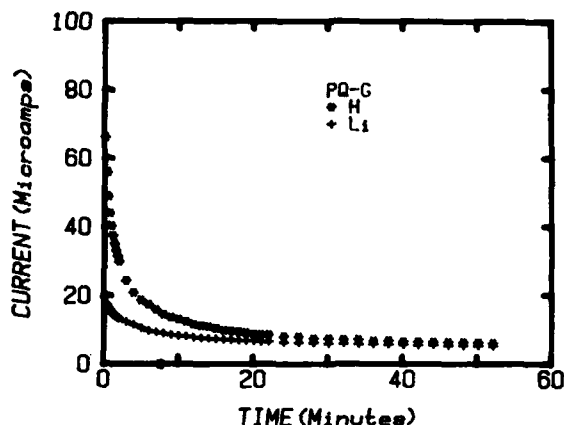


Figure 1. The electrodiffusion currents for H^+ and Li^+ sweeps on samples from Sawyer Premium Q bar PQ-G are shown as functions of time. We believe that the large initial decay during a H^+ sweep is caused by the replacement of the mobile alkalis with protons. The applied fields were 2000 V/cm for H^+ and 15 V/cm for Li^+ .

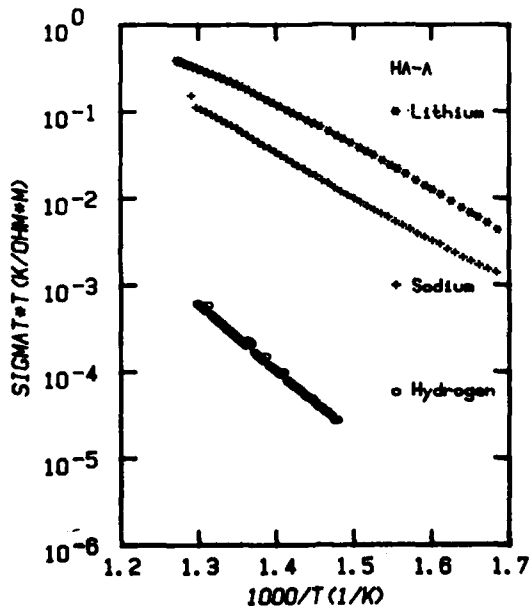


Figure 2. The product of the apparent ionic conductivity times temperature in samples taken from the aluminum doped quartz HA-A. In all cases the conductivity is thermally activated.

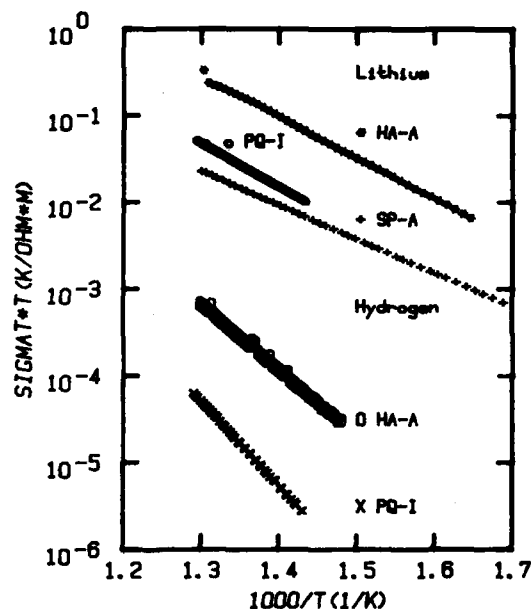


Figure 3. The upper portion of the graph compares σT vs $1000/T$ for Li^+ for samples with approximately 50 ppm Al (HA-A), 10 ppm Al (PQ-I), and less than 1 ppm Al (SP-A). The lower curve compares σT vs $1000/T$ for samples with approximately 50 ppm Al and 10 ppm Al.

The activation energy, E , which includes both the migration energy and the association energy can be determined from the slope of the log σT vs $1000/T$ curves. Table I gives the activation energies determined for Li^+ , Na^+ , and H^+ for a number of different sweeping runs.

Table I. Sweeping Activation Energies (eV)

Sample	Li^+	Na^+	H^+
SP-A	0.76	0.99	1.75
PQ-G	1.14	1.25	1.95
PQ-I	0.98	1.13	1.93
HA-A	0.94	1.00	1.52
Literature	0.76-0.86 ²⁴	1.03-1.2 ²⁴	1.71-1.75 ²⁵

The activation energies vary considerably between the different quartz samples. The migration energy for a given ion is expected to be the same for all quartz, but different samples will have different traps so the association energy may fluctuate considerably. Bar SP-A is a very clean bar of quartz while both PQ-G and PQ-I have quite a bit of aluminum and other defects. Bar HA-A has a lot of aluminum but it is otherwise relatively defect free. Since the Al should dominate over other traps in bar HA-A, the activation energies reported for it may represent the effects of Al as the only trap. Our activation energies for the alkali ions are slightly lower than the 1.3 to 1.4 eV found by Jain and Nowick²⁰ in their ionic conductivity studies.

Vig¹⁷ has reported that sweeping lowers the etch tunnel density and improves the mechanical strength of quartz resonator blanks. We have carried out etching and four-point bend fracture tests on a number of unswept and H⁺-swept synthetic quartz samples. Table II gives the etch tunnel density for unswept and H⁺-swept samples from three bars of synthetic quartz.

Table II. Etch Tunnel Densities (Tunnels/cm²)

	EG-C	PQ-C	TD-A
Unswept	1300	1365	1710
H ⁺ -swept	310	380	820

These results show that sweeping significantly reduces the etch tunnel density. However, the basic extended defect is still present. The sweeping has most likely replaced the alkali ions trapped in the dislocation network with protons.

Four-point bend fracture tests were made on samples cut parallel (X-parallel) and perpendicular (Z'-parallel) to the X axis of AT-cut synthetic quartz plates. Figure 4 shows the fracture stress results for unswept, unswept and etched, H⁺-swept, and H⁺-swept and etched X-parallel samples from Sawyer swept bar PQ-D. Figure 5 shows the results for Z'-parallel samples. Our results as shown in Figure 4 and 5 indicate that swept and etched samples have the highest fracture strength. Etching alone might be expected to improve fracture resistance since it should improve the surface finish. However, the unswept-etched X-parallel samples showed a reduced strength while for Z'-parallel the unswept-etched samples were stronger.

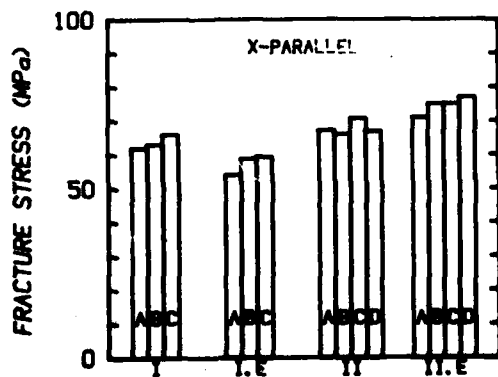


Figure 4. The fracture stress as measured under four-point bend is shown for test sample bars cut from AT-plates parallel to the X axis. I corresponds to as-received material while II corresponds to H⁺-swept; E indicates etched. The samples are A (Sawyer EG-C), B (Sawyer PQ-C), C (Thermodynamics, TD-A) and D (Sawyer Swept PQ-D). Etched H⁺-swept samples show a higher fracture strength.

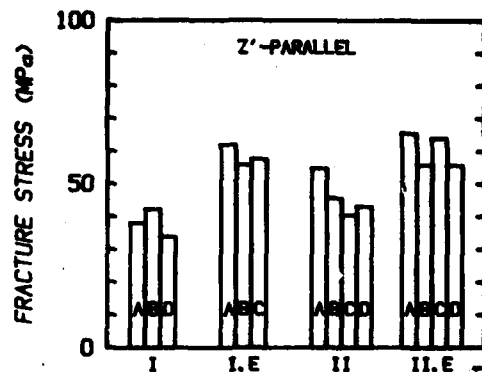


Figure 5. The fracture stress as measured under four-point bend is shown for test sample bars cut from AT-plates perpendicular to the X axis. The sample designations are the same as in Fig. 4.

Conclusions

The mobility of H⁺ in the electrodiffusion process is much less than that of the alkali ions. As expected, the apparent ionic conductivity for a specific ion scales roughly with the aluminum content of the quartz. Since the activation energy for a specific ion depends upon the quality of the quartz, more than one type of trap must be involved. H⁺-sweeping significantly reduces the formation of etch tunnels. Etched H⁺-swept samples showed an improved fracture strength.

References

1. L. E. Halliburton, N. Koumvakalis, M. E. Marques and J. J. Martin, *J. Appl. Phys.* **52**, 3565 (1981).
2. D. R. Koehler and J. J. Martin, *Proc. 37th Ann. Freq. Control Symposium*, (present volume).
3. D. B. Fraser, In *Physical Acoustics*, W. P. Mason, ed. Vol. V, pp. 54-110, Academic Press, NY (1968).
4. D. R. Koehler, *Proc. of the 35th Ann. Freq. Control Symposium*, USAERADCOM, Ft. Monmouth, NJ, p. 322, May (1981).
5. H. G. Lipson, A. Kahan, R. N. Brown and F. K. Euler, *Proc. 35th Ann. Freq. Control Symposium*, USAERADCOM, Ft. Monmouth, NJ, p. 329, May (1981).
6. J. C. King, *Bell System Techn. J.* **38**, 583 (1959).
7. A. Kats, *Phillips Res. Rpts.* **17**, 133 (1962).
8. D. B. Fraser, *J. Appl. Phys.* **35**, 2913 (1964).
9. G. Kreft, *Rad. Effects* **26**, 249 (1975).
10. R. M. Brown, J. O'Connor and A. F. Arrington, *Nat. Res. Bull.* **15**, 1063 (1980).
11. T. M. Flanagan and T. F. Wrobel, *IEEE Trans. Nucl. Sci.* **NS-16**, 130, Dec. 1969).

12. J. C. King and H. H. Sanders, IEEE Trans. Nucl. Sci. NS-19 23 (1972).
13. P. Pellegrini, F. Euler, A. Kahan, J. M. Flanagan and T. F. Wrobel, IEEE Trans. Nucl. Sci. NS-25, 1267 (1978).
14. G. W. Arnold, Proc. 11th Ann. Freq. Control Symposium, p. 119, May (1957).
15. J. W. Nielsen and P. G. Foster, Am. Mineral 45, 299 (1960).
16. W. J. Spencer and K. Haruta, J. Appl. Phys. 35, 2368 (1964).
17. J. R. Vig, J. W. LeBus and R. L. Filler, Proc. 31st Ann. Freq. Control Symposium, USAERADCOM, Ft. Monmouth, NJ, p. 131, May 1977.
18. A. D. Franklin, in Point Defects in Solids, J. H. Crawford and L. M. Slifkin, eds., Plenum, NY, Vol. 1, Ch. 1, (1972).
19. A. B. Lidiard, Hand. of Physik 20, 246 (1957).
20. H. Jain and A. S. Nowick, J. Appl. Phys. 53, 477 (1982).
21. J. J. Martin, L. E. Halliburton, R. B. Bossoli and A. F. Armington, Proc. 36th Ann. Freq. Control Symposium, USAERADCOM, Ft. Monmouth, NJ, p. 77, May 1982.
22. M. E. Markes and L. E. Halliburton, J. Appl. Phys. 50, 8172 (1979).
23. W. A. Sibley, C. T. Butler, J. R. Hopkins, J. J. Martin and J. A. Miller "Growth and Hardening of Alkali Halides for use in Infrared Laser Windows," Technical Report, AFCRL-TR-0342 (1973).
24. J. Verhoogen, Am. Mineral, 37, 637 (1952).
25. H. Wendon, Am. Mineral, 42, 859 (1957).

Acknowledgements

This work was supported in part by U.S. Air Force Contract F 19628-80-C-0086, monitored by A. F. Armington, RADC/ESM, Hanscom AFB, MA, 01731. Brinda Subramaniam was a Visiting Fulbright Scholar.



AD P 002471

Aluminum and Hydroxide Defect Centers in Vacuum Swept Quartz

Herbert G. Lipson, Alfred Kahan, and John O'Connor

Solid State Sciences Division
Rome Air Development Center
Hanscom Air Force Base, Bedford, MA 01731

Summary

Vacuum swept quartz is characterized by infrared spectroscopic measurements both along the sweeping axis and by scanning small regions of the sample perpendicular to the sweeping axis. We observe large variations in the strengths of the as-grown OH^- and Al-OH^- defect related bands between the anode and cathode. In a vacuum swept sample, initially containing aluminum-sodium centers, a region depleted of as-grown OH^- extends a short distance into the sample, with a corresponding sharp gradient in the Al-OH^- related band. Electron spin resonance data indicates that in this region of the sample 35 percent of the defects are in the form of aluminum-hole centers. A sample, initially containing aluminum-lithium centers, was air swept and then vacuum swept. Successive periods of vacuum sweeping deplete hydroxide defects near the anode and substantially decrease them near the cathode. Vacuum sweeping effects are accelerated by removing the surface layers and repolishing the crystal after successive short sweeping periods.

Key Words: Quartz, Impurities, Defect Centers, Sweeping, Infrared Spectroscopy.

Introduction

The most common characterization techniques of point defect centers in quartz include optical spectroscopy, anelastic loss measurements, and electron spin resonance. It is generally accepted that the low temperature 3581 and 3366 cm^{-1} infrared bands in α -quartz are associated with as-grown hydroxide and Al-OH^- defect centers, respectively. These bands have been used extensively to monitor changes induced by sweeping (electrodifusion) or radiation.¹⁻⁴ The customary experimental technique is to measure band strengths along the sample growth or sweeping axis, the z -axis. This type of measurement averages the band strengths of defect centers along the z -axis. The local effects of sweeping are not revealed by these measurements. Kats,¹ and subsequently Krefft,⁵ have studied the sweeping process by scanning the infrared spectrum of small regions of the sample perpendicular to the sweeping axis. In Ref. 6 we have described this non-destructive characterization technique for determining the spatial distribution of the defect centers between the anode and cathode. This measurement technique is also useful in evaluating the uniformity of radiation induced defects.

Several considerations indicate that sweeping quartz in a vacuum atmosphere improves material and device properties. However, previous investigations indicated that effects are non-uniform, and the use of crystals processed in this manner may lead to unpredictable results. To properly characterize vacuum sweeping, we performed infrared measurements perpendicular to the sweeping axis and scanned small regions of the sample between the anode and cathode. Infrared spectra of vacuum swept synthetic crystals, initially containing interstitial lithium or sodium ions compensating substitutional aluminum impurity, Al-Li^+ or Al-Na^+ , Ref. 6, showed:

- (1) As-grown OH^- concentration near the anode is reduced, and subsequently, with increasing sweeping time, it is depleted. Band strength varies greatly between the anode and cathode, and after 12 to 24 hours of sweeping the depleted region extends approximately 3 to 4 mm into the crystal.
- (2) Al-OH^- is formed near the anode, and it is progressively swept towards the cathode. Eventually, with increasing sweeping time, Al-OH^- is reduced in the region where the as-grown OH^- is depleted.
- (3) The compensation processes are the same for Al-Na^+ and Al-Li^+ , but longer sweeping times are necessary to observe the same magnitude of effects in crystals with large aluminum concentration.

This paper reports additional results on the spatial distribution of defect centers in vacuum swept quartz. After vacuum sweeping, the anode region of a sample with an initial large Al-Na^+ concentration was completely depleted of as-grown OH^- , but Al-OH^- was still present. Several sections were cut from this crystal region and were examined for aluminum-hole centers, Al-e^+ , by ESR. Results show the largest percentage of Al-e^+ ever observed for vacuum swept quartz. Another crystal, containing Al-Li^+ , was air swept for several weeks and then vacuum swept for different time periods to a total of 45 days. Results show large defect center gradients along the sweeping axis, and depletion of both as-grown OH^- and Al-OH^- near the anode. We also include some results on spatial distribution of defects in irradiated samples.

Experimental Procedures

The unswept and swept Premium-Q and High-Q samples used in this investigation were grown at Sawyer Research Products (SARP) by the same process, except that for Premium-Q, lithium salt was added to the mineralizer. The primary impurity in both High-Q and Premium-Q quartz is substitutional aluminum, Al^{3+} . In High-Q material Al^{3+} is compensated with sodium, $\text{Al}-\text{Na}^+$, and in Premium-Q with lithium, $\text{Al}-\text{Li}^+$. The unswept samples were rectangular sections, approximately $1.6 \times 1.8 \times 2.0$ cm, with parallel x-, y-, and z-faces. An air swept bar, $1.0 \times 1.2 \times 5.0$ cm was also used for vacuum sweeping studies. Platinum foil electrodes were placed on the z-faces, and the sweeping was performed at 10^{-5} torr and 500°C , with an electric field of 1000 V/cm .

Infrared transmissions were measured between 3100 and 3700 cm^{-1} at 85 K with a Digilab FTS-14 Fourier spectrophotometer. The full beam was utilized in the direction of sweeping. Specific areas normal to the sweeping direction were scanned by moving a $3.2 \times 1.0\text{ mm}$ aperture along the x- or y-face. Additional measurements were made with the 3 mm focused beam of a Nicolet SX-170 Fourier spectrophotometer. The results obtained with the Digilab and Nicolet spectrophotometers are in good agreement for band strengths measured along the optic axis, but instrument polarization differences affect measurements normal to the axis. The Nicolet instrument data are normalized to correct for these polarization differences.

Figure 1 shows the absorption spectra of samples measured normal to the growth (sweeping) axis obtained with unpolarized and polarized radiation. At 85 K , four principal bands associated with as-grown OH^- impurities in synthetic quartz have vibration frequency peaks at 3348 , 3396 , 3438 , and 3581 cm^{-1} . The four bands behave very similarly as a function of sweeping or irradiation. Figures 1a and 1c show that for samples with relatively low OH^- concentration, only the 3581 cm^{-1} band has a strong narrow peak and an easily determined low background absorption, whereas the absorption of the other three bands are broad and are superimposed on Si-O lattice bands. For these reasons we select the 3581 cm^{-1} band to monitor changes in as-grown OH^- .

Figures 1b and 1d show that after irradiating the Premium-Q sample with 1 Mrad of ^{60}Co , or vacuum sweeping the High-Q sample for 24 hours, the OH^- bands are depleted in the regions examined. At the same time, two bands associated with $\text{Al}-\text{OH}^-$, at 3306 and 3366 cm^{-1} , are formed. We choose the stronger 3366 cm^{-1} band to monitor changes in this defect center. Also, we are mainly interested in relative changes in as-grown OH^- and $\text{Al}-\text{OH}^-$ induced by sweeping or irradiation, and for this purpose it is sufficient to compare absorption peak heights instead of calculating changes in the integrated areas.

Both the 3581 and 3366 cm^{-1} bands have large dichroic ratios with maximum strengths when the electric vector is perpendicular to the optic axis, $E \perp c$. The band strengths shown in Fig. 1 measured with unpolarized radiation are close to those ob-

tained with a wire grid infrared polarizer set at 45° . For the specific sample orientation used, maximum band strengths, approximately equivalent to the values measured along the optic axis, are obtained with a 0° setting of the polarizer, and are designated as $E \parallel c$. Minimum band strength is observed for 90° rotation of the polarizer, indicating that the bands are completely polarized. Unpolarized and $E \perp c$ absorption data differ only in band strength, but for reasonable signal-to-noise the polarized measurements require long scanning times owing to low polarizer transmission. In this paper we are interested in relative changes in the defect centers, and we present only the results of the unpolarized radiation.

Results and Discussion

Point Defects

We are concerned with the point defect structure of as-received high quality synthetic quartz, the effects of ionizing radiation, and changes due to sweeping in an air or vacuum atmosphere. The point defect models associated with the most common substitutional and interstitial impurities in α -quartz, and the influence of sweeping or irradiation on these defects, have been reviewed by a number of investigators.⁷⁻¹¹

The major point defect in as-received high quality quartz is the aluminum-metal defect center, $\text{Al}-\text{M}^+$ ($\text{M} = \text{Li}, \text{Na}, \text{or K}$). Irradiating a sample at room temperature will dissociate $\text{Al}-\text{M}^+$ and form $\text{Al}-\text{OH}^-$ and aluminum-hole centers, $\text{Al}-e^+$. In this process the internal as-grown OH^- is the source of hydrogen ions. The dissociated alkali metal ions have not been observed by any of the standard characterization techniques. The relative ratio of $\text{Al}-\text{OH}^-$ to $\text{Al}-e^+$ defects depends on the type of electron traps available in the crystal. The $\text{Al}-e^+$ center thermally decays near 270°C , and there are indications that in this process the defect is converted to $\text{Al}-\text{OH}^-$. In turn, the $\text{Al}-\text{OH}^-$ center thermally decays in the 350°C to 450°C range, and the original $\text{Al}-\text{M}^+$ center is re-formed.

There is general agreement, that sweeping synthetic quartz in an air atmosphere, where water vapor is present, or in a hydrogen atmosphere, at temperatures between 500°C and 550°C , will:

- (1) Dissociate $\text{Al}-\text{M}^+$ and charge compensate the substitutional Al^{3+} with hydrogen to form an $\text{Al}-\text{OH}^-$ center which is stable below the α - to β -quartz inversion temperature 573°C .
- (2) Sweep the alkali metal ions towards the cathode, and physically remove most of them from the crystal.
- (3) Not significantly affect the as-grown OH^- bands.

It is, however, uncertain whether $\text{Al}-\text{OH}^-$ is formed in a one-step process, with the hydrogen provided by the external source, or whether the hydrogen ion is derived internally from the as-grown OH^- , with the external hydrogen replenishing the as-grown OH^- .

In principle, one expects that sweeping a crystal in a vacuum atmosphere, with no external source for hydrogen ions, will dissociate Al-M^+ and form the Al-e^+ center. The role of as-grown OH^- in this process is uncertain. There are several possibilities which must be considered:

- (1) Al-e^+ will form depending on the availability of electron traps.
- (2) As-grown OH^- acts as an internal source of hydrogen ions and stable Al-OH^- is formed at the dissociated Al-M^+ center.
- (3) Both defects, Al-OH^- and Al-e^+ , are formed in the same crystal region. Undetermined factors govern the preference of forming one defect center over the other.
- (4) The Al-OH^- center exists as an intermediate state, and after the as-grown OH^- is depleted this center becomes unstable and is also converted to Al-e^+ . For this process to occur it is necessary to assume that unfilled electron traps exist in the crystal.
- (5) Vacuum sweeping effects are either uniform across the sample, or, defect center gradients arise between the anode and cathode. These gradients will depend on the initial defect concentration and the sweeping parameters.

Experimental verification of some of these considerations will determine whether vacuum sweeping is effective in removing hydroxide defects and produce room temperature stable Al-e^+ .

Radiation Effects

Figure 2 shows radiation induced changes in as-grown OH^- and Al-OH^- bands on a crystal containing Al-Li^+ after exposure to 7 Mrads of ^{60}Co radiation.¹² The sample was measured with the infrared beam parallel to the crystal growth axis, between the +z and -z-faces. The figure shows the usual decrease of as-grown OH^- and the formation of Al-OH^- bands, and that the effects are saturated by 1 Mrad.

Figure 3 shows band strength data after 1 Mrad irradiation measured perpendicular to the crystal growth axis. The initial as-grown OH^- distribution is relatively uniform for about one half of the sample, and there is an approximately a 50 percent band strength variation for the other half of the crystal. The effect of room temperature irradiation is an almost uniform decrease in the as-grown OH^- across the entire sample. At one end of the crystal where the initial OH^- concentration was low, OH^- is depleted after irradiation. The Al-OH^- concentration is relatively uniform throughout the entire sample. The combination of uniform Al-OH^- and residual as-grown OH^- in some parts of the sample indicates that the formation of Al-OH^- is limited by the aluminum ion concentration. In another sample, with a high aluminum ion concentration, 1 Mrad irradiation depletes all the OH^- , but Al-OH^- varies by a factor of two across the sample. Considering the fact that OH^- is depleted across the entire sample, an explanation of the Al-OH^-

variation must invoke additional mechanisms, for example, radiation induced hydrogen migration to remote sites.

These considerations are complicated by the fact that Al-e^+ centers are also produced during the irradiation. The sequence of formation and the factors that determine the relative concentration of Al-OH^- and Al-e^+ has not been established. Does Al-e^+ form first and then Al-OH^- when the electron traps are saturated, or are both defects created simultaneously? Regardless of the sequence, the final percentage of Al-e^+ centers is more probably determined by the concentration of the electron traps rather than the aluminum concentration. The interplay between the defect centers may be clarified by studying infrared, ESR, and the optical bands between 2 eV to 4 eV as a function of radiation dose and annealing.

Aluminum-hole Centers

Figure 4 shows the effects of sweeping a crystal containing Al-M^+ . This figure is an updated version of Fig. 11 of Ref. 6. The air swept portion of the figure is identical with the original data. Sweeping the sample in an air atmosphere forms Al-OH^- , but it has no significant effect on as-grown OH^- . The sweeping process consists of a gradual movement of ions towards the cathode. This migration of alkali and hydrogen ions towards the cathode is retarded by a barrier such as a space charge build up, or ion clogging in the channels close to or on the cathode surface. This can be inferred by comparing the 3, 6, and 12 hour sweeping time curves. After 3 hours of sweeping a 0.5 mm thick surface layer was ground from the cathode side of the sample, and the crystal repolished. The extent of Al-OH^- penetration within the sample, between the 3 and 6 hours of sweeping, is attributed to the physical removal of this layer. The layer was not removed after the 6 hour sweeping and the change in penetration between 6 and 12 hours is small. Measurement of this sample along the sweeping axis indicates Al-OH^- , but does not reveal the Al-OH^- non-uniformity or the incomplete sweeping of the sodium ions. The radiation response of a resonator fabricated from this bar, with a highly non-uniform sodium distribution, would be unpredictable.

The right hand portion of Fig. 4 shows the corresponding data for sweeping in a vacuum environment. We include results for an additional 12 hours of sweeping, a total of 24 hours. As-grown OH^- is depleted near the anode and Al-OH^- is formed in this region, with effects extending through one half of the sample, 8 mm. The surface layers were not removed after the 12 hours of sweeping, and changes between 12 and 24 hours are insignificant. This again demonstrates the need for removing the surface layers after each sweeping period.

The primary experimental technique used to measure Al-e^+ is electron spin resonance. This topic has been reviewed extensively by Hall¹³ and more recently by Markes and Halliburton.¹⁴ The existence of Al-e^+ in irradiated crystals, initially containing Al-M^+ , has been confirmed by ESR. However, to the best of our knowledge, there is no reported ESR evidence for Al-e^+ in synthetic quartz.

swept in a vacuum environment below the α - to β -quartz inversion temperature, 573 °C. The optical absorption in quartz, observed for irradiated crystals between 2.0 eV and 4.0 eV, which gives rise to a smoky color, has been correlated with the Al-e^+ center.¹⁵ The darkening of vacuum swept quartz, and the propagation of a color front between the anode and cathode as a function of sweeping time and temperature, has been reported.^{5,16} (References to previous investigations are given in Ref. 5). In addition to direct ESR measurements, the coloration of vacuum swept quartz may also be used as an indicator of Al-e^+ .

In our experiments, after 24 hours of vacuum sweeping, the sample did not show any darkening. Several 2.5 mm thick sections of this crystal, containing the anode region which is depleted of as-grown OH^- , were cut parallel to the z-face, and examined by ESR.¹⁷ The results show that over most of the anode the sweeping is 98 percent complete, that is, the Al-Na^+ is dissociated, and that 35 percent of the defects are in the form of Al-e^+ . This is the largest percentage of aluminum-hole centers ever observed for vacuum swept crystals. Previous inconclusive ESR results on vacuum swept crystals may be attributed to the fact that the ESR sample position relative to the electrodes was not selected properly.

The results of this study show that vacuum swept quartz, at least the anode section, contains Al-e^+ , Al-OH^- , and is depleted of as-grown OH^- . Al-e^+ coexists with Al-OH^- in the same region of the crystal, and Al-e^+ is formed before all Al-OH^- is exhausted. However, from infrared measurements only, we can not establish whether both defect centers are formed simultaneously, or whether Al-OH^- acts as an intermediate state, and Al-e^+ is formed after the internal hydrogen source replenishing the Al-OH^- is depleted.

The results of vacuum sweeping high quality synthetic crystals containing Al-Na^+ are similar to those obtained by Krefft⁵ for natural quartz. Krefft investigated changes in the optical and infrared spectra induced by vacuum sweeping, and found that the depletion of the as-grown OH^- , the relative increase in Al-OH^- , and the propagation and position of a color front between the anode and cathode are correlated. The investigations carried out by Krefft were performed on natural quartz, which had much higher overall absorption levels than the synthetic crystals used for our investigations. On this basis it is not surprising that we do not observe the visual coloration associated with Al-e^+ . Natural crystals also show substantial Al-OH^- bands before irradiation or sweeping, whereas, no Al-OH^- exists in our as-received synthetic quartz.

Sweeping Effects

Figure 5 shows experimental results for a Premium-Q, Li^+ -compensated sample, which was air swept for several weeks at Sawyer Research Products, and then vacuum swept at our facility for defined time periods. We assume that after this period of air sweeping the as-grown OH^- and Al-OH^- centers were fully saturated. We show data for the relevant bands for the anode and cathode regions. Both

the as-grown OH^- and Al-OH^- bands show strength differences between the anode and cathode, with the cathode region having a larger percentage of OH^- . After vacuum sweeping for 14 days, the OH^- content of the crystal is lowered both at the anode and the cathode. An additional 14 days of sweeping did not affect the bands, indicating the presence of a barrier against migration of the hydrogen ions. Following this sweeping period, a 100 μm surface layer was removed from the anode and cathode, and the sample was swept for an additional 17 days. The anode region has been depleted of both as-grown OH^- and Al-OH^- , and there is also a substantial decrease in the OH^- band at the cathode. Figure 6 shows more detailed results of band strength variation between the anode and cathode at the end of 45 days of sweeping. The region free of hydrogen related defects extends 3 to 4 mm into the crystal, a section large enough for resonator fabrication.

The depletion of both as-grown OH^- and Al-OH^- bands in the anode region has also been observed by Krefft⁵ in a sample that was completely colored and presumably contains Al-e^+ . Our sweeping experiments were carried out at 500 °C while those of Krefft were at 700 °C, well above the 573 °C inversion temperature. The trigonal symmetry of α -quartz changes to a sixfold symmetry for β -quartz, and may allow β -quartz to accommodate different, or additional, point defects.¹⁸ Also, Fraser¹⁹ showed that for natural quartz the anelastic losses associated with Al-M^+ are discontinuous through the inversion, and that Al-M^+ thermally dissociates in the β -phase between 573 °C and 700 °C without sweeping. Consequently, defect centers created by vacuum sweeping at 700 °C may be different from those induced by sweeping below 573 °C.

The removal of the surface layers after short sweeping times, for vacuum as well as air sweeping, is a critical element in facilitating the removal of hydrogen related defects. This is especially true for samples containing heavy concentration of Al-M^+ centers. Higher electric fields will also accelerate sweeping. In our experiments, the electric field was maintained consistently at 1000 V/cm, but it is known that higher fields can be safely applied. It is also known¹⁶ that vacuum sweeping proceeds more rapidly at elevated temperatures, but from a practical point of view it is critical that the crystal be maintained below 573 °C, because the α - to β -quartz inversion introduces electrical twinning, which renders the material useless for piezoelectric resonators.

Conclusions

1. It is possible to obtain crystal regions completely free of hydrogen related defects by sweeping the sample in a vacuum atmosphere.
2. Stable aluminum-hole centers are introduced by vacuum sweeping at 500 °C.
3. For vacuum swept quartz aluminum-hole centers coexist with Al-OH^- centers in the same region of the crystal.
4. The most efficient sweeping process is to sweep for repeated short time periods, and remove sample surface layers after each sweeping period.

References

1. A. Kats, Phillips Res. Reports, **17**, 133(1962).
2. D.M. Dodd and D.B. Fraser, J. Phys. Chem. Solids, **26**, 673(1965).
3. H.G. Lipson, F. Euler, and A.F. Armington, 32nd Annual Symposium on Frequency Control, (ASFC), **11**(1978).
4. W.A. Sibley, J.J. Martin, M.C. Wintersgill, and J.D. Brown, J. Appl. Phys., **50**, 5449(1979).
5. G.B. Krefft, Radiation Effects, **26**, 249(1975).
6. F. Euler, H.G. Lipson, A. Kahan, and A.F. Armington, 36th ASFC, 115(1982).
7. D. B. Fraser, Physical Acoustics, W. P. Mason, ed., (Academic Press, NY), Vol. V, 59(1968).
8. D.L. Griscom, 33rd ASFC, 98(1979).
9. L.E. Halliburton, M.E. Markes, and J.J. Martin, 34th ASFC, 1(1980).
10. L. E. Halliburton, N. Koumvakalis, M.E. Markes, and J.J. Martin, J. Appl. Phys., **52**, 3565(1981).
11. F. Iwasaki and M. Kurashige, Ferroelectrics, **43**, 43(1982).
12. The authors thank Lester Lowe, RADC/ESR, for irradiating the crystals used in this investigation.
13. J.A. Weil, Radiation Effects, **26**, 261(1975).
14. M.E. Markes and L.E. Halliburton, J. Appl. Phys., **50**, 8172(1979).
15. N. Koumvakalis, J. Appl. Phys., **51**, 5528(1980).
16. J. Flesche and J. Leitz, Neues Jahrb. fur Miner. Abh., **109**, 238(1968).
17. The ESR measurements on this and other samples were performed by L. Halliburton at Oklahoma State University.
18. G. Dolino and J.P. Bachheimer, Ferroelectrics, **43**, 77(1982).

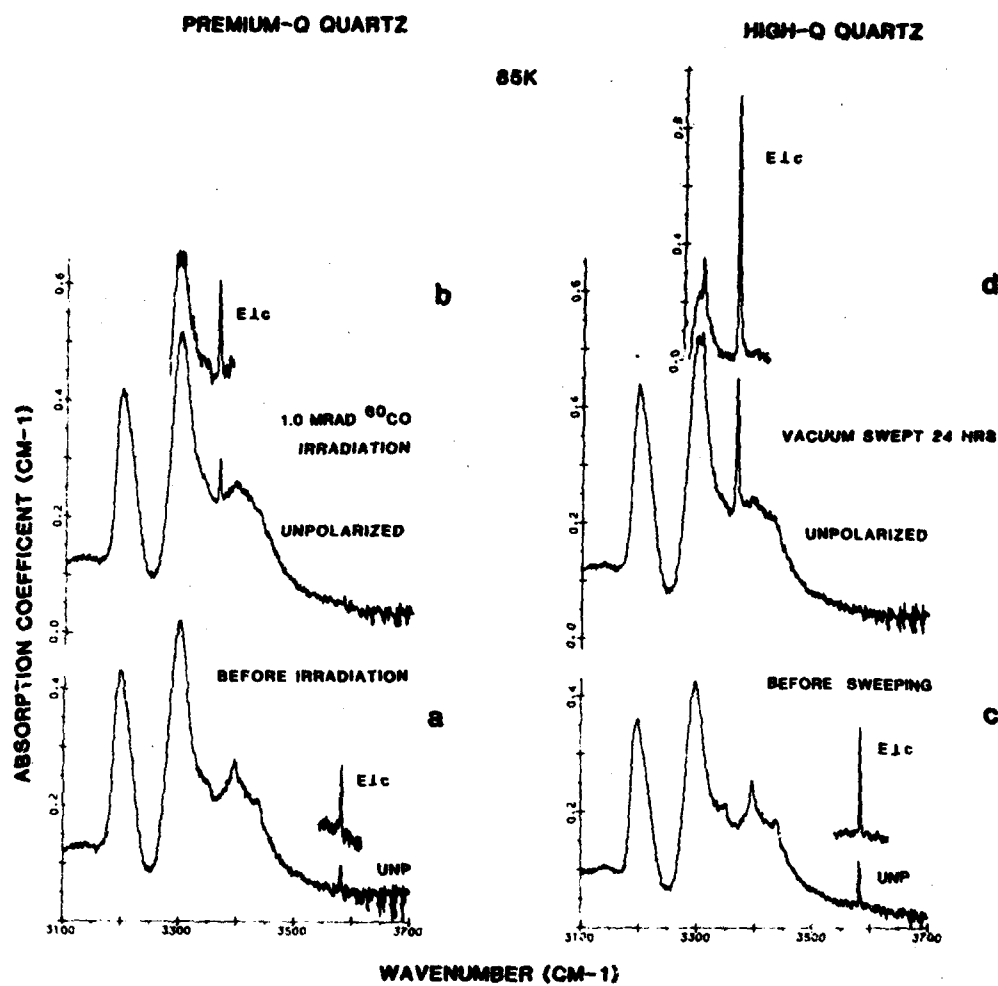


Figure 1. Unpolarized and polarized infrared absorption spectra for Premium-Q and High-Q quartz measured perpendicular to the crystal growth axis, (a) Premium-Q as-received, (b) Premium-Q after 1 Mrad ^{60}Co irradiation, (c) High-Q as-received, and (d) High-Q after vacuum sweeping for 24 hours.

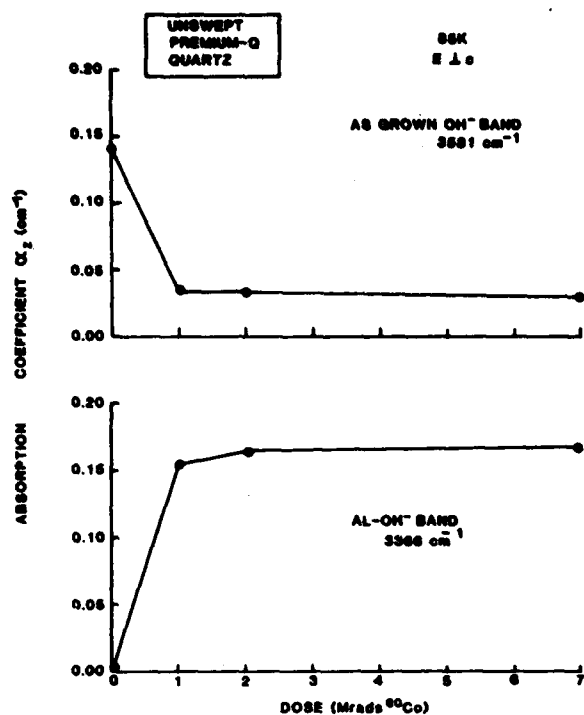
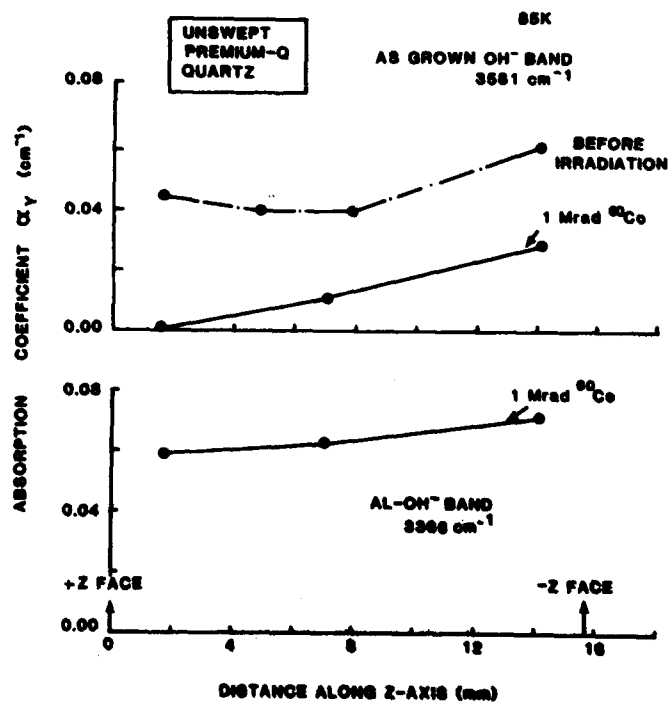


Figure 2. Hydrogen related defect band absorptions in unswept Premium-Q quartz, measured parallel to the crystal growth axis, as a function of irradiation dose.

Figure 3. Hydrogen related defect band absorptions in unswept Premium-Q quartz, measured perpendicular to the crystal growth axis, before and after 1 Mrad ⁶⁰Co irradiation.



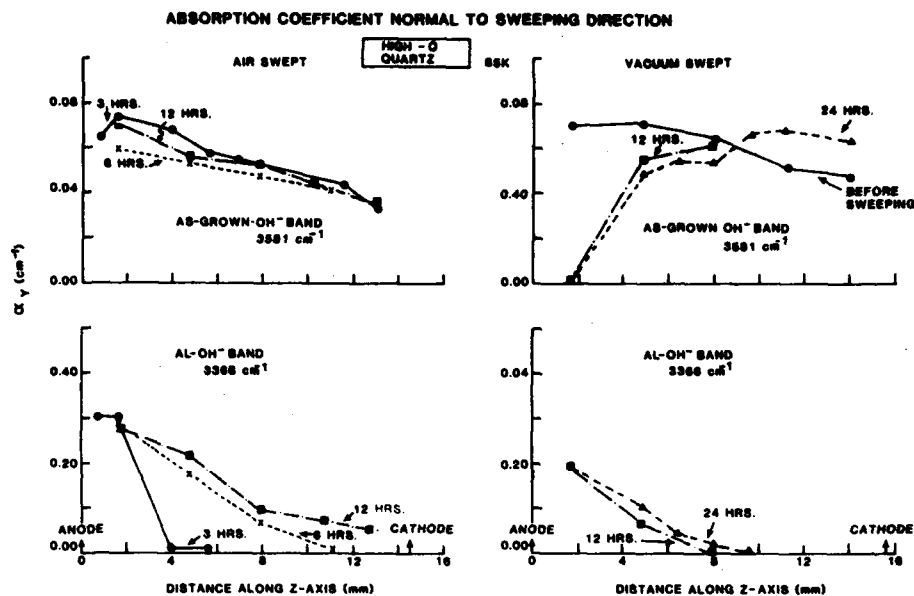


Figure 4. Hydrogen related defect band absorptions in High-Q quartz as a function of position between the anode and cathode after air or vacuum sweeping for the indicated time periods.

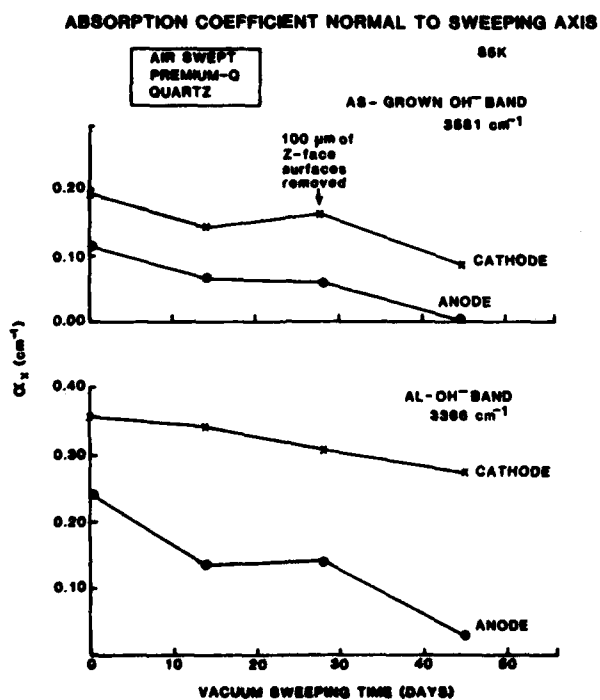


Figure 5. Hydrogen related defect band absorptions for air swept Premium-Q quartz measured at the anode and cathode after vacuum sweeping for the indicated time periods.

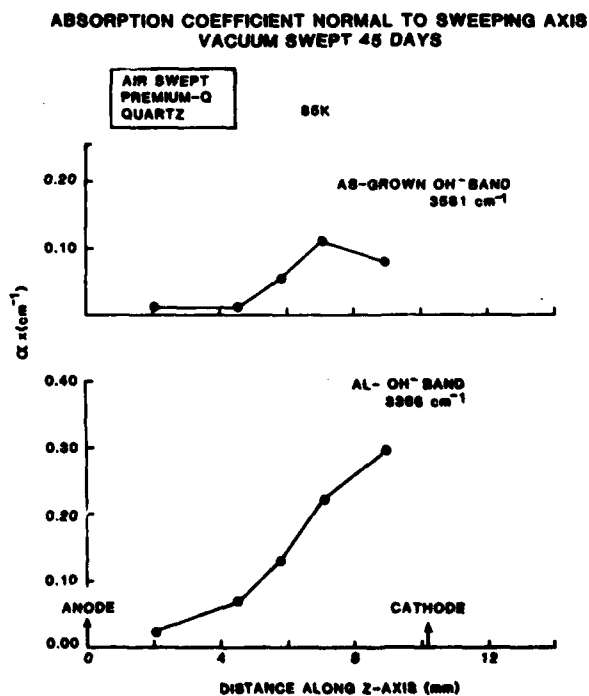


Figure 6. Hydrogen related defect band absorptions as a function of position between the anode and cathode for an air swept Premium-Q quartz sample after vacuum sweeping for 45 days.

AD P 002472

EFFECT OF SEED TREATMENT ON QUARTZ DISLOCATIONS

Alton F. Armington, John J. Larkin, John J. O'Connor, J. Emery Cormier and Jane A. Horrigan
 Rome Air Development Center
 Solid State Sciences Division
 Hanscom Air Force Base, MA 01731

Abstract

This work is a continuation of the work on the preparation of high purity, low dislocation quartz being performed at RADC. Several methods have been investigated for the preparation of high purity materials. The effect of seed pretreatment is also discussed.

Introduction

The major portion of the Air Force quartz growth program is the development of the production of high purity, low dislocation quartz based on the assumption that this material will have low frequency offset and low etch tunnel density. Previously we have reported on purity studies performed in our laboratory. In this presentation we will present additional work on purity and some initial studies on the effects of seeds and seed pretreatment.

The system used for this work has been described previously¹ and consists of ten autoclaves varying in diameter from one to four inches which are under computer control. Most of the studies reported here were in the three inch (by three feet) units. The mineralizer was generally hydroxide and the nutrient cultured quartz with an aluminum concentration of about 15 parts per million atomic. Some runs were performed in silver liners and some in unlined autoclaves.

Results

A series of runs were performed to determine the effect of using a silver liner in the autoclave without a top. The hope was that the purity would not be seriously affected in this case, based on the assumption that the material deposited on the autoclave wall would not be transported to the seeds as rapidly as without the liner. Runs were performed using an unlined autoclave, a lined autoclave, and a silver liner without the top which is usually welded on. The results of chemical analysis by atomic absorption are shown in Figure 1. It is apparent from the results that using the same conditions there is little to be gained using a partial liner, since the unlined results are as good as those from the partial liner. The higher results in the partial liner are probably due to a higher growth rate in this case. We are at present, growing with an unwelded top to see if this will reduce the impurities in a growing crystal.

A series of runs was also performed with and without the use of the lithium additive to the mineralizer. The results in this case are shown in Figure 2. Both lined and unlined autoclaves were tested, as well as a carbonate mineralizer. The aluminum content is somewhat higher without the lithium addition but not significantly higher. It is to be expected that the difference between crystals with and without lithium addition would be more pronounced at higher growth rates².

All of these studies seems to indicate that growth rate is the most important parameter to consider in all cases where low purity is desired, at least down to the parts per million range. In the case of

ultrapure nutrients³, where even lower purity is evidenced, both the nutrient and the liner material are also considerations. Figure 3 shows the effect of growth rate on the aluminum content of several samples. The upper curve shows aluminum content without the addition of lithium while the lower curve is for runs with lithium addition. It can be seen that the difference is more pronounced at higher growth rates. In this set of experiments we found the aluminum (not plotted) to be higher in the carbonate runs but these were only performed at growth rates above 30 mils/day. While the lithium content increases also with growth rate, the values are not as well defined as for aluminum. We could find no correlation with sodium either in terms of growth rate or lithium addition. We have also found a relationship between the growth rate and point defects in the crystal⁴. Also about half of the points shown were crystals grown using seeds from the +x region of quartz crystals again indicating that growth rate is more significant than type of seed.

The condition of the autoclave wall may have an effect on the purity of the crystal. Figure 4 shows some analysis on the coating material on the wall. It can be seen that aluminum is a significant impurity, probably precipitating early from the hot mineralizer after it enters the growth zone in the convective current above the baffle. Also shown in this figure are the results before and after an extensive cleaning in the autoclave. This does indicate that some aluminum in the crystal may be picked up from the wall during growth.

The purity of the nutrient also affects the purity of the crystal as has been shown previously^{1,3}. However, there are at least three ways aluminum can be limited to less than a part per million without the use of a liner. Even better purity can be produced with a liner but this is both expensive and complicated in a commercial size autoclave. The results of some of these techniques are shown in Figure 5. The first result, devitrification, was produced by the recrystallization of high purity glass (in a silver liner), followed by growth in an unlined vessel. The second result, recrystallization, was the result of the second recrystallization of quartz from a high purity (3.3 ppm Al) cultured quartz in an unlined vessel. The third result was produced by using only nutrient cut from the Z growth area of cultured samples. Analysis has shown that this is 5-10 times lower in aluminum than the other areas of the crystal. Similar analyses have recently been reported³ for the S and Z regions.

Seed Studies

Our effort in improving seed quality has concentrated on two areas, seed pretreatment and the use of seeds from the X growth area. Using Z face seeds both from the Z growth and the X growth area in the same autoclave, Figure 6, we obtain better strain control with the X area seed. The growth rate of the seeds was about 20 mils per day, in a silver liner. The mineralizer in this case was sodium hydroxide with no lithium added. The aluminum content of the crystal grown from the X seed was 5 ppm, with the Z area seed having a somewhat higher aluminum content. The use of this figure is somewhat optimistic since we do not

AD-A136 673

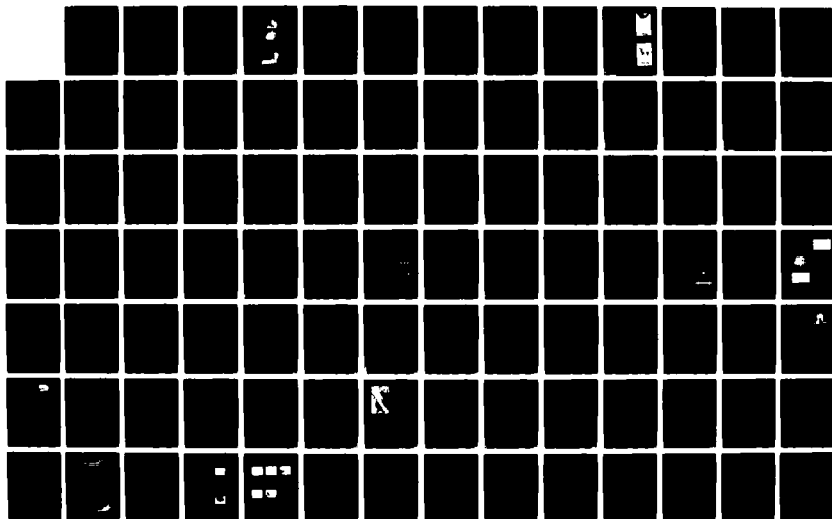
PROCEEDINGS OF THE ANNUAL SYMPOSIUM ON FREQUENCY
CONTROL (37TH) 1-3 JUNE 1..(U) ARMY ELECTRONICS
RESEARCH AND DEVELOPMENT COMMAND FORT MONMOUTH 1983

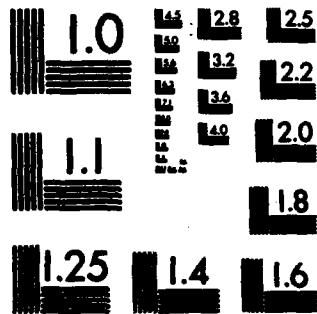
F/G 9/1

3/4

NL

UNCLASSIFIED





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

always get strain-free crystals using the X growth seeds. However, they are always more strain-free than the Z area seeds used in the same run. We are not certain why some seeds do produce strain, but suspect it is due to the growth rate or seed treatment. In one case (Figure 7), bubbles were produced at the seed interface with dislocation (strain) lines propagating from them.

We have completed a series of runs on the effect of the heating-up cycle on the seed condition prior to growth. In this case we heated the seeds up for twenty-four hours to operating temperature, cooled down and removed the seed. We have done this for untreated seeds, HF-etched seeds, ammonium bifluoride seeds and chemically etched seeds. There does not appear to be any difference in the results whether or not a pretreatment is employed.

Summary

The growth rate appears to be the most important factor in the preparation of low aluminum quartz crystals, at least down to the parts per million range. For purer crystals, liners and select nutrients are required. Studies using seeds from the X-growth area show promise but more studies must be done to produce reliable results. Pretreatment of seeds does not ap-

pear to offer any advantages for the growth of dislocation-free crystals.

Acknowledgements

The authors are indebted to Mackie Harris for his assistance in the growth of the crystals. Characterization was also performed by L. Halliburton, Oklahoma State and by A. Fluesmeier, Eagle Picher Inc. The authors are also indebted to Joseph Balascio and Baldwin Sawyer for their helpful discussions and assistance.

References

1. A.F. Armington, J.J. Larkin, J.J. O'Connor and J. A. Horrigan, Proceedings of the 35th Annual Frequency Control Symposium, U.S. Army Electronics Command, Fort Monmouth, N.J., p297, Electronics Industry Association, 2001 Eye Street, N.W., Washington, D.C. 20006.
2. N.C. Lias, et al, J. Crystal Growth, 18, 1 (1973).
3. S.P. Doherty, S.E. Morris, D.C. Andrews and D.F. Croxall, to be published in Radiation Effects.
4. J.J. Martin and A.F. Armington, to be published in J. Crystal Growth.

	Al*	Fe	Li	Na	K	GROWTH RATE MILS/DAY
CLOSED SILVER LINER	3.8	<0.5	1.0	<0.5	<0.5	16
PARTIAL SILVER LINER	13	<0.5	4.4	2.7	<0.5	30
UNLINED (X-35)	6.9	2.7	—	1.7	0.7	19

*IN PARTS PER MILLION ATOMIC

Figure 1. Analytical Results of Full and Partial Liners.

	Al*	Fe	Li	Na	K
UNLINED RUN					
WITH LITHIUM	6.9	2.7	—	1.7	0.7
NO LITHIUM	6.5	3.4	1.5	2.1	<0.5
PARTIAL LINER					
WITH LITHIUM	13	<0.5	4.4	2.7	<0.5
NO LITHIUM	17	0.5	4.7	3.3	<0.5
LINED RUN					
WITH LITHIUM	3.8	<0.5	1.0	<0.5	<0.5
NO LITHIUM	5.3	<0.5	1.5	0.9	<0.5
LINED RUN (CARBONATE)					
WITH LITHIUM	35	1.2	17	11	0.5
NO LITHIUM	43	0.8	14	11	1.4

*IN PARTS PER MILLION ATOMIC

Figure 2. Analytical Results of Runs with and without Lithium Additive.

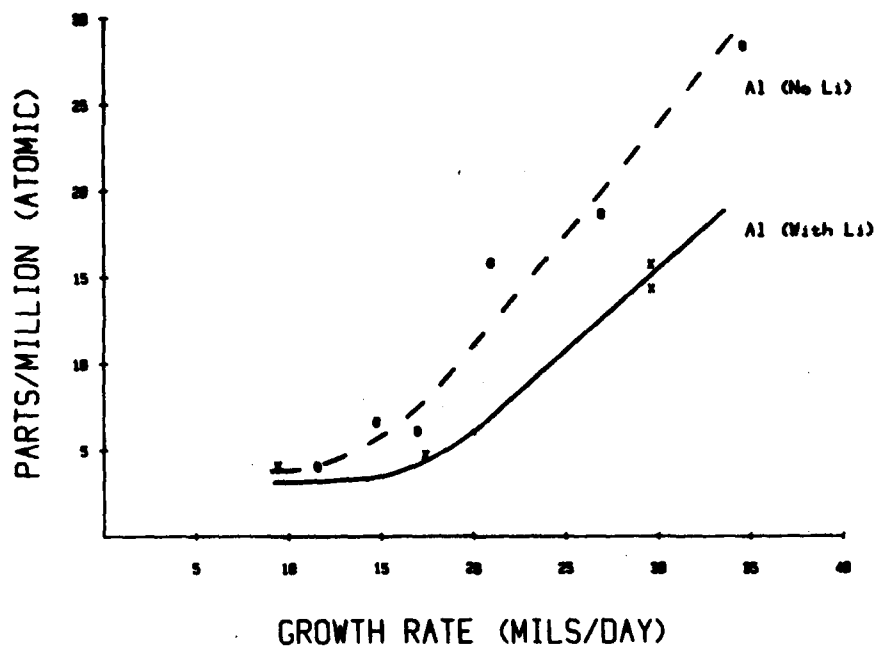


Figure 3. Effect of Growth Rate on Impurity Concentration in Grown Crystals.

ANALYSIS OF WALL DEPOSITS*

AL	Fe	Li	Na	K
920	22%	127	6.6%	32
780				
2590				

ALUMINUM IN CRYSTALS

BEFORE CLEANING - 5-8

AFTER CLEANING - 3.7-6.5

*PARTS PER MILLION ATOMIC UNLESS SPECIFIED

Figure 4. Results of Cleaning an Unlined Autoclave.

ALUMINUM* IN CRYSTALS FROM VARIOUS NUTRIENTS

- 1 - DEVITRIFICATION - 0.6
- 2 - RECRYSTALLIZATION - <1
- 3 - Z GROWTH - 0.6

*PARTS/MILLION ATOMIC

Figure 5. Effect of Nutrient on Purity.

X
SEED



Z
SEED



Figure 6. Topographs of Crystals Grown on Z-area and X-area seeds.



Figure 7. Topograph of X-area seed showing Interface Bubbles.

AD P 002473

→ THERMOLUMINESCENCE FROM DIFFERENT GROWTH SECTORS IN SYNTHETIC QUARTZ CRYSTALS

S. Katz, A. Halperin and M. Ronen

Racah Institute of Physics
The Hebrew University of Jerusalem
Israel

Summary

Many electronic devices containing quartz crystals operate in an environment exposed to radiation, and it is important therefore to investigate the effects of radiation on quartz crystals. In the present work we show that the technique of thermoluminescence TL can serve as a good tool for such an investigation. We report on TL measured in the temperature range of 15 - 300 K. The measurements have been carried out separately on the various growth sectors (namely +X, -X, +Z, -Z) of Y plates cut from a Y grown bar of unswept Electronic Grade Synthetic Quartz. The well resolved glow peaks at 72 K, 80 K, 87 K and 104 K were observed for the first time in the present work. We suggest that the glow peaks appearing only in +X growth sector samples (namely 135 K, 162 K, 178 K) after pre-irradiations at room temperature (RT) by x-rays (white radiation, 50 KV, 17 ma), might be related to $[Al_{Si}]^0$ centers. The glow peaks above 104 K in the +Z sector samples are completely different from those observed in +X sector samples. The glow peaks in this temperature range progressively fade out with increasing pre-doses at room temperature. From the measurements and the results described in this paper it can be seen that one or more glow peaks correspond to a certain type of defect present in various growth sectors and that their TL intensities are proportional to defect concentration present in quartz.

Introduction

Alpha-quartz is used in a variety of electronic devices, e.g. high precision oscillators, filters, etc., where high frequency stability is of prime importance. It is known that acoustic losses and frequency instability in quartz crystal oscillators are affected by radiation induced charge trapping at point defects. It is therefore important to understand thoroughly the origin and the nature of these defects. Many electronic devices containing quartz crystals operate in an environment exposed to radiation, and it is important therefore to investigate the effects of radiation on quartz crystals. In the present work we show that the technique of thermoluminescence TL can serve as a good tool for characterization of certain defects present in quartz. Several investigators¹⁻³ have measured the TL of quartz above room temperature, in a wide variety of natural and/or synthetic crystals. Their results were found to be sample-dependent. Low temperature TL has received less attention and has been studied among others by Madlin⁴ and Malik et al.⁵ after irradiations at 95 K. In the present work we report on TL measured in the temperature range of 15-300 K. The measurements have been carried out separately on the various growth sectors in the quartz crystals, which helped in the characterization of the defects and impurities involved.

Experimental

Y-cut plates were cut from a single-crystal Y grown bar of unswept Electronic Grade Synthetic Quartz. Samples from the different growth sectors of Y plates (namely -X, +X, -Z, +Z) were cut in the form of plates having dimensions of 10x6x3 mm³. The measurements were carried out separately on the different growth sectors. The samples were mounted in a Displex cryostat produced by Air Products and Chemicals Inc., designed for measurements in the range 10-350 K. The crystal was fitted into a rectangular hole in the cold finger with a copper cover screwed to the holder using indium to improve the thermal contact between the crystal and its holder. For the TL measurements the samples were irradiated at low temperature for a period of 100 minutes, through a thin (0.3 mm) aluminum window using white radiation from a tungsten target at 50 KV and 15 ma. The total x-ray dose in each irradiation was close to 2 Mrad. After irradiation the cryostat was placed in front of an optical system comprising a 0.25 m Jarrell-Ash monochromator and a cooled E.M.I. 9658R photomultiplier tube operated at 1160 V. Its output was fed into a 610 Keithley electrometer and recorded on a chart recorder. This arrangement enabled us to take fast scans (15-20 seconds) of the TL spectra over the range 250-700 nm. A heating rate of 2.5 K/min was obtained by means of a Malabs driving motor potentiometer with its output connected to a precision temperature controller. Temperatures were measured by a chromel vs gold - 0.07% iron thermocouple pair connected to a digital multimeter with its reference at 0°C. Glow curves were plotted from the spectra recorded at intervals of about 2 K. In order to get a better insight into the origin of the TL changes in the various growth sectors, the crystals were also X - irradiated (50 KV, 17 ma) at room temperature for different periods of time. This radiation resulted in a smoky coloration mainly in +X sector samples and left the Z sector samples uncolored. Optical absorption measurements after room temperature pre-irradiations were carried out on a Carry 14 spectrophotometer at 77 K on +X and +Z growth sectors.

Results

Fig. 1 shows glow curves obtained from the +X sector before room temperature irradiation (dashed curve) and after a pre-irradiation at room temperature for 2.5 h (solid curve). The glow curve for the non-irradiated crystal shows glow peaks located at 72 K, 80 K, 87 K, 104 K and 127 K, and a complex of glow peaks at higher temperatures up to about 220 K. After pre-irradiation at room temperature for 2.5 h (solid curve in Figure 1) the glow curve for +X growth sector shows some decrease in intensity of the peaks at 72 K, 80 K and 87 K and 104 K, the disappearance of the glow peak at 127 K and of the complex of glow bands in the temperature range of 137 K - 165 K and an increase in intensity of the complex of glow peaks belonging to

the range 180 K - 230 K. Also shown is the appearance of 3 new strong peaks at 135 K, 162 K, 178 K. These 3 peaks increase in intensity with the time of pre-irradiations at room temperature as can be seen in Fig. 2, which depicts the dependence of the maximal glow peak intensity upon the total time of irradiations, at room temperature (RT) for the different glow peaks present in the +X sector sample. The strongest glow peaks in the +X region, after prolonged pre-irradiations at room temperature, were at 162 K and 178 K. The 72 K and 80 K glow peaks showed an emission maximum near 430 nm and that of the shoulder at 87 K was at 470 nm. Other glow peaks showed in general a maximum emission near 380 nm. In parallel it was found by us that curves of the optical absorption spectrum recorded in the wavelength range 250-750 nm show after every pre-irradiation at RT, intensity increase, with the maximal TL intensities of the glow peaks at 135 K, 162 K and 178 K. Our measurements of visible absorption spectra show a maximum near 2.5 eV. This kind of spectra was reported by Koumvakalis⁶ for the aluminum - associated hole center $[Al_{eq}]^0$. Fig. 3 shows marked differences in the glow curves for the sample cut from +Z sector. The glow curve for the non-irradiated crystal at room temperature (dashed curve) shows strong peaks at 112 K, 145 K, 162 K and 200 K. The glow curve traced after pre-irradiation at room temperature for 7 h (solid curve) shows that the intensity of the glow peaks appearing at temperatures greater than 104 K are greatly suppressed. The glow peak at 104 K shows an emission maximum at 470 nm and its behaviour with the time of pre-irradiation at room temperature is similar to that of the glow peaks at 72 K, 80 K and 87 K. The dependence of the maximal glow peak intensities upon the total time of irradiations at RT for the various glow peaks encountered in the TL curves characteristic of +Z sectors is shown in Fig. 4. For the peaks at 72 K, 80 K, 87 K, 104 K there is a reduction in intensity after 2.5 h of irradiation at RT followed by a slight increase in intensity after total irradiation times of 4.5 h, 7 h, 10 h. As shown in Fig. 4 other glow peaks are considerably reduced after prolonged pre-irradiation at RT. It was found by us from the corresponding optical curves at 77 K that there is no measured absorption at all for the sample cut from a +Z region after the various irradiations at RT.

In order to analyse some glow peaks we applied the initial rise method to both +X and +Z crystals. This allowed us to obtain the activation energies for glow peaks where the initial rise was clearly resolved. These results together with the emission wavelengths are summarized in Table I.

TABLE I. Glow peak temperature, activation energies and emission wavelengths for the different glow peaks found in the quartz samples

Glow Peak Temperature (K)	Activation Energy (eV)	Emission Wavelength (nm)
72	0.16	430
135	0.213	380
162	0.35	380

Discussion and Conclusions

To obtain a thorough understanding of radiation effects in quartz crystals it is necessary to consider data obtained from measurements below room temperature.

To the best of our knowledge the well resolved peaks at 72 K, 80 K, 87 K and 104 K were observed for the first time in the present work. These peaks appear in both +X and +Z sectors, which indicates that

they are associated with defect centers present in both sectors at these temperatures. The glow peaks at 72 K and 78 K emit at 430 nm and those at 87 K and 104 K at 470 nm. So there are at least two types of centers associated with these glow peaks. It has been shown by Markes and Halliburton⁸ that irradiation at low temperatures produces hydrogen atoms H^0 . The atoms are detrapped and disappear if the quartz crystal is heated much above 100 K. We can assume tentatively that the previously mentioned glow peaks are due to the presence of hydrogen atoms H^0 .

Halliburton et al.⁹ have shown that the aluminum in as received unswept quartz is primarily in the form of $Al-M^+$ centers (where M^+ represents an alkali ion). In addition they have indicated that irradiation at temperatures above 200 K converts the $Al-M^+$ centers into $Al-OH^-$ centers and aluminum hole trapped centers $[Al_{eq}]^0$. They have also shown that $Al-OH^-$ centers are converted to aluminum hole trapped centers by irradiation at 77 K. In our case it seems that the glow peak at 127 K and the complex of glow peaks between 137 K and 165 K present in +X growth sector samples (see Fig. 1. dashed curve) can be correlated with $Al-M^+$ centers because in this sector both glow peaks completely disappear after room temperature irradiations. We suggest that the glow peaks appearing in +X growth sector samples (namely 135 K, 162 K and 178 K) might be related to $[Al_{eq}]^0$ centers. Their TL intensities progressively increase together with the intensities of the optical absorption spectra measured after various pre-irradiations at room temperature. Therefore the TL for these 3 glow peaks results from recombination of electrons and holes at recombination centers connected with defects of aluminum hole trapped centers. This suggestion is corroborated by the fact that the 3 glow peaks at 135 K, 162 K and 178 K do not appear in +Z crystals and are characteristic of +X sector crystals. The glow peaks above 104 K in the +Z sector samples are completely different from those observed in +X sector samples. All glow peaks in this temperature range progressively fade out with increasing predoses at room temperature. These glow peaks may be related to dissociation at $Al-M^+$ centers which may be present in +Z sector samples. A final conclusion on this point can not be reached unless a sweeping experiment previous to room temperature irradiations is carried out. In that case marked differences should appear in the glow curve for +Z sector samples measured after completion of the sweeping process.

The TL technique is a nondestructive method for measuring impurity concentrations and various radiation doses. From the measurements and the results described in this paper it can be shown that one or more glow peaks correspond to a certain type of defect present in various growth sectors and that their TL intensities are proportional to defect concentration present in quartz.

The above results were obtained on electronic grade quartz crystals. Additional experiments are in progress and measurements are planned on premium high-Q grade crystals. It is hoped that the investigation will enable us to continue to clarify the nature of the various defects and impurities incorporated in these growth sectors.

References

1. R. Yokota, Phys. Rev., 91 (1958) 1013.
2. D.W. Mc Morris, J. Geophys. Res., 76 (1971) 7875.
3. S.A. Durrani et al., J. Phys. D, 10 (1977) 1351.
4. W.L. Madlin, J. Chem. Phys., 38 (1963) 1132.
5. D.M. Malik et al., J. Appl. Phys., 52 (5) (1981) 3600.
6. N. Koumvakalis, J. Appl. Phys., 51 (10) (1980) 5528.
7. J.C. King and H.H. Sander, IEEE Trans. on Nuc. Sci. NS-19 (1972) 23.
8. M.E. Markes and L.E. Halliburton, J. Appl. Phys., 50 (1979) 8172.
9. L.E. Halliburton et al., J. Appl. Phys., 52 (5) (1981) 3565.

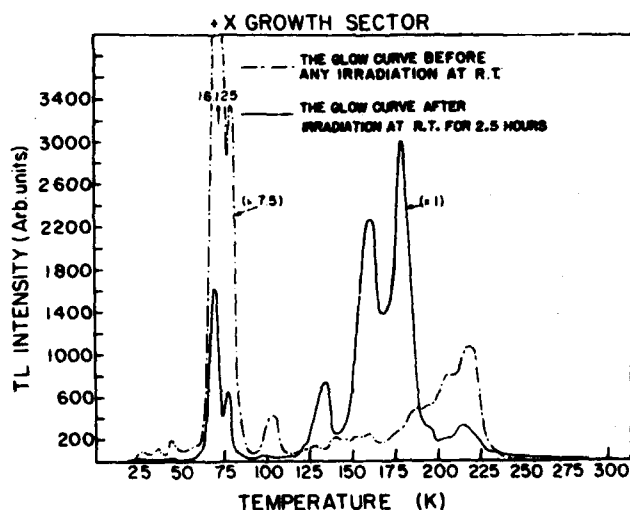


Figure 1. Glow curve for the +X growth sector sample irradiated at low temperature only is given by the dashed curve. The solid curve shows a glow curve for the same sample except that the low temperature irradiation is preceded by a 2.5 hours irradiation at room temperature. Off scale value of the 72 K peak intensity is indicated.

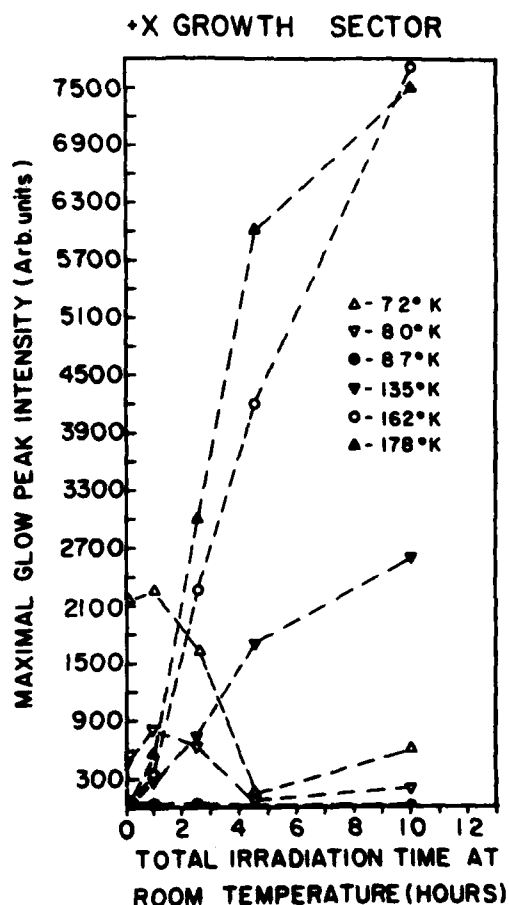


Figure 2. The dependence of the maximal glow peak intensity upon the total irradiation time at room temperature for the different glow peaks present in the +X growth sector sample.

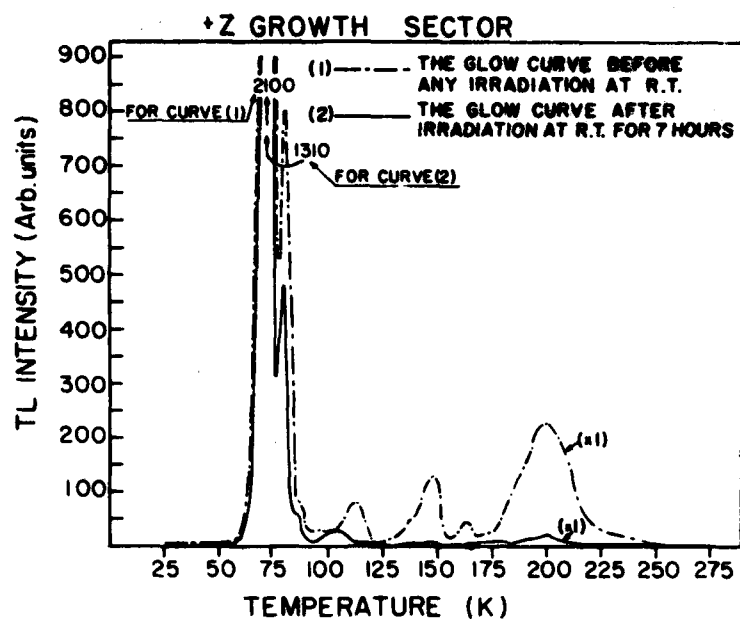


Figure 3. Glow curve for the +Z growth sector sample irradiated at low temperature only is given by the dashed curve (1). The solid curve (2) shows a glow curve for the same sample except that the low temperature irradiation is preceded by a 7 hours irradiation at room temperature. Off scale values of the 72 K peak intensities for curve (1) and curve (2) are indicated.

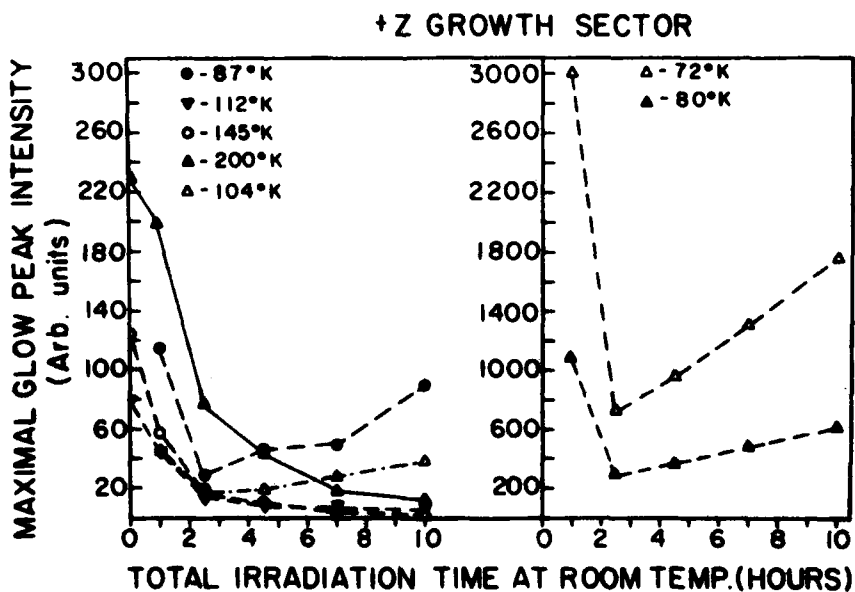


Figure 4. The dependence of the maximal glow peak intensity upon the total irradiation time at room temperature for the different glow peaks present in the +Z growth sector sample.

CHARACTERIZATION OF QUARTZ CRYSTALS BY CATHODOLUMINESCENCE

S. Katz and A. Halperin

Racah Institute of Physics
The Hebrew University of Jerusalem, Israel

and

M. Schieber
School of Applied Science and Technology,
Hebrew University, Jerusalem, Israel.

Summary

Cathodoluminescence (CL) in a scanning electron microscope (SEM) is used in the present work for investigating various growth sectors found in synthetic quartz crystals. This is a relatively simple, rapid and non-destructive method which gives a very high resolution. Our results using CL correlate well with those obtained by other methods, e.g. x-ray topography and light scattering tomography. The advantages of CL technique over other methods are described in detail and discussed.

Introduction

The various crystal defects and impurities found in synthetic quartz crystal resonators influence considerably their frequency stability and quality factor Q . It is therefore important to understand the origin and the nature of such defects and to find simple experimental methods for their investigation. So far these defects have been investigated by X-ray topography¹⁻³ and by light scattering tomography⁴⁻⁵.

In the present work we describe a relatively simple and rapid method for investigating various growth sectors found in synthetic quartz crystals by use of cathodoluminescence (CL) in a SEM⁶. This is a non-destructive method which gives a very high resolution and it is suitable for the investigation of layers of the crystal close to the surface. Our results using cathodoluminescence correlate well with those obtained by other methods, e.g. X-ray topography and light scattering tomography. The advantages of cathodoluminescence technique over other methods will be described in detail and discussed.

Experimental

The investigation was carried out on several quartz plates cut from a single crystal Y grown bar of unswept Electronic Grade Quartz perpendicular to the Y axis. These crystals comprised the different growth sectors of Y plates, namely +X, -X, +Z, -Z. The samples were covered by a thin layer of carbon (80 Å thickness) for avoiding specimen charging during the electron beam scanning of the surface. Cathodoluminescence studies were made in a Cambridge Stereoscan S4 SEM. The crystals were scanned by an electron beam of 30 KeV (penetration depth of about 7 μm). The diameter of the electron beam was about 1 μm and the tilting angle of the sample was near 45°. The emitted radiation was collected by a quartz lens which focused it onto the photocathode of an RMI 6235 S photomultiplier.

In order to gain better insight into the origin of the luminescence in the various growth sectors some of the crystal plates were irradiated by x-rays (white radiation, 50 KV, 17 ma) at room temperature for a period of 5 hours.

Results and Discussion

Fig. 1 shows the C.L. panchromatic micrograph of the various growth sectors found in a quartz Y plate. This is a composite micrograph formed from 6 pictures of different integrated cathodoluminescence patterns taken from a quartz sample before irradiation at room temperature. Strong and well defined luminescence patterns were obtained from +X growth zone including growth striations decorated by impurities. The luminescence of -X growth zone was in general similar to that of the +X zone but sometimes lower in intensity. The +Z and -Z growth sectors gave only faint luminescence, if at all. In the later growth sections, growth striations are seen as continuations of the corresponding ones in the +X growth sector. It has been shown that the Z growth sectors have the lowest impurity content and this is compatible with the fade out of luminescence in these sectors. Similar growth striations are often observed by X-ray topography¹ and by light scattering tomography⁵ in the growth sectors of the Y-cut plates. They are in general caused by variations of lattice parameter associated with changes in the impurity segregations during the variation of the growth rate. Growth sector boundaries are also observed. It has been reported¹ that the presence of growth striations is associated with a higher equivalent series resistance of the resonator and that the different deleterious defects change the nature of the resonator vibrations.

The cathodoluminescence method used in the present investigation has some advantages over other methods: a) Unlike other methods, for example X-ray topography this is a non-destructive method and only the surface of the sample is bombarded for very short periods of time (a few seconds) by low energy electrons. Therefore the damage to the crystal is much less significant. b) The samples are used as grown crystals and it is not necessary to slice them to a thickness of 1 mm. In the X-ray topography one must consider the dependence of the topographic contrast of X-rays upon the thickness of the samples and the structure of defects in them. The specimens must always be sliced because X-ray beams are not transmitted by very thick samples. c) The method gives a very high resolution. It is worth mentioning that the easiest and most convenient way to achieve an optimum spatial resolution is to use electron beam excitation in a SEM. Focussing and positioning of the beam are much easier in a scanning electron microscope than in systems based on optical excitations (as in the case of light scattering tomography). The recording of micrographs is also much easier in such an instrument. d) It is very suitable for the investigation of layers of the crystal close to the surface.

Fig. 2 shows the CL panchromatic micrograph of the various growth sectors of the same crystal as in Fig. 1 after irradiation at room temperature by X-ray white radiation for a period of 5 hours. This radiation caused a smoky coloration in +X and -X growth sectors

and left the Z sectors uncolored. After this irradiation the cathodoluminescence patterns obtained from +X and -X growth sectors under the same experimental conditions as for Fig. 1 (e.g. the same accelerating voltage, beam current, the same voltage on the photomultiplier etc.) are much weaker than those observed before the irradiation. It has been shown by Halliburton et al.⁷ that the aluminum in as received unswept quartz is primarily in the form of Al-M⁺ centers (M⁺ - represents an alkali ion). After room temperature irradiation these centers are converted into Al-OH⁻ and aluminum hole trapped centers. So it may be suggested that the strong luminescence in the +X and -X growth sectors of the sample before any irradiation at room temperature arises from recombination of electrons with holes trapped adjacent to Al-M⁺ centers. This is because the CL fades out in the both +X and -X growth sectors after room temperature irradiation and so it may be related to dissociation of Al-M⁺ centers present eventually in these growth sectors. A final conclusion can be reached only after carrying-out a sweeping experiment on the sample before any room temperature irradiation and thorough investigations of C.L. spectra from low to room temperatures carried out separately on the various growth sectors in the quartz crystals.

Acknowledgement

We acknowledge with thanks the technical help of Engineer Marcel Ronen during the present work.

References

1. K.L. Bye et al., *Journal of Materials Science*, **14** (1979) 800.
2. J. Yoshimura and K. Kohra, *Journal of Crystal Growth*, **33** (1976) 311.
3. D.Y. Parpia, *Journal of Materials Science*, **12** (1977) 844.
4. K. Moriya and T. Ogawa, *Phil. Mag.*, **A44** (1981) 1085.
5. K. Moriya and T. Ogawa, *Journal of Crystal Growth*, **44** (1978) 53.
6. S. Katz, A. Halperin and M. Schieber, *Ultramicroscopy*, **8** (1982) 452.
7. L.E. Halliburton et al., *J. Appl. Phys.*, **52**(5) (1981) 3565.



Figure 1. The panchromatic CL micrograph of the various growth sectors present in a quartz Y plate. This is a composite micrograph of different integrated CL patterns taken before irradiation at room temperature.



Figure 2. The panchromatic CL micrograph of the various growth sectors of the same crystal as in Figure 1 after irradiation at room temperature for 5 hours by X-ray white radiation.

AD P 002475

IMPROVEMENTS OF LASER INTERFEROMETRIC MEASUREMENT SYSTEM OF VIBRATION DISPLACEMENTS

Takeshiko ABE, Masahiko OKIZAKI, and Tetsuo TSUBUKI

Yokohama National University
Tokiwada, Hodogaya-ku
Yokohama, Japan 240

SUMMARY

Improvements have been made on the laser interferometric measurement system of vibration displacements of quartz crystal resonators. In this paper, the following modifications have been made: the DC phase control and the small-deviation low-frequency phase modulation are introduced on a laser beam. The amplitude of the vibration displacement is obtained as a digital data. A measurement and processing of data are made by the computer system. The improved system can be used to the measurement of vibration displacement down to five angstroms within a few percent repeatability in the frequency range from 1 MHz to 10 MHz. Using this system, the detailed vibration displacement distributions of a plano convex SC-cut quartz crystal resonators have been measured.

Introduction

We have previously developed the laser interferometric measurement system of vibration displacements introducing a low-frequency phase modulation into one of the laser beams and proved it to be a powerful means for measuring displacements of plano-convex AT-cut quartz crystal resonators. In this measurement system, the peak-to-peak values of the vibration frequency component and the modulation frequency component in the output of the photodetector are measured, and the vibration displacement is calculated from the ratio of these values.

In this paper, the following modifications have been made in order to apply the digital measurement technique to the measurement system mentioned above. In the improved system, DC phase control and the small-deviation low-frequency modulation are adapted by the PZT-driven mirror phase modulator, so that the two output signals of the photo detector become the pure sinusoidal waves. The voltage ratio of the two signals is measured and converted to the digital signal by a digital level meter. The corresponding amplitude of the vibration displacement is calculated from the digital signal by a microcomputer system.

The features of the improved measurement system are:

- (1) The measurement of data and the calculation of displacements are made automatically, therefore time needed for the measurement can be reduced and the high accuracy can be obtained.
- (2) Data processing -- determination of

the optimally approximated displacement distribution functions and the drawing a graph of the displacement distributions can be easily made.

In order to evaluate the performance of the measurement system, measurements have been made on a sample resonator whose displacement amplitude has been already known well. By using the new measurement system, the distributions of the three dimensional vibration displacement components of the fundamental, the third overtone, and the fifth overtone plano-convex SC-cut quartz crystal resonators have been measured.

Principle of Measurement

Fig. 1 shows the fundamental arrangement of the laser interferometric measurement system in the case of the measurement of vibration displacements perpendicular to the surface of the resonator under study. The laser beam, emitted from a laser unit, separated into two beams by a half mirror. One of these beams is projected onto the mirror, and the other beam is focused onto one point on the resonator surface. The reflected two beams from the mirror and the resonator are superposed by the half mirror and are collected to the avalanche photo diode by a condenser lens. In the presence of a vibration displacement $U \cos \omega t$ to be measured, the beam A undergoes a phase modulation of $(4\pi/\lambda) U \cos \omega t$, where λ is the wavelength of the laser light. When the DC phase control ϕ_{DC} and the low-frequency phase modulation $\phi \cos \omega t$ are introduced into the laser beam B, the beam B undergoes a phase modulation of $\phi_{DC} + \phi \cos \omega t$. Hence, the light intensity incident upon the photo detector is expressed as the following equation.

$$I = E_a^2 + E_b^2 + 2E_a E_b \cos(\phi_{DC} + \phi \cos \omega t + \frac{4\pi}{\lambda} U \cos \omega t) \quad (1)$$

where E_a and E_b are the amplitude of the beam A and B, respectively. If DC phase shift of beam B is adjusted to $\pi/2$, the light intensity of Eqn. (1) becomes as follows.

$$I = E_a^2 + E_b^2 - 2E_a E_b \sin(\phi \cos \omega t + \frac{4\pi}{\lambda} U \cos \omega t)$$

Since the range of displacement U of our concern is not more than several hundred angstroms, and the deviation of the low-frequency phase modulation is controlled to be sufficiently small, the approximations $U \ll \lambda$, and $\phi \ll 1$ can be applied. Then, the Equa.(2) becomes as follows.

$$I = E_a^2 + E_b^2 - 2E_a E_b \phi \cos \pi t - \frac{8\pi}{\lambda} E_a E_b \cdot U \cos \pi t \quad (3)$$

The output of the photo detector is proportional to the light intensity incident upon the detector. Therefore, the output of the photo detector contains the two components; the low-frequency component I_0 , the 3rd term of Equa. (3) which depends only on the intensity of the two beams, and the vibration frequency component I_1 , the 4th term of Equa. (3), which depends not only on the intensity of the beams but also on the amplitude of the vibration displacement. By using the ratio of the amplitudes of these two components, the amplitude of vibration displacement is expressed as follows:

$$U = \frac{\lambda \phi |I_1|}{4\pi |I_0|} \quad (4)$$

Therefore, if the deviation of low frequency phase modulation is estimated beforehand, the amplitude of vibration displacement can be calculated from the ratio of I_0 and I_1 . The estimation of ϕ , i.e. the estimation of the absolute magnitude of the displacement can be made by the method proposed in the previous paper.¹

In the case of the measurement of vibration displacements parallel to the resonator surface, the measurement can be made simply by changing the arrangement to the Mach-zehnder interferometer. In this case, vibration displacement U parallel to the resonator surface is expressed as follows:

$$U = \frac{\sqrt{2} \lambda \phi |I_1|}{4\pi |I_0|} \quad (5)$$

Measurement system

Fig.2 shows the laser interferometric measurement system. Measurement system is made up of the three parts: the laser interferometer, the optical detection system, and the microcomputer and minicomputer system. The optical arrangement depicted in Fig.2 is Michelson interferometer for the measurement of the vibration displacement perpendicular to the resonator surface. The interferometer is formed on an optical bench and covered with a plastic case in order to reduce the interference of environmental mechanical disturbances and fluctuations of the air. The laser beam emitted from a He-Ne laser is separated into two beams by a half mirror.

One of these beams is projected onto the PZT-driven mirror and the other beam is focused onto one point on the resonator surface. The reflected beams from the mirror and the resonator surface are superposed by the half mirror and focused on to the avalanche photo diode by a condenser lense. The resonator under study is attached to the measurement jig which can be moved along the bench surface and the direction perpendicular to it and can be rotated in any direction around the center of the resonator surface. The output signal of the avalanche photo diode is separated into two signals. One of these signals is filtered by LPF and the other is filtered at the frequency of the measured displacement and converted to the 20 kHz signal by a selective level meter. The output signals of the LPF and the selective level meter are led to the digital level meter. The ratio of these two signals is converted to the digital signal in dB value by the digital level meter. A microcomputer system, which is composed mainly of a 8 bit microprocessor, makes control of the digital level meter and storage of the data. A minicomputer system can make a determination of the optimally approximated displacement distribution function and make a graph of the displacement distributions.

Fig. 3 show the optical arrangement, which is called the Mach-zehnder interferometer, for the measurement of vibration displacements parallel to the resonator surface. In this case, the laser beams, one of which undergoes DC phase control and the low-frequency phase modulation by the PZT-driven mirror, are focused on one point of the resonator surface at the incident angle 45 degrees. Scattered beams perpendicular to the surface are collected by the condenser lense and focused on the avalanche photo diode.

Performance of Measurement system

In order to test the performance of the improved system, the measurement of a plano-convex AT-cut quartz crystal resonator (1.5MHz) has been made. Fig.4 shows the relation between the driving current of the resonator and the displacement component parallel to the resonator surface. The good linearity has been obtained between these two values. The lower limit of measuring range, in this case 4 angstroms, has been determined by the total gain of the optical detection system. The repeatability of measurement of a few percent has been obtained in the range from 5 angstroms to 600 angstroms. From the measurement of the third and the fifth overtone mode of the same resonator, it has been shown that the performances of the system are nearly equal at 4.5MHz and 7.5 MHz.

Measurement of SC-cut quartz crystal resonators

As an application of the measurement system, measurements of plano-convex SC-cut quartz crystal resonators have been made. Fig. 5 shows schematic view of the samples and Table 1 shows the dimensions of the three samples. The results measured on the

fundamental mode of sample 1 are shown in Fig. 6. U_x is the displacement component parallel to the x-direction, U_z parallel to the z-direction, and U_y perpendicular to the resonator surface, respectively. The distribution of the largest component U_x is confined in the center region of the plate and is similar to the thickness-shear mode of plano-convex AT cut resonators. However, the magnitudes of U_z and U_y are about one fourth of that of U_x , and the relative amplitudes of U_z and U_y are nearly five times of those of the AT-cut resonators. Fig. 7 and 8 show the distributions of the 3rd overtone SC-cut resonator (sample2) and 5th overtone SC-cut resonator (Sample3) respectively. In all cases, the relative magnitudes of U_z and U_y are much larger than those of the AT-cut resonator. Consequently, it can be said that the dependency of Q factor of the SC-cut resonator upon the gas pressure in the holder should be fairly large compared with that of the AT-cut resonator.

Conclusion

Improvements have been made on the laser interferometric measurement system of vibration displacement of quartz crystal resonators. It has been shown from the measurement of a sample resonator that the measurement of vibration displacement can be made down to five angstroms within a few percent repeatability in the frequency range from 1 MHz to 10 MHz. By using the new measurement system, the distributions of the three dimensional vibration displacement of the plano-convex SC-cut resonators have been studied. The results show that the vibration mode is similar to the thickness-shear mode of the plano convex AT-cut resonator, however, the relative amplitude of the vibration component perpendicular to the surface is nearly five times of that of the AT-cut resonator. From these results, it is concluded that this measurement system can be a powerful means for the study of quartz crystal resonators.

Acknowledgement

The authors wish to acknowledge Fujitsu Ltd. for the offer of sample SC-cut resonators.

References

- (1) K. Iijima, Y. Tsuzuki, Y. Hirose, and M. Akiyama; "Laser Interferometric Measurement of The Vibration displacements of a Plano-convex AT-cut Quartz Crystal Resonator", Proc. of 30th APCS, p.65-70, June 1976

TABLE 1 PLATE DIMENSIONS OF THREE SAMPLES

SAMPLE NO.	1	2	3
FREQUENCY (MHz)	2.5	2.5	2.5
OVERTONE ORDER	1	3	5
CUT ANGLE (YXW) φ/θ	22°33'/34°00'	22°30'/33°50'	22°30'/33°50'
BLANK RADIUS R [mm]	7.0	11.0	15.0
THICKNESS t [mm]	0.73	2.18	3.65
CURVATURE R ₁ [mm]	200	150	300
ELECTRODE R _e RADIUS [mm]	4.0	6.0	8.0

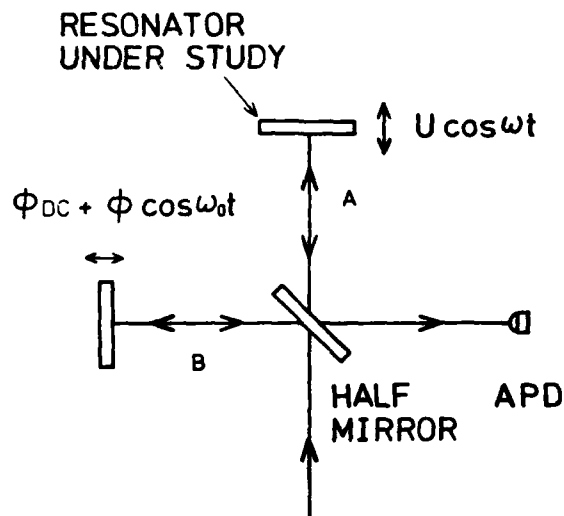


FIG.1 PRINCIPLE OF LASER INTERFEROMETRIC MEASUREMENT SYSTEM

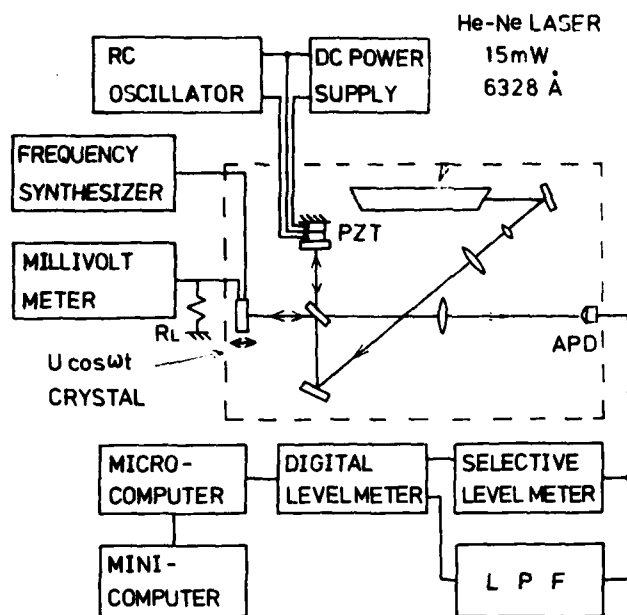


FIG.2 MEASUREMENT SYSTEM
(MICHELSON INTERFEROMETER)

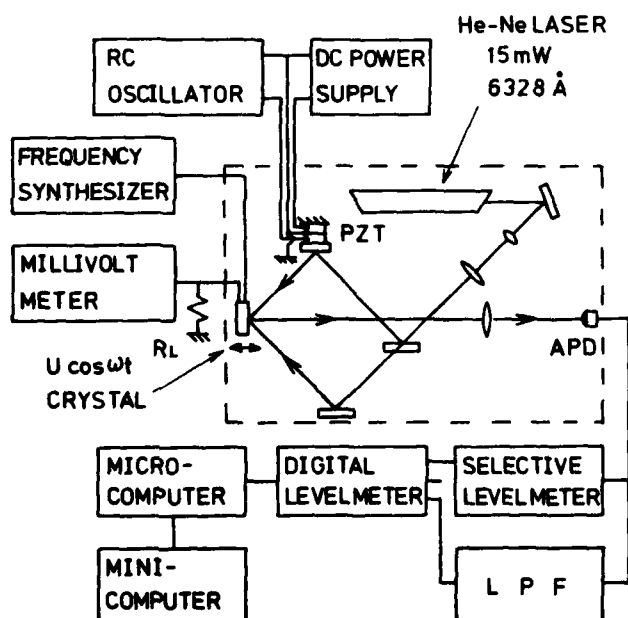


FIG.3 MEASUREMENT SYSTEM
(MACH-ZEHNDER INTERFEROMETER)

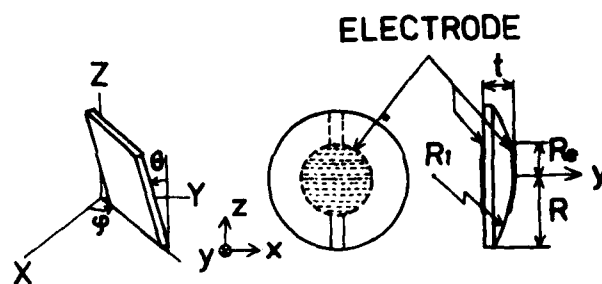


FIG.4 SCHEMATIC VIEW
OF SC-CUT RESONATOR

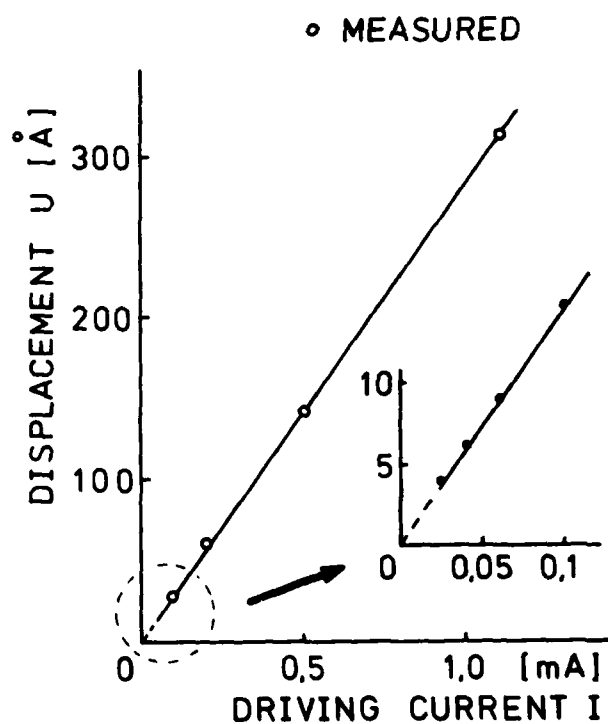
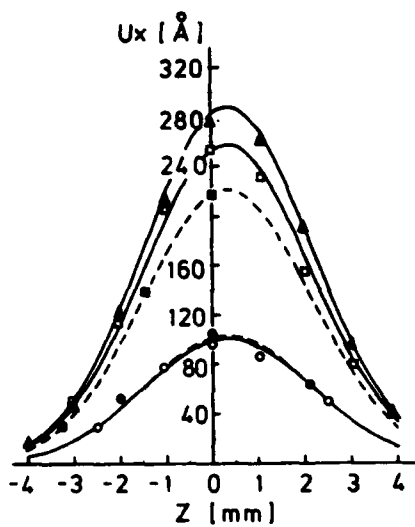
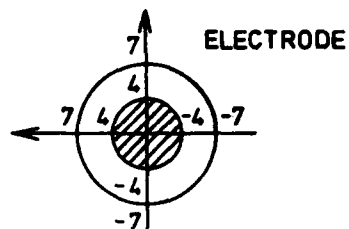


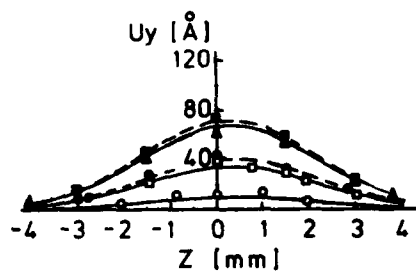
FIG.5 RELATION BETWEEN
DISPLACEMENT AND
DRIVING CURRENT



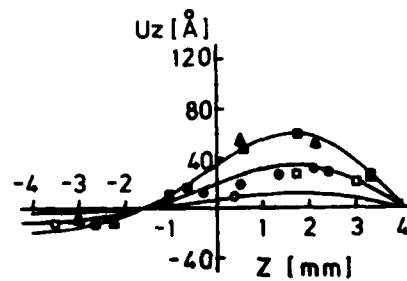
(a) DISTRIBUTION OF
DISPLACEMENT COMPONENT U_x



- $X = -3.0$ [mm]
- $X = -1.5$ [mm]
- ▲ $X = 0$ [mm]
- $X = 1.5$ [mm]
- ◆ $X = 3.0$ [mm]

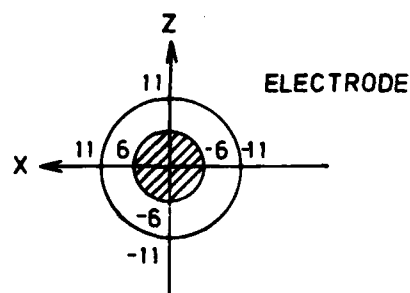
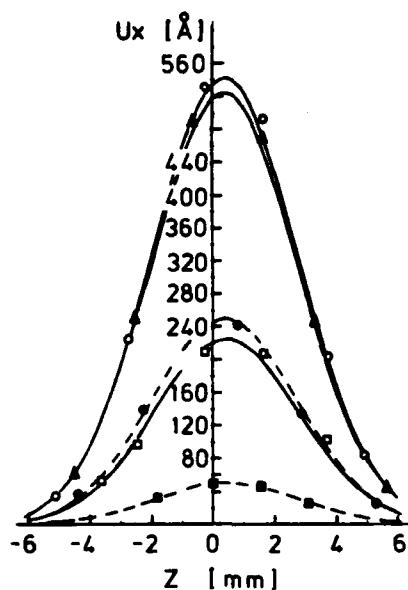


(b) DISTRIBUTION OF
DISPLACEMENT COMPONENT U_y



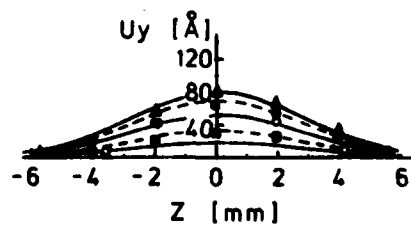
(c) DISTRIBUTION OF
DISPLACEMENT COMPONENT U_z

FIG.6 DISTRIBUTIONS OF DISPLACEMENTS OF FUNDAMENTAL
SC-CUT RESONATOR

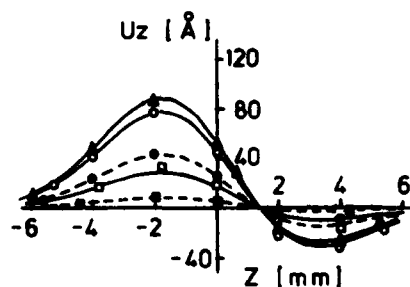


- $X = -4$ [mm]
- $X = -2$ [mm]
- △— $X = 0$ [mm]
- -○- - $X = 2$ [mm]
- -□- - $X = 4$ [mm]

(a) DISTRIBUTION OF
DISPLACEMENT COMPONENT U_x

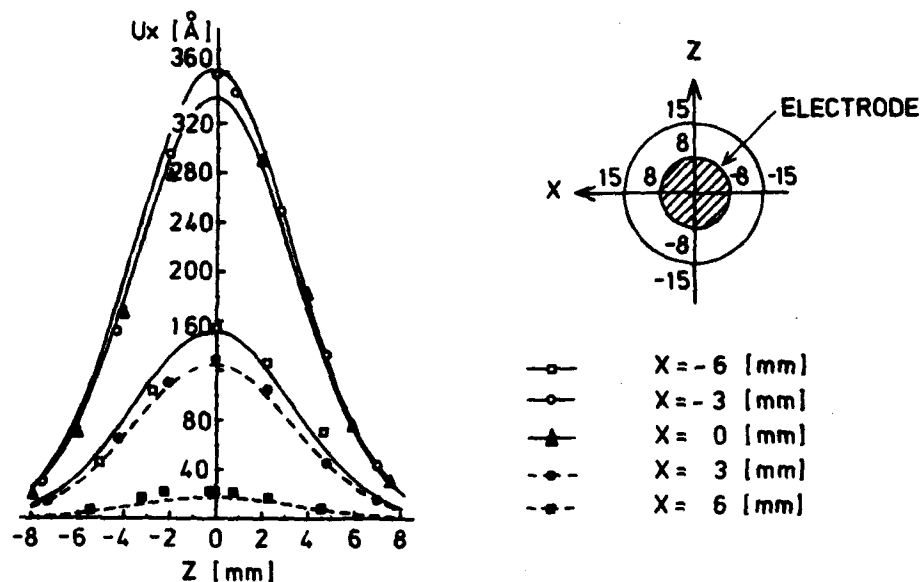


(b) DISTRIBUTION OF
DISPLACEMENT COMPONENT U_y

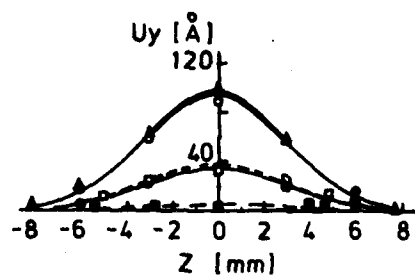


(c) DISTRIBUTION OF
DISPLACEMENT COMPONENT U_z

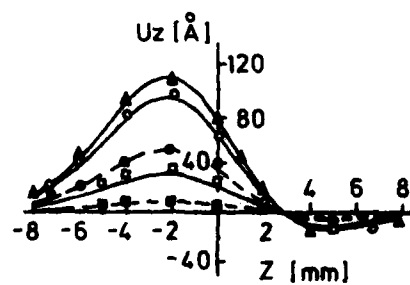
FIG.7 DISTRIBUTIONS OF DISPLACEMENTS OF 3RD OVERTONE
SC-CUT RESONATOR



(a) DISTRIBUTION OF
DISPLACEMENT COMPONENT U_x



(b) DISTRIBUTION OF
DISPLACEMENT COMPONENT U_y



(c) DISTRIBUTION OF
DISPLACEMENT COMPONENT U_z

FIG.8 DISTRIBUTIONS OF DISPLACEMENTS OF 5TH OVERTONE
SC-CUT RESONATOR

THE STRESS COEFFICIENT OF FREQUENCY OF QUARTZ PLATE RESONATORS

Muhammad Mizan & Arthur Ballato

U.S. Army Electronics Technology & Devices Laboratory (ERADCOM)
Fort Monmouth, New Jersey 07703

Abstract

The stress-frequency effect exists for both bulk and surface acoustic wave devices, and has been studied for more than thirty years. Eight years ago these studies led to the discovery of the SC-, or "stress-compensated" cut, which has found widespread acceptance for many frequency control applications. The calculations required to obtain $K_f(\psi)$ characterizing the effect are very elaborate, incorporating as they do the influences of lattice deformation and the third-order elastic constants.

These calculations have been carried out previously by EerNisse for the three thickness plate modes at six ψ values on the upper zero temperature coefficient (ZTC) locus. The results have now been used to extract simple polynomial coefficients that permit computation of $K_f(\psi)$, as well as several other pertinent measures, in simple engineering form for intermediate ψ values.

Key Words

Resonators, Frequency Control, Acoustic Waves, Quartz Crystals, Doubly Rotated Cuts, Bulk Acoustic Waves, Force-frequency Effects, Piezoelectric Crystals.

Introduction

The stress-frequency effect was observed more than thirty years ago for radial edge forces applied to quartz discs. References 1 through 4 contain in their bibliographies a rather complete listing of work in this area. The recent paper by Oura, et al.⁵ indicates that interest in, and practical application of the effect have not waned.

Studies related to this effect led to the discovery of the stress-compensated SC cut a few years ago. The introduction of this doubly rotated cut and others close in orientation to it is beginning to yield very appreciable practical dividends, and to spur additional work on electrode effects. Sherman's paper⁶ is a case in point.

Because the stress-frequency effect depends upon rather elaborate calculations involving nonlinear elastic coefficients, it is often inconvenient to "build in" its influence in analytic expressions. For this reason, a simple expression for $K_f(\psi)$ was developed for the AT cut,⁷ based on a least-squares curve fit. In this paper the method is extended to doubly rotated quartz discs vibrating in the three thickness modes. The present formulation is useful primarily because of its simplicity. The engineering coefficients obtained contain the information derived from the complicated ab initio calculations, but they may be applied

much more easily to practical problems.

It might be thought that with the advent of the SC cut, where electrode stress effects are absent, or negligible, one would have no need for considering any other cuts. This is not so. Cuts other than the SC are currently in use for a number of reasons:

- In oven controlled crystal oscillator (OCXO) applications, achieving a desired turnover temperature requires extremely tight θ tolerances; these are lessened by changing ϕ .
- At ϕ angles other than that of the SC, one or more of the following can be improved: amplitude-frequency effect, c-mode phase noise, b-c mode separation, piezoelectric coupling, etc.

In these cases it is necessary to know the effect on stress-frequency of the departure from the SC cut. From our results the necessary measure of sensitivity may be obtained. Additionally, since the results cover all three modes, the influence of stress effects on multi-mode oscillators is available for all ϕ values.

Edge-Stress

In Fig. 1 are shown the definitions for angles ϕ , θ , and ψ . The angles ϕ and θ specify the doubly rotated orientation, while ψ specifies the edge-force azimuth angle.

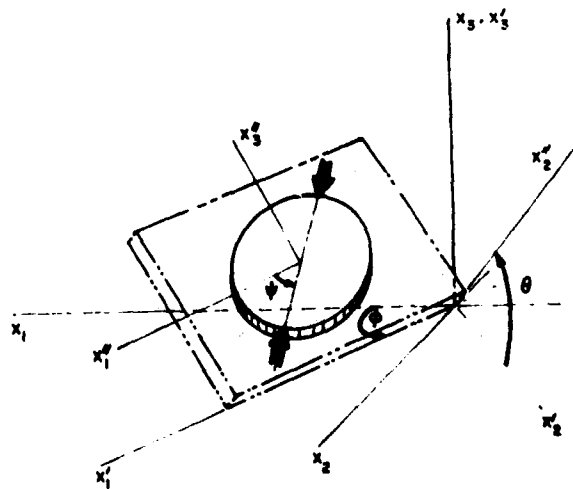


Figure 1. Reference angles ϕ and θ specifying doubly rotated plates, and force azimuth angle ψ .

AD P 002476

Figures 2 through 7 graph the input data obtained by EerNisse using a variational approach². The data are for the three thickness modes (a, b, & c), and represent values of K_T at roughly 50 ψ increments. The ϕ angles are 0° (AT cut), 10° , 15° (FC cut), 19.1° (IT cut), 22.4° (SC cut), and 30° . The θ angles are such as to place the orientations on the zero temperature coefficient (ZTC) locus, with $|\theta| \approx 34^\circ - 35^\circ$. It should be noted that the results to be given below are for this so-called upper ZTC locus; they should not be extrapolated to the lower locus. The figures also contain the least-squares fitted curves passed through the data points.

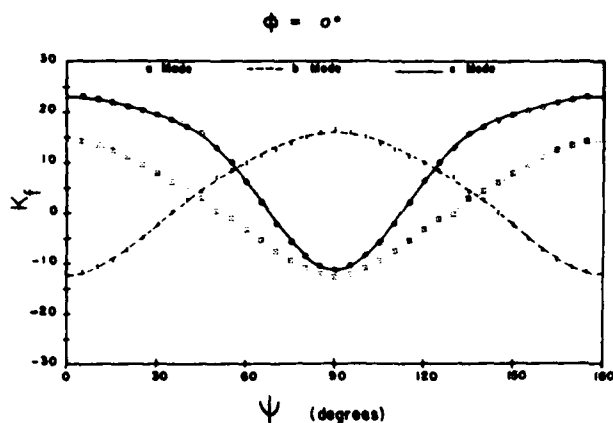


Figure 2. Input data points and resulting curve fits, $\phi = 0^\circ$.

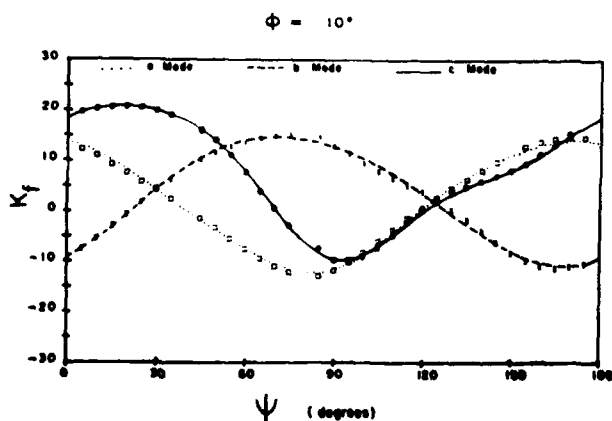


Figure 3. Input data points and resulting curve fits, $\phi = 10^\circ$.

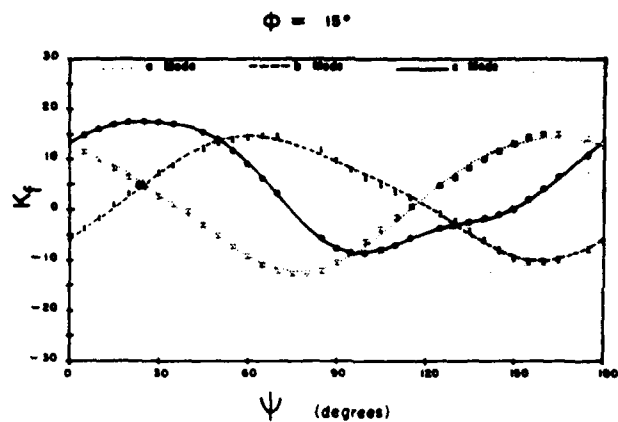


Figure 4. Input data points and resulting curve fits, $\phi = 15^\circ$.

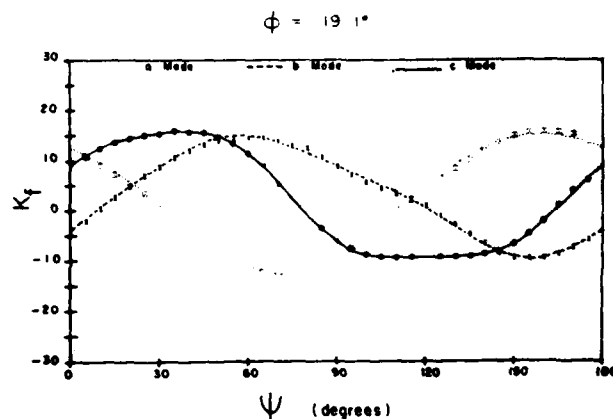


Figure 5. Input data points and resulting curve fits, $\phi = 19.1^\circ$.

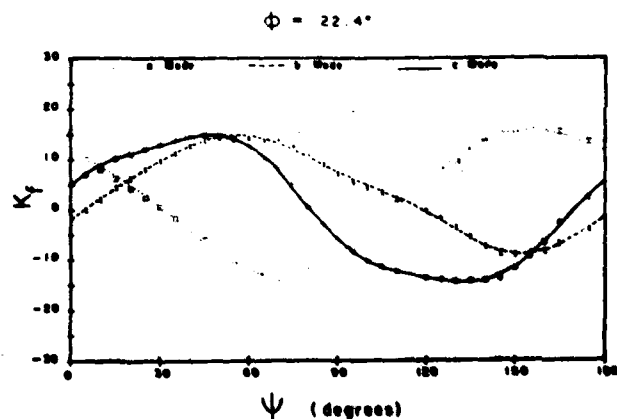


Figure 6. Input data points and resulting curve fits, $\phi = 22.4^\circ$.

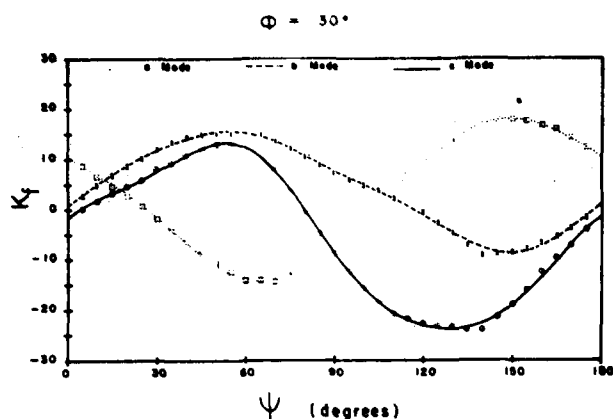


Figure 7. Input data points and resulting curve fits, $\phi = 30^\circ$.

Formulas and Coefficients

With due regard for the symmetries involved, and for the accuracies of both the input data and of the resulting curve fit, the following expressions were used for the fitting functions:

$$K_f(\phi, \psi) = \sum_{n=0}^3 A_n(\phi) \cos 2n\psi + \sum_{p=1}^3 B_p(\phi) \sin 2p\psi \quad (1)$$

$$A_n(\phi) = \sum_{q=0}^3 a_{nq} \cos 3q\phi \quad (2)$$

$$B_p(\phi) = \sum_{r=0}^3 b_{pr} \cos 3r\phi \quad (3)$$

The values of a_{nq} and b_{pr} differ for each mode; they are listed in Table 1 and Table 2, respectively. Units are 10^{-15} ms/N.

Table 1 - Values of a_{nq} for modes a, b, and c.

nq	MODE		
	a	b	c
00	4.1537	2.8717	-9.8628
01	-3.5254	0.9429	22.8714
02	1.5713	-1.3902	-4.4999
03	-0.4678	0.7174	2.1115
10	7.0787	-4.2748	14.6146
11	6.9915	-7.3642	-9.3982
12	-1.9925	-0.2015	12.3242
13	0.6338	-1.6466	-1.6736
20	-0.2821	-0.1120	-0.2374
21	-0.7029	-1.1791	-3.0783
22	0.2493	-0.1053	-0.4793
23	-0.0375	0.0759	-0.8885
30	0.1096	1.6201	0.2056
31	-0.0197	-1.4423	1.9022
32	0.8179	0.2886	-1.1293
33	-0.1552	-0.4597	0.3414

Table 2 - Values of b_{pr} for modes a, b, and c.

b_{pr}	MODE		
	a	b	c
b_{10}	-25.557	41.4561	21.8756
b_{11}	32.622	-58.8359	-21.6335
b_{12}	-13.177	30.7243	3.9990
b_{13}	5.958	-13.0410	-4.0039
b_{20}	2.1263	3.4438	5.4115
b_{21}	-2.7042	-4.6983	-9.7486
b_{22}	1.0297	2.2568	7.9220
b_{23}	-0.4576	-0.9930	-3.5330

Properties of $K_f(\phi)$

Once the coefficients a_{nq} and b_{pr} have been determined, the simple form of (1) permits the rapid computation of various parameters of interest. Examples are given in Figures 8, 9, and 10. Figure 8 shows the maximum and minimum values of K_f for the c mode as function of ϕ . In Figure 9 are plotted the ψ values at which these extrema are attained. An altitude chart of K_f is shown in Fig. 10, again for the c mode, as functions of ψ and ϕ angles. (Cf. Fig. 8 in Reference 2). Units are 10^{-15} ms/N.

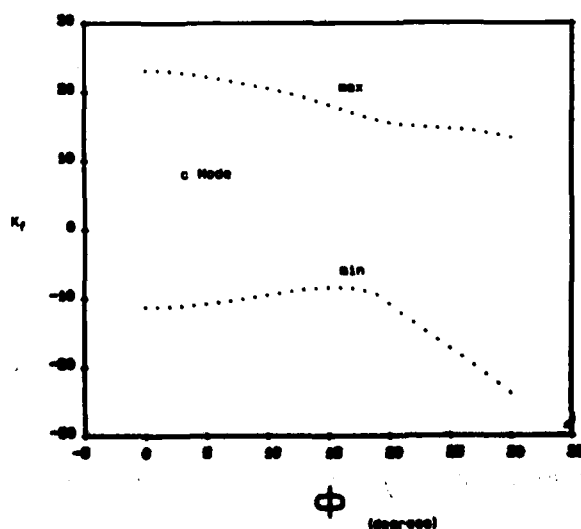


Figure 8. Values of the c-mode $K_f(\phi)$ extrema versus ϕ .

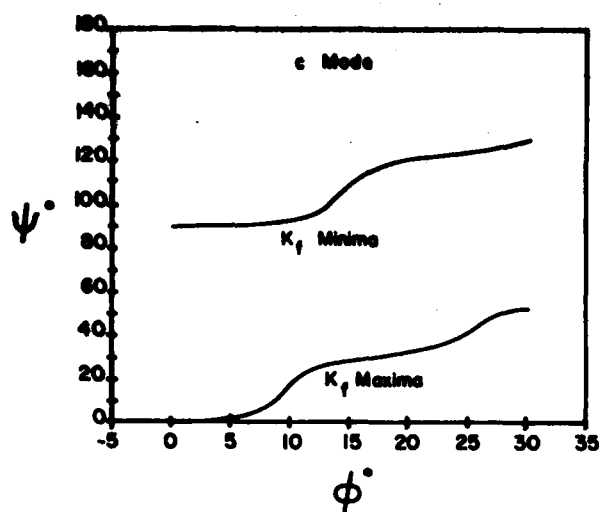


Figure 9. Psi angles of the c-mode $K_f(\phi)$ extrema versus ϕ .

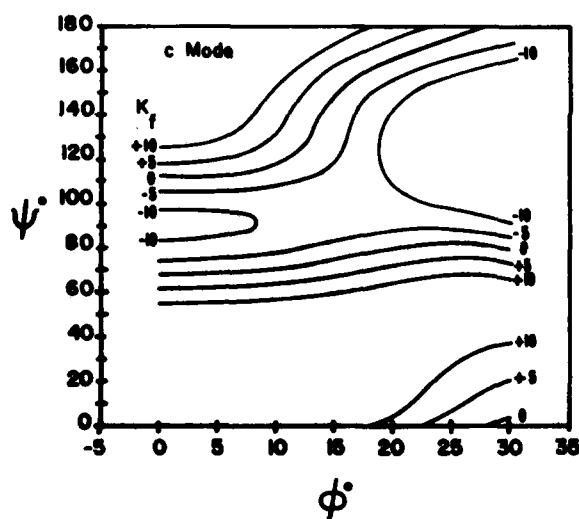


Figure 10. Altitude chart of $K_f(\phi, \psi) = \text{constant}$, for the c mode.

Film-Stress

The equivalence between a superposition of edge forces and film stress is indicated schematically in Fig. 11. The superposition is expressed in terms of the film-stress coefficient $\langle K_f(\phi) \rangle$; it is determined as the azimuth-averaged value of the edge-stress coefficient $K_f(\phi, \psi)$ as follows:

$$\langle K_f(\phi) \rangle = \frac{1}{\pi} \int_0^\pi K_f(\phi, \psi) d\psi. \quad (4)$$

Insertion of (1) into (3) gives

$$\langle K_f(\phi) \rangle = A_0(\phi). \quad (5)$$

The coefficient $A_0(\phi)$ for the c mode is plotted versus ϕ in Fig. 12; where it equals zero defines the SC cut. Also plotted are A_1 and A_2 . Figure 13 graphs the coefficients A_3 , B_1 , and B_2 versus ϕ for the c mode. Units are 10^{15} ms/N.

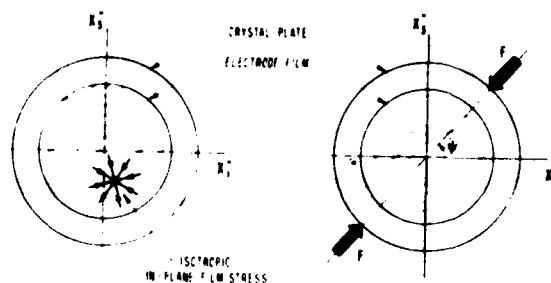


Figure 11. Electrode film stress considered as equivalent to a superposition of edge forces.

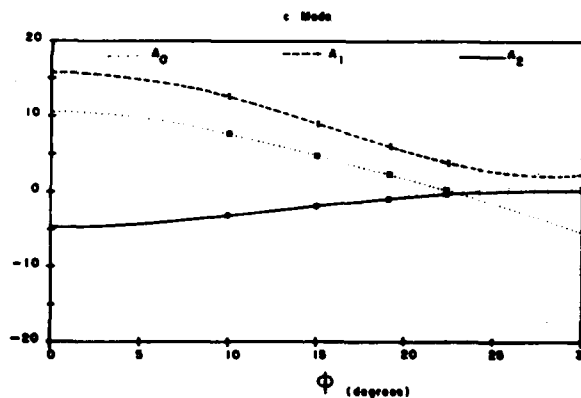


Figure 12. Coefficients $A_0(\phi)$, $A_1(\phi)$, and $A_2(\phi)$ for the c mode versus ϕ . $A_0(\phi) = 0$ defines the SC cut.

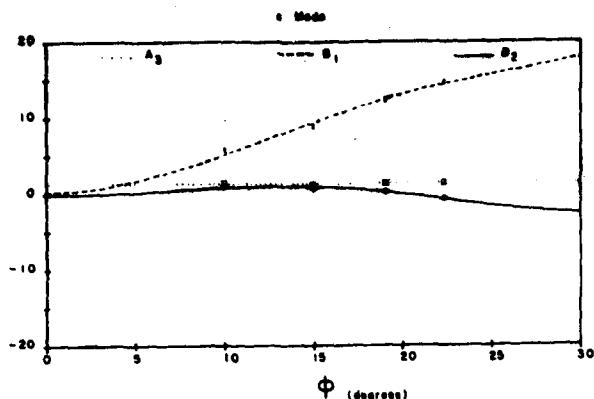


Figure 13. Coefficients $A_3(\phi)$, $B_1(\phi)$, and $B_2(\phi)$ for the c mode versus ϕ .

Table 3 lists the computed values of $A_0(\phi)$ for all three modes at six specific ϕ values. Units are 10^{15} ms/N. The nominal SC angle of $\phi = 22.40$ produces a computed c-mode A_0 value of 0.13×10^{-15} rather than zero; this is due to a number of factors: uncertainty in the input material values; accuracy limitations in the variational procedure used to obtain the input $K_f(\phi, \psi)$ data; residuals in the least-squares due to the simple fitting function chosen; the disparity between the flat plate geometry for which the calculations of $K_f(\phi, \psi)$ were carried out, and the finite, contoured resonators used to define the SC cut by means of its thermal-transient behavior.

From the coefficients in Tables 1 and 2, the zero of $A_0(\phi)$ occurs at

$$\phi \approx 22.70;$$

at this value, the slope is

$$\left. \frac{\partial A_0}{\partial \phi} \right|_{SC} \approx -0.65 \times 10^{-15} \text{ m.s/N, } ^\circ\phi.$$

Table 3 - $\langle K_f(\phi) \rangle$ for modes a, b, and c.

ϕ°	mode a	mode b	mode c
0	1.7267	3.1308	10.6357
10	1.8994	3.0443	7.6708
15	2.0035	2.9303	4.7859
19.1	2.0327	3.3507	2.3644
22.4	2.1377	3.4999	0.1314
30	2.5811	4.2643	-5.3582

Relative sensitivities of film-stress coefficients are plotted in Fig. 14 for the a, b, and c modes, versus ϕ , compared to the AT-cut c-mode value. The ordinate scale is logarithmic. From this graph the sensitivity of the c-mode to changes in ϕ angle about the SC value may be determined. It is also seen that the b- and a-mode values are somewhat less than that of the c-mode AT cut, and are nearly invariant with ϕ .

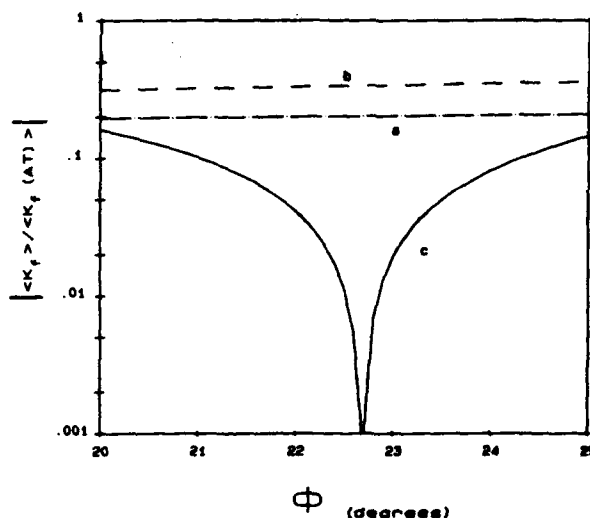


Figure 14. Absolute value of the film stress coefficient $\langle K_f(\phi) \rangle$, normalized to that of the AT cut, as function of ϕ , for modes a, b, and c.

Conclusion

• Results of complex calculations determining the stress-frequency effect from fundamental material constants have been put into simple engineering terms.

• The edge-stress-frequency coefficient $K_f(\phi, \psi)$ and the film-stress-frequency coefficient $\langle K_f(\phi) \rangle$ are available as simple analytic expressions for cuts on the upper zero temperature coefficient of frequency locus, for the a, b, and c thickness modes of plates.

Acknowledgement

We thank E.P. EerNisse of QUARTEX, Inc. for the kind permission to use his calculations of $K_f(\phi, \psi)$ as our input data.

References

1. A. Ballato, E.P. EerNisse, and T. Lukaszek, "The force-frequency effect in doubly rotated quartz resonators," Proc. 31st Annual Frequency Control Symposium, U.S. Army Electronics Command, Fort Monmouth, NJ 07703, June 1977, pp. 8-16.
2. E.P. EerNisse, T.J. Lukaszek, and A. Ballato, "Variational calculation of force-frequency constants of doubly rotated quartz resonators," IEEE Trans. Sonics Ultrason., vol. SU-25, no. 3, May 1978, pp. 132-138.
3. A. Ballato, E.P. EerNisse, and T.J. Lukaszek, "Experimental verification of stress compensation in the SC cut," 1978 Ultrasonics Symposium Proceedings, IEEE Cat. #78CH 1344-ISU, September 1978, pp. 144-147.
4. T.J. Lukaszek and A. Ballato, "Resonators for severe environments," Proc. 32nd Annual Frequency Control Symposium, U.S. Army Electronics R&D Command, Fort Monmouth, NJ 07703, May-June 1979, pp. 311-321.
5. N. Oura, M. Kuramochi, J. Nakamura, Y. Miyazaki, and M. Yamashina, "Frequency changes in fifth-overtone 5-MHz circular plano-convex AT- and IT-cut plates due to radially applied three forces," IEEE Trans. Sonics Ultrason., vol. SU-30, no. 2, March 1983, pp. 111-114.
6. J.H. Sherman, Jr., "Temperature coefficient of the frequency shift arising from electrode film stress," IEEE Trans. Sonics Ultrason., vol. SU-30, no. 2, March 1983, pp. 104-110.
7. A. Ballato, "Force-frequency compensation applied to four-point mounting of AT- cut resonators," IEEE Trans. Sonics Ultrason., vol. SU-25, no. 4, July 1978, pp. 223-226.

TEMPERATURE DERIVATIVES OF ELASTIC STIFFNESSES DERIVED
FROM THE FREQUENCY-TEMPERATURE BEHAVIOR OF QUARTZ PLATES

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

Abstract

Linear field equations for small vibrations superposed on thermally induced deformations by steady and uniform temperature changes are derived from the nonlinear field equations of thermoelasticity in Lagrangian formulation.

From the solutions of these equations for the thickness-vibrations, the temperature derivatives of elastic stiffnesses are related analytically to the known or measured properties such as the second and third order elastic stiffnesses, thermal expansion coefficients, and temperature coefficients of frequency of quartz plates.

Six values of the first temperature derivative $C_{pq}^{(1)}$ and six values of the effective second temperature derivative $\tilde{C}_{pq}^{(2)}$ are calculated from the temperature coefficients of frequency measured by Bechmann, Ballato and Lukaszek for various doubly-rotated quartz plates.

The presently calculated values are compared with the first temperature derivatives obtained by Sinha and Tiersten. In the incremental stress-strain-temperature relations, certain expressions involving the elastic stiffnesses, temperature derivatives and thermal expansion coefficients can be identified as having similar significance as the temperature coefficients of C_{pq} defined by Bechmann. Values of these expressions are calculated and compared with the existing values. The loci of the zeros of the first and second order temperature coefficients of frequency for thickness shear (B and C) modes and the frequency-temperature characteristics of the LC-cut are studied and compared with experimental values.

Introduction

The dependence of resonance frequencies of thickness-vibrations in doubly-rotated quartz plates on the steady and uniform temperature changes was studied in a comprehensive paper by Bechmann, Ballato and Lukaszek (1962,1963) in which the linear elastic stress-strain relations were employed.^{1,2} Their values of the temperature coefficients of the second-order elastic stiffnesses have since been applied to predict frequency shifts owing to temperature changes in various crystal vibrators of finite extent.

In 1966, a complete set of values of the third-order elastic stiffnesses C_{pqr} of quartz was published by Thurston, McSkimin and Andreatch.³ The presence of the nonlinear terms associated with C_{pqr} in the stress-strain relations was found to be essential in the studies of the force and acceleration sensitivities of crystal plates.⁴ Holland realized there is a need to extend the work by Bechmann, et al, for the design of resonators subject simultaneously to thermal and stress biases.⁶ By a nonlinearly based formalism and referring the position of material points to a fixed reference frame, Sinha and Tiersten calculated the first temperature derivatives of the second-order elastic stiffnesses of quartz,⁷ and have shown that predictions of temperature coefficient of velocity of surface waves based on

these values are in substantially better agreement with experimental data.⁸

In the present paper, nonlinear field equations of thermoelasticity in Lagrangian formulation are employed at the final and initial states, respectively. By taking the differences of the corresponding field equations in the final and initial states, a set of linear field equations for small vibrations superposed on the thermally induced deformations owing to steady and uniform temperature changes in crystal are obtained.

Solutions of these equations for the thickness-vibrations are obtained by a perturbation method, from which the first temperature derivatives $C_{pq}^{(1)}$ and the effective second temperature derivatives $\tilde{C}_{pq}^{(2)}$ are related by analytical expressions to the first and second order temperature coefficients of frequency, second and third order elastic stiffnesses, and the thermal expansion coefficients of quartz.

By employing the above mentioned relations and the least squares regression analysis, six values of $C_{pq}^{(1)}$ and six values of $\tilde{C}_{pq}^{(2)}$ are obtained. In performing calculations, we have used the values of the first and second order frequency coefficients for thickness vibrations of quartz plates by Bechmann, Ballato & Lukaszek (1963),² the values of thermal expansion coefficients and second order elastic stiffnesses by Bechmann, Ballato & Lukaszek (1962),¹ and the values of the third order elastic stiffnesses by Thurston, McSkimin & Andreatch.³

These presently calculated values are compared with the first order temperature derivatives obtained by Sinha and Tiersten.⁷ By regrouping the terms in the incremental stress-strain-temperature relations, certain expressions in terms of the elastic stiffnesses, temperature derivatives and thermal expansion coefficients can be identified as having similar significance as the temperature coefficients of C_{pq} defined by Bechmann.

Therefore values of these expressions are calculated and are listed with the values of the temperature coefficients of C_{pq} obtained by Bechmann, Ballato and Lukaszek,¹ Adams, Enslow, Kusters and Ward,⁹ and Kahan.¹⁰ The loci of the zeros of the first and second order frequency coefficients for thickness shear (B & C) modes as a function of ϕ and θ , and the frequency-temperature characteristics of the LC-cut are predicted and compared with experimental data.

Equations of Motion for Small Vibrations
Superposed on Thermally Induced Deformations

The process of superposing small vibrations on thermally induced deformations in a crystal may be described by the three consecutive states the crystal goes through as follows:

1. Natural State: At first, the crystal is said to be at a natural state when it is at rest, free of stress and strain, and has a uniform temperature T_0 . Referred to a common rectangular Cartesian frame of reference as shown in Fig. 1, x_i ($i = 1,2,3$) denotes the position

AD P002477

of a generic material point, ρ_0 the mass density, C_{ijkl} , C_{ijklmn} , and $C_{ijklmnpq}$ the second-, third-, and fourth-order elastic stiffnesses of the crystal.

2. Initial State: The crystal is now subject to a steady and uniform temperature increase from T_0 to T , and allowed to expand freely. At this state, the position of a material point is moved, due to thermal expansion, from x_i to y_i .

The nonlinear governing equations of thermoelasticity in Lagrangian formulation are given as follows:

$$\begin{aligned} U_i &= y_i - x_i, \\ E_{ij} &= \frac{1}{2}(U_{j,i} + U_{i,j} + U_{k,i}U_{k,j}), \\ T_{ij} &= C_{ijkl}^0 E_{kl} + \frac{1}{2}C_{ijklmn}^0 E_{kl}E_{mn} + \frac{1}{6}C_{ijklmnpq}^0 E_{kl}E_{mn}E_{pq} - \lambda_{ij}^0, \\ (T_{ij} + T_{jk}U_{i,k})_{,j} &= \rho_0 \ddot{U}_i, \\ P_i &= n_j(T_{ij} + T_{jk}U_{i,k}) \text{ on } S, \end{aligned} \quad (1)$$

where U_i , E_{ij} , T_{ij} , and P_i denote initial displacement, Lagrangian Strain Tensor, Kirchhoff-Piola stress tensor (of second kind), and surface traction all referred to the natural state. Also, n_i are the components of the unit outward normal to the bounding surface S of the crystal and θ is the temperature increase from T_0

$$\theta = T - T_0. \quad (2)$$

In the third of (1), the stress-strain-temperature relations, material properties are assumed to be temperature dependent, i.e.,

$$\begin{aligned} C_{ijkl}^0 &= C_{ijkl} + C_{ijkl}^{(1)} \cdot \theta + \frac{1}{2}C_{ijkl}^{(2)} \cdot \theta^2, \\ C_{ijklmn}^0 &= C_{ijklmn} + C_{ijklmn}^{(1)} \cdot \theta, \\ C_{ijklmnpq}^0 &= C_{ijklmnpq}, \\ \lambda_{ij}^0 &= \lambda_{ij}^{(1)} \cdot \theta + \lambda_{ij}^{(2)} \cdot \theta^2, \end{aligned} \quad (3)$$

where λ_{ij}^0 is the stress coefficient of temperature and

$$\begin{aligned} C_{ijkl}^{(1)} &= \left. \frac{\partial C_{ijkl}^0}{\partial T} \right|_{T_0}, \\ C_{ijkl}^{(2)} &= \left. \frac{\partial^2 C_{ijkl}^0}{\partial T^2} \right|_{T_0}, \\ C_{ijklmn}^{(1)} &= \left. \frac{\partial C_{ijklmn}^0}{\partial T} \right|_{T_0}, \end{aligned} \quad (4)$$

are temperature derivatives of the elastic stiffnesses evaluated at temperature T_0 in the natural state. It is the values of these derivatives that we intend to compute from the solutions of the thickness-vibrations and existing experimental data.

When values of the linear thermal expansion coefficient

$$\alpha_{ij}^0 = \alpha_{ij}^{(1)} \cdot \theta + \alpha_{ij}^{(2)} \cdot \theta^2 \quad (5)$$

are measured instead of those of λ_{ij}^0 , we have by its definition and symmetry

$$\frac{1}{2}(U_{j,i} + U_{i,j}) = \alpha_{ij}^0 = \alpha_{ji}^0 \quad (6)$$

and by the second equation of (1)

$$E_{ij} = \alpha_{ij} + \frac{1}{2}\alpha_{ki}\alpha_{kj}. \quad (7)$$

There is no difficulty in using (7). However, the quadratic terms, estimated from Bechmann's values² of $\alpha_{ij}^{(1)}$ and $\alpha_{ij}^{(2)}$ for quartz, are of the order of 10^{-3} smaller than the first term in (7) and hence are neglected. Therefore for initial fields, we have

$$\begin{aligned} U_{j,i} &= U_{i,j} = E_{ij} = \alpha_{ij}^0, \\ T_{ij} &= 0, \\ \ddot{U}_i &= 0, \end{aligned} \quad (8)$$

since the plate is at rest and allowed free expansion.

3. Final State: When small-amplitude vibrations are superposed on thermally induced deformations, the plate is said to be at the final state. The position of the material point is moved from y_i to z_i . We let

$u_i = z_i - y_i$ be the incremental displacement due to the vibrations. Then the total displacement from the natural state to the final state is, as shown in Fig. 1,

$$\bar{U}_i = U_i + u_i = z_i - x_i \quad (9)$$

Similarly, we let

$$\begin{aligned} \bar{E}_{ij} &= E_{ij} + e_{ij}, \\ \bar{T}_{ij} &= T_{ij} + t_{ij}, \\ \bar{P}_i &= P_i + p_i, \end{aligned} \quad (10)$$

where \bar{E}_{ij} , \bar{T}_{ij} , and \bar{P}_i are the total strain, stress and traction, and e_{ij} , t_{ij} and p_i the incremental strain, stress and traction, respectively.

The governing equations of the total fields should necessarily be the same as (1) for initial fields. Therefore we have

$$\begin{aligned} \bar{E}_{ij} &= \frac{1}{2}(\bar{U}_{j,i} + \bar{U}_{i,j} + \bar{U}_{k,i}\bar{U}_{k,j}), \\ \bar{T}_{ij} &= C_{ijkl}^0 \bar{E}_{kl} + \frac{1}{2}C_{ijklmn}^0 \bar{E}_{kl}\bar{E}_{mn} + \frac{1}{6}C_{ijklmnpq}^0 \bar{E}_{kl}\bar{E}_{mn}\bar{E}_{pq} - \lambda_{ij}^0, \\ (\bar{T}_{ij} + \bar{T}_{jk}\bar{U}_{i,k})_{,j} &= \rho_0 \ddot{\bar{U}}_i, \\ \bar{P}_i &= n_j(\bar{T}_{ij} + \bar{T}_{jk}\bar{U}_{i,k}) \text{ on } S. \end{aligned} \quad (11)$$

In (11), we have assumed that the temperature of the crystal remains at T , or the temperature variation caused by the incremental vibrations is negligible.

Subtraction of (1) from (11), respectively, gives the governing equations for incremental fields,

$$\begin{aligned} e_{ij} &= \frac{1}{2}(u_{j,i} + u_{i,j} + u_{k,j}u_{k,i} + u_{k,i}u_{k,j}) \\ t_{ij} &= \left[C_{ijkl}^{(0)} + C_{ijklmn}^{(0)}E_{mn} + \frac{1}{2}C_{ijklmnpq}^{(0)}E_{mn}E_{pq} \right] e_{kl} \\ (t_{ij} + t_{jk}u_{i,k} + t_{jk}u_{i,k})_{,j} &= \rho_0 \ddot{u}_i \\ p_i &= n_j(t_{ij} + t_{jk}u_{i,k} + t_{jk}u_{i,k}) \text{ on } S. \end{aligned} \quad (12)$$

In (12), the terms containing the products of incremental fields have been disregarded, since incremental deformations are assumed to be small. We note (12) are very similar to Eqs. (16)-(18) of Reference 4 except the constitutive equations.

Further substitution of the material property-temperature relations (3) and (5) and the thermally induced initial fields (8) into (12) gives the incremental stress-strain-temperature relations

$$t_{ij} = \left[C_{ijkl}^{(1)} + D_{ijkl}^{(1)}\theta + D_{ijkl}^{(2)}\theta^2 \right] e_{kl} \quad (13)$$

The displacement equations of motion

$$G_{ijkl}u_{k,jl} = \rho_0 \ddot{u}_i \quad (14)$$

and the traction boundary conditions

$$p_i = n_j G_{ijkl}u_{k,l} \text{ on } S \quad (15)$$

where

$$\begin{aligned} D_{ijkl}^{(1)} &= C_{ijkl}^{(1)} + C_{ijklmn}^{(1)}\alpha_{mn}^{(1)} \\ D_{ijkl}^{(2)} &= \frac{1}{2}C_{ijkl}^{(2)} + C_{ijklmn}^{(2)}\alpha_{mn}^{(2)} \\ \tilde{C}_{ijkl}^{(2)} &= C_{ijkl}^{(2)} + 2C_{ijklmn}^{(1)}\alpha_{mn}^{(1)} + C_{ijklmnpq}^{(1)}\alpha_{mn}^{(1)}\alpha_{pq}^{(1)} \\ G_{ijkl} &= C_{ijkl} + G_{ijkl}^{(1)}\theta + G_{ijkl}^{(2)}\theta^2 \\ G_{ijkl}^{(1)} &= C_{sjkl}\alpha_{is}^{(1)} + C_{ijsl}\alpha_{ks}^{(1)} + C_{ijklmn}\alpha_{mn}^{(1)} + C_{ijkl}^{(1)} \\ G_{ijkl}^{(2)} &= C_{sjtl}\alpha_{is}^{(1)}\alpha_{kt}^{(1)} + C_{sjkl}\alpha_{is}^{(2)} + C_{ijsl}\alpha_{ks}^{(2)} \\ &\quad + C_{sjkl}\alpha_{is}^{(1)} + C_{ijsl}\alpha_{ks}^{(1)} + C_{sjklmn}\alpha_{is}^{(1)}\alpha_{mn}^{(1)} \\ &\quad + C_{ijslan}\alpha_{ks}^{(1)}\alpha_{mn}^{(1)} + C_{ijklmn}\alpha_{mn}^{(2)} + \frac{1}{2}C_{ijkl}^{(2)} \end{aligned} \quad (16)$$

We see from (16) that $D_{ijkl}^{(n)}$ has the same symmetry as C_{ijkl} , but

$$G_{ijkl}^{(n)} = G_{klij}^{(n)} \quad (17)$$

In (13)-(15) terms containing θ^3 or higher powers are dropped. We note that among the unknown temperature derivatives of elastic stiffnesses in (3), only $C_{ijkl}^{(1)}$ and $\tilde{C}_{ijkl}^{(2)}$ appear explicitly in $D_{ijkl}^{(n)}$ and $G_{ijkl}^{(n)}$ and, in turn, in the field equations (13)-(15),

$\tilde{C}_{ijkl}^{(2)}$ is defined in the third equation of (16) and is called the effective second temperature derivative of C_{ijkl} .

We see therefore in problems of small vibrations superposed on homogeneous thermal strains, only $C_{ijkl}^{(1)}$ and $\tilde{C}_{ijkl}^{(2)}$ are needed to describe the temperature dependence of elastic stiffnesses up to the second power of θ ; $C_{ijkl}^{(2)}$, $C_{ijklmn}^{(1)}$ and $C_{ijklmnpq}^{(1)}$ do not appear separately and hence cannot be determined separately either.

Thickness-Vibrations

Let X_i ($i = 1, 2, 3$) be the crystallographical axes of quartz. A doubly-rotated quartz plate is defined by $(y, x, \omega, \ell)\phi, \theta$ in the conventional IEEE notation.¹¹ Let $2b$ be the thickness and n (or n_i) the unit outward normal to the faces of the plate in the natural state. Referred to X_i coordinates, the components of the normal in terms of rotation angles ϕ and θ are

$$n_1 = -\cos\theta \sin\phi, \quad n_2 = \cos\theta \cos\phi, \quad n_3 = \sin\theta \quad (18)$$

For traction-free face conditions, (15) becomes

$$p_i = n_j G_{ijkl}u_{k,l} = 0 \text{ at } n_j X_j = \pm 2b \quad (19)$$

The solution for harmonic, antisymmetric thickness-vibrations

$$u_k = A_k \sin \xi n_p X_p e^{i\omega t} \quad (20)$$

satisfies (14) and (19) provided

$$\begin{aligned} (Q_{ik} - \lambda \delta_{ik})A_k &= 0 \\ \xi &= \frac{n\pi}{2b}, \quad n = 1, 3, 5, \dots \end{aligned} \quad (21)$$

where

$$\begin{aligned} Q_{ik} &= Q_{ki} = G_{ijkl}n_j n_l \\ \lambda &= \rho_0 \frac{\omega^2}{\xi^2} = \rho_0 \left(\frac{2b\omega}{n\pi} \right)^2 \end{aligned} \quad (22)$$

In the eigenvalue problem (21), $Q_{ik} = G_{ijkl}n_j n_l$ and G_{ijkl} , defined in (16), depends on the unknown temperature derivatives $C_{ijkl}^{(1)}$ and $\tilde{C}_{ijkl}^{(2)}$. In order to obtain analytical solutions from which $C_{ijkl}^{(1)}$ and $\tilde{C}_{ijkl}^{(2)}$ can be calculated from other known or measured properties, the Rayleigh-Schrodinger perturbation method is employed.¹² We let the amplitude and eigenvalue be

$$\begin{aligned} A_k &= A_k^{(0)} + A_k^{(1)}\theta + A_k^{(2)}\theta^2 \\ \lambda &= \lambda^{(0)} + \lambda^{(1)}\theta + \lambda^{(2)}\theta^2 \end{aligned} \quad (23)$$

where $\theta = T - T_0$ is assumed to be small.

By inserting (23) into (21) and equating coefficients of like powers of θ , we obtain

$$\theta^0: \left(Q_{ik}^{(0)} - \lambda^{(0)} \delta_{ik} \right) A_k^{(0)} = 0 \quad (24)$$

$$\Theta^1: \left(Q_{ik}^{(0)} - \delta_{ik} \lambda^{(0)} \right) A_k^{(1)} + \left(Q_{ik}^{(1)} - \delta_{ik} \lambda^{(1)} \right) A_k^{(0)} = 0 \quad (25)$$

$$\Theta^2: \left(Q_{ik}^{(0)} - \delta_{ik} \lambda^{(0)} \right) A_k^{(2)} + \left(Q_{ik}^{(1)} - \delta_{ik} \lambda^{(1)} \right) A_k^{(1)} + \left(Q_{ik}^{(2)} - \delta_{ik} \lambda^{(2)} \right) A_k^{(0)} = 0 \quad (26)$$

where

$$\begin{aligned} Q_{ik}^{(0)} &= G_{ijkl} n_j n_l \\ Q_{ik}^{(1)} &= G_{ijkl}^{(1)} n_j n_l \\ Q_{ik}^{(2)} &= G_{ijkl}^{(2)} n_j n_l \end{aligned} \quad (27)$$

From the zero-order equation (24), we can solve for $\lambda^{(0)}$ and $A_k^{(0)}$ uniquely by normalization, i.e., requiring

$$A_k^{(0)} A_k^{(0)} = 1 \quad (28)$$

Multiplication of (25) by $A_i^{(0)}$ and utilization of (24) and (28) leads to

$$\lambda^{(1)} = A_i^{(0)} Q_{ik}^{(1)} A_k^{(0)} \quad (29)$$

Solutions of $A_k^{(1)}$ from (25) exist,¹³ since the second term $(Q_{ik}^{(1)} - \delta_{ik} \lambda^{(1)}) A_k^{(0)}$ is orthogonal to the eigenvalue of $(Q_{ik}^{(0)} - \delta_{ik} \lambda^{(0)})$, and $A_k^{(1)}$ can be determined uniquely by requiring $A_k^{(1)}$ orthogonal to $A_k^{(0)}$, i.e.,

$$A_k^{(0)} A_k^{(1)} = 0 \quad (30)$$

Similarly, multiplication of (26) by $A_i^{(0)}$, and employment of (28) and (30) results to

$$\lambda^{(2)} = A_i^{(0)} Q_{ik}^{(2)} A_k^{(0)} + A_i^{(0)} Q_{ik}^{(1)} A_k^{(1)} \quad (31)$$

Normally, we compute the first- and second-order eigenvalues due to temperature changes from (29) and (31). However in the present work, we want to compute $C_{ijkl}^{(1)}$ and $C_{ijkl}^{(2)}$ from these equations by regarding $\lambda^{(1)}$ and $\lambda^{(2)}$ as known quantities.

Substitution of $\omega = 2\pi f$ into the second equation of (22) gives

$$f^2 = \frac{n^2}{4\rho_0 b^2} (\lambda^{(0)} + \lambda^{(1)} \cdot \Theta + \lambda^{(2)} \cdot \Theta^2) \quad (32)$$

and $f_0^2 = n^2 \lambda_0^2 / 4\rho_0 b^2$. By taking the first and second derivatives of (32) with respect to temperature T and substituting Bechmann's definition of temperature coefficient of frequency

$$Tf^{(n)} = \frac{1}{n f_0} \frac{\partial^n f}{\partial T^n} \bigg|_{T_0}$$

into the resulting expressions, we obtain

$$\begin{aligned} \lambda^{(1)} &= 2\lambda^{(0)} T f^{(1)} \\ \lambda^{(2)} &= 2\lambda^{(0)} T f^{(2)} + \lambda^{(0)} (T f^{(1)})^2 \end{aligned} \quad (34)$$

By substituting (34) into (29) and (31) and rearranging so as to have the unknown temperature derivatives on the left hand side, we have

$$\begin{aligned} C_{ijkl}^{(1)} n_j n_l A_i^{(0)} A_k^{(0)} &= 2\lambda^{(0)} T f^{(1)} - \left[C_{ijsl}^{(1)} \alpha_{ks}^{(1)} \right. \\ &\quad \left. + C_{sjkl}^{(1)} \alpha_{is}^{(1)} + C_{ijklmn}^{(1)} \alpha_{mn}^{(1)} \right] n_j n_l A_i^{(0)} A_k^{(0)} \end{aligned} \quad (35)$$

$$\begin{aligned} \tilde{C}_{ijkl}^{(2)} n_j n_l A_i^{(0)} A_k^{(0)} &= 4\lambda^{(0)} T f^{(2)} + 2\lambda^{(0)} (T f^{(1)})^2 \\ &\quad - 2 \left[C_{sjkl}^{(2)} \alpha_{is}^{(2)} + C_{ijsl}^{(2)} \alpha_{ks}^{(2)} + C_{sjtl}^{(1)} \alpha_{kt}^{(1)} \right. \\ &\quad \left. + C_{sjkl}^{(1)} \alpha_{is}^{(1)} + C_{ijsl}^{(1)} \alpha_{ks}^{(1)} + C_{sjklmn}^{(1)} \alpha_{mn}^{(1)} \right. \\ &\quad \left. + C_{ijslmn}^{(1)} \alpha_{mn}^{(1)} + C_{ijklmn}^{(2)} \right] n_j n_l A_i^{(0)} A_k^{(0)} \\ &\quad - 2 \left[C_{sjkl}^{(1)} \alpha_{is}^{(1)} + C_{ijsl}^{(1)} \alpha_{ks}^{(1)} + C_{ijklmn}^{(1)} \alpha_{mn}^{(1)} \right. \\ &\quad \left. + C_{ijkl}^{(1)} \right] n_j n_l A_i^{(0)} A_k^{(1)} \end{aligned} \quad (36)$$

Temperature Derivatives of Elastic Stiffnesses of Quartz

Equations (35) and (36) can be applied to calculations of temperature derivatives for crystals of any symmetry. For α quartz, a trigonal crystal, $C_{ijkl}^{(1)}$ and $C_{ijkl}^{(2)}$ each has six independent components referring to the crystallographic axes. By expanding¹⁴ the left side of (35) and (36) and adopting the abbreviated notions, these two equations can be rewritten as

$$\begin{aligned} C_{66}^{(1)} x_1 + C_{22}^{(1)} x_2 + C_{44}^{(1)} x_3 + C_{14}^{(1)} x_4 \\ + C_{13}^{(1)} x_5 + C_{33}^{(1)} x_6 = y^{(1)} \end{aligned} \quad (35)'$$

$$\begin{aligned} \tilde{C}_{66}^{(2)} x_1 + \tilde{C}_{22}^{(2)} x_2 + \tilde{C}_{44}^{(2)} x_3 + \tilde{C}_{14}^{(2)} x_4 \\ + \tilde{C}_{13}^{(2)} x_5 + \tilde{C}_{33}^{(2)} x_6 = y^{(2)} \end{aligned} \quad (36)'$$

where $y_1^{(1)}$ and $y_2^{(2)}$ are the right hand side expressions of (35) and (36), respectively, and

$$\begin{aligned} x^1 &= A_1^{(0)2} n_2^2 - 2A_1^{(0)} A_2^{(0)} n_1 n_2 + A_2^{(0)2} n_1^2 \\ x^2 &= A_1^{(0)2} n_1^2 + 2A_1^{(0)} A_2^{(0)} n_1 n_2 + A_2^{(0)2} n_2^2 \\ x^3 &= 2 \left[A_1^{(0)2} n_2 n_3 + 2A_1^{(0)} A_2^{(0)} n_1 n_3 + 2A_1^{(0)} A_3^{(0)} n_1 n_2 \right. \\ &\quad \left. - A_2^{(0)2} n_2 n_3 + A_2^{(0)} A_3^{(0)} n_1^2 - A_2^{(0)} A_3^{(0)} n_2^2 \right] \\ x^4 &= A_1^{(0)2} n_3^2 + 2A_1^{(0)} A_3^{(0)} n_1 n_3 + A_2^{(0)2} n_3^2 + 2A_2^{(0)} A_3^{(0)} n_2 n_3 \end{aligned} \quad (33)$$

$$\begin{aligned}
& + A_3^{(0)2} n_1^2 + A_3^{(0)2} n_2^2 \\
x^5 &= 2 \left(A_1^{(0)} A_3^{(0)} n_1 n_3 + A_2^{(0)} A_3^{(0)} n_2 n_3 \right) \\
x^6 &= A_3^{(0)2} n_3^2
\end{aligned}
\tag{37}$$

From (18) and (24), n_i and $A_k^{(0)}$ are obtained. Then x_i are calculated according to (37). In computing $y^{(1)}$ and $y^{(2)}$, we have used the measured values of temperature coefficients of frequency $Tf^{(n)}$ for various doubly rotated cuts by Bechmann, Ballato and Lukaszek,² thermal expansion coefficients $\alpha_{ij}^{(n)}$, second-order elastic stiffnesses C_{ijkl} (or C_{pq}) of Bechmann, Ballato and Lukaszek,¹ and the third-order elastic stiffnesses C_{ijklmn} (or C_{pqr}) of Thurston, McSkimin, and Andreatch.³

We determine the six values of $C_{pq}^{(1)}$ in (35)' by employing the method of least squares¹⁵ according to the following algorithm:

1. Obtain the least squares estimate of $C_{pq}^{(1)}$ and their standard errors from a given data set.
2. Construct a 99.9% confidence interval.
3. Reject data points outside the interval.
4. Obtain a subset of data.
5. Go back to Step 1.

We begin the calculation with 117 data points given in Table 5 of Ref. 2. After five repeated calculations as described above, the number of data points is reduced to 28. The data of these selected points from Ref. 2 are given in Table I. In the last column of Table I, we have assigned a value of 1.0 (one) as the weight for observations from the well known orientations like AT, X, Y, and IT cuts, and a value of 0.5 for the rest of the observations. Among these 28 data points, 11 intersect with the 24 data points listed in Table III of Ref. 1. The same 28 data points have also been used in (36)' for calculating $\tilde{C}^{(2)}$.

Least squares estimated values of the first temperature derivatives and the effective second temperature derivatives are tabulated, respectively, in Table II and III with their standard deviations. Also shown in Table II are the first temperature derivatives obtained by Sinha and Tiersten.

In the incremental stress-strain-temperature relations (13), we observe that $D_{ijkl}^{(1)}/C_{ijkl}$ and $D_{ijkl}^{(2)}/C_{ijkl}$ have similar significance to the temperature coefficients of elastic stiffnesses, $TC_{ijkl}^{(1)}$ and $TC_{ijkl}^{(2)}$ employed in Refs. 1 and 2. We note these coefficients are not equivalent, for they are defined in two different formulations. However, for purposes of comparison, we report the values of $D_{pq}^{(1)}/C_{pq}$ and $D_{pq}^{(2)}/C_{pq}$, respectively, in Tables IV and V along with the temperature coefficients of elastic stiffnesses given by Bechmann, Ballato and Lukaszek,¹ Adams, Enslow, Kusters and Ward,⁹ and Kahan.¹⁰

Based on our values of $C_{pq}^{(1)}$ and $\tilde{C}^{(2)}$, the first- and second temperature coefficients of frequency are predicted and compared with the 117 observed values¹ in Table VI. The loci of $Tf^{(1)}$ and $Tf^{(2)}$ for C mode and B mode ($f_B > f_C$) are also computed and shown in Figs. 2 and 3, respectively. It may be seen from Table VI and Figs. 2 and 3 that results of our present prediction are either comparable or slightly better than those predicted in Ref. 1.

Finally, the frequency-temperature characteristics for LC cut of quartz vibrating in A, B, and C modes ($f_A > f_B > f_C$) are studied. Fig. 4 shows that our predicted curves agree well with measured values.¹⁶

Acknowledgments

We wish to thank Dr. A. Ballato of the U.S. Army Electronics Technology and Devices Laboratory for many valuable discussions.

This work was partially supported by the U.S. Army Research Office, Contract No. DAAG29-79-C-0019.

References

1. R. Bechmann, A. D. Ballato, and T. J. Lukaszek, "Higher-Order Temperature Coefficients of the Elastic Stiffnesses and Compliances of the α -Quartz," *Proc. IRE* 50, 1962, 1812.
2. R. Bechmann, A. D. Ballato, and T. J. Lukaszek, "Higher-Order Temperature Coefficients of the Elastic Stiffnesses and Compliances of α -Quartz," *USAEIRD Tech. Report* 2261, September 1963.
3. R. N. Thurston, H. J. McSkimin and P. Andreatch, Jr., "Third-Order Elastic Constants of Quartz," *J. Appl. Phys.* 37, 1966, 267.
4. P. C. Y. Lee, Y. S. Wang, and X. Markenscoff, "High Frequency Vibrations of Crystal Plates Under Initial Stresses," *J. Acoust. Soc. Am.*, 57, No. 1, 1975, pp. 95-105.
5. P. C. Y. Lee, K. M. Wu and Y. S. Wang, "Effects of Acceleration on the Resonance Frequencies of Crystal Plates," *J. Acoust. Soc. Am.* 63(4), 1978, pp. 1039-1047.
6. R. Holland, "Temperature Coefficients of Stiffness in Quartz," *IEEE Trans. Sonics Ultrason.*, SU-23, 1976, 72.
7. B. K. Sinha and H. F. Tiersten, "First Temperature Derivatives of the Fundamental Elastic Constants of Quartz," *J. Appl. Physics*, 50(4), 1979, pp. 2732-2739.
8. B. K. Sinha and H. F. Tiersten, "On the Temperature Dependence of the Velocity of Surface Waves in Quartz," *Ultrasonics Symposium Proceedings*, 1978, pp. 662-665.
9. C. A. Adams, G. M. Enslow, J. A. Kusters, and R. W. Ward, "Selected Topics in Quartz Crystal Research," *Proc. 24th Ann. Freq. Control Symp.*, 1970, pp. 55-63.
10. Alfred Kahan, "Elastic Constants of Quartz and Their Temperature Coefficients," *Proc. 36th Ann. Freq. Control Symp.*, 1982, pp. 159-167.
11. "Standards on Piezoelectric Crystals, 1949," *Proc. IRE*, 37, 1949, p. 1378.

12. A. Nayfeh, "Perturbation Methods," John Wiley & Sons, Inc., New York 1973.
13. Francis B. Hildebrand, "Methods of Applied Mathematics," 2nd Edition, 1965, Prentice-Hall, Inc., Englewood Cliffs, NJ, pp. 30-34.
14. Reduce 2, a system and language for algebraic manipulation by Anthony C. Hearn, 1973, may be used to perform the tedious expansion.
15. N. R. Draper and H. Smith, "Applied Regression Analysis," John Wiley & Sons, New York 1966, Chapter 2.
16. M. Nakazawa, H. Ito, A. Usui, A. Ballato, and T. Lukaszek, "New Quartz Resonators with Precision Frequency Linearity over a Wide Temperature Range," Proc. 36th Ann. Freq. Control Symp., 1982, pp. 290-296.

TABLE III

Second Temperature Derivatives of the Elastic Stiffnesses
for Alpha-Quartz at 25°C, $\tilde{C}_{pq}^{(2)} 10^3 \text{ N/m}^2/\text{C}^2$

pq	PRESENT	
	$\tilde{C}_{pq}^{(2)}$	Standard Error
66	11.021	1.68
22	-13.121	1.19
44	-26.593	2.04
14	4.2596	.715
13	-18.787	3.50
33	-20.926	10.2

TABLE I
Selected 28 data points from Ref. 1

ORIENTATION (IN DEGREES)	MODE	$T_f^{(1)}$ $10^{-6}/\text{C}^\circ$	$T_f^{(2)}$ $10^{-9}/\text{C}^\circ$	WEIGHTS IN LEAST SQUARES ANALYSIS
10.0	-38.000	A	-94.70	0.5
30.0	0.0	A	-20.50	1.0
0.0	-49.217	B	0.0	0.5
0.0	-55.000	B	-13.70	0.5
0.0	-60.000	B	-24.50	0.5
5.0	-48.000	B	-1.50	0.5
30.0	20.000	B	-9.30	0.5
30.0	40.000	B	-22.20	0.5
30.0	50.000	B	-28.00	0.5
0.0	35.250	C	0.0	1.0
0.0	0.0	C	92.50	1.0
5.0	-24.000	C	33.10	0.5
5.0	-28.000	C	-1.70	0.5
10.0	-30.000	C	0.80	0.5
10.0	-31.000	C	0.61	0.5
10.0	-32.000	C	0.26	0.5
10.0	-38.000	C	-5.50	0.5
10.0	-40.000	C	-8.80	0.5
10.0	-48.000	C	-27.40	0.5
12.5	-33.500	C	-0.40	0.5
20.0	34.333	C	-0.06	1.0
30.0	10.000	C	13.30	0.5
30.0	20.000	C	7.00	0.5
30.0	30.000	C	5.00	0.5
30.0	32.000	C	6.30	0.5
30.0	34.000	C	0.75	0.5
30.0	36.000	C	-4.55	0.5
30.0	60.000	C	-72.80	0.5

TABLE IV

First Temperature Coefficients of Elastic Stiffnesses
for Alpha-Quartz at 25°C, $10^{-6}/\text{C}^\circ$

pq	Present $D_{pq}^{(1)}/\text{C}^\circ$	Bechmann, et al. ¹ $T_{Cpq}^{(1)}$	Adams, et al. ⁹ $T_{Cpq}^{(1)}$	Kahan ¹⁰ $T_{Cpq}^{(1)}$
11	-68.2	-48.5	-49.6	-35.6
13	-705	-550	-651	-612
14	84.2	101	89	93.2
33	-197	-160	-192	-205
44	-186	-177	-172	-184
66	158	178	167	180

TABLE II

First Temperature Derivatives of the Elastic Stiffnesses
for Alpha-Quartz at 25°C, $\tilde{C}_{pq}^{(1)} 10^6 \text{ N/m}^2/\text{C}^\circ$

pq	PRESENT $\tilde{C}_{pq}^{(1)}$	STANDARD ERROR	Sinha and Tiersten ⁷
66	5.0747	.029	5.053
22	1.5976	.035	1.575
44	-5.3780	.045	-5.198
14	0.91675	.018	0.8814
13	-2.1983	.061	-2.127
33	-6.3255	.19	-7.140

TABLE V

Second Temperature Coefficients of Elastic Stiffnesses
for Alpha-Quartz at 25°C, $10^{-9}/\text{C}^2$

pq	Present $D_{pq}^{(2)}/\text{C}^\circ$	Bechmann, et al. ¹ $T_{Cpq}^{(2)}$	Adams, et al. ⁹ $T_{Cpq}^{(2)}$	Kahan ¹⁰ $T_{Cpq}^{(2)}$
11	-117	-107	-107	-117
13	-1019	-1150	-1021	-988
14	-34.6	-48	-19	-46.6
33	-157	-275	-162	-100
44	-272	-216	-261	-273
66	130	118	164	172

TABLE VI

Comparison of the observed and calculated first- and second-order temperature coefficients of frequency $Tf^{(n)}$ for the thickness modes of quartz plates.

ORIENTATION (IN DEGREES)	MODE	$Tf^{(1)}$ $10^{-6}/^{\circ}\text{C}$		$Tf^{(2)}$ $10^{-9}/^{\circ}\text{C}^2$	
		OBSERVED	CALCULATED	OBSERVED	CALCULATED
10.0	-28.000	A	-78.80	-88.01	
10.0	-30.000	A	-84.70	-94.80	
12.5	-34.000	A	-81.03	-88.61	-129.00
15.0	-33.000	A	-78.00	-78.98	-110.00
15.0	-35.000	A	-72.00	-81.20	-88.30
20.0	-36.000	A	-70.00	-74.19	-93.20
30.0	0.0	A	-20.50	-20.42	-53.20
30.0	10.000	A	-29.30	-31.68	-67.00
30.0	20.000	A	-42.00	-44.96	-86.00
30.0	30.000	A	-56.80	-57.31	-105.00
0.0	-40.000	B	20.70	19.35	-15.20
0.0	-45.000	B	10.70	8.27	-32.00
0.0	-49.237	B	0.0	-0.98	-40.00
0.0	-50.000	B	-1.30	-2.69	-44.40
0.0	-55.000	B	-12.70	-12.46	-71.70
0.0	-60.000	B	-24.50	-24.02	-75.00
5.0	-43.000	B	35.70	8.33	-9.50
5.0	-45.000	B	3.20	4.52	-44.60
5.0	-47.000	B	-0.90	0.68	-61.10
5.0	-48.000	B	-1.50	-1.24	-63.10
5.0	-49.000	B	-5.10	-3.20	-53.50
10.0	-24.000	B	18.20	-0.54	-34.00
10.0	-26.000	B	16.60	2.91	-43.00
10.0	-30.000	B	10.00	15.34	-27.00
10.0	-32.000	B	9.80	12.97	-33.80
10.0	-33.000	B	7.50	7.79	-31.90
10.0	-34.000	B	5.70	10.61	-41.50
10.0	-35.000	B	5.40	9.41	-42.90
10.0	-38.000	B	1.10	5.74	-39.00
10.0	-40.000	B	-2.40	3.19	-41.90
10.0	-42.000	B	-5.90	0.54	-47.50
12.5	-30.000	B	2.22	7.18	-27.20
12.5	-31.000	B	2.50	6.59	-32.30
12.5	-32.000	B	2.10	5.83	-33.80
12.5	-33.000	B	1.80	4.97	-34.90
12.5	-35.000	B	1.50	4.33	-43.40
12.5	-36.000	B	-2.10	4.07	-44.70
15.0	-31.000	B	-11.20	-0.73	-21.70
15.0	-32.000	B	-8.80	-0.74	-39.10
15.0	-33.000	B	-7.50	-1.13	-33.20
15.0	-34.000	B	-7.30	-1.72	-37.40
15.0	-34.500	B	-7.40	-2.06	-28.10
15.0	-35.000	B	-7.45	-2.42	-33.50
15.0	-35.500	B	-8.50	-2.79	-39.30
20.0	-36.000	B	-19.90	-11.67	-30.50
30.0	10.000	B	-14.90	-15.73	-62.50
30.0	20.000	B	-9.20	-9.22	-29.90
30.0	30.000	B	-19.10	-17.87	-24.50
30.0	40.000	B	-22.20	-21.99	-39.30
30.0	50.000	B	-28.00	-27.75	-43.90
0.0	50.000	C	-71.50	-65.54	-82.60
0.0	60.000	C	-17.50	-23.96	-16.00
0.0	35.250	C	0.0	-0.18	0.16
0.0	10.000	C	96.30	90.64	
0.0	0.0	C	82.50	92.80	
0.0	-10.000	C	89.70	81.48	
0.0	-20.000	C	65.80	63.04	
0.0	-24.000	C	33.10	32.88	-58.50
5.0	-25.000	C	-0.57	2.41	-82.20
5.0	-26.000	C	-6.00	-5.25	-28.30
5.0	-28.000	C	-1.70	-1.69	-16.40
10.0	-28.000	C	3.30	-0.16	-40.00
10.0	-30.000	C	0.80	0.56	-11.50
10.0	-31.000	C	0.61	0.43	-10.70
10.0	-32.000	C	0.26	0.10	-9.80
10.0	-33.000	C	-0.87	-0.41	-7.81
10.0	-34.000	C	-1.70	-1.10	-6.30
10.0	-35.000	C	-0.20	-1.95	-5.80
10.0	-36.000	C	-3.40	-2.98	-6.00
10.0	-36.000	C	-5.30	-5.64	-5.90
10.0	-38.000	C	-8.80	-8.78	-6.50
10.0	-48.000	C	-27.40	-27.52	-19.60
12.5	-25.000	C	12.10	14.49	-35.60
12.5	-26.000	C	12.30	14.38	-37.90
12.5	-27.000	C	12.50	14.72	-38.60
12.5	-28.000	C	14.60	7.80	-39.70
12.5	-31.000	C	2.70	1.41	-10.50
12.5	-31.500	C	2.00	1.07	-9.30
12.5	-32.000	C	1.90	0.73	-8.80
12.5	-32.500	C	1.40	0.36	-8.70
12.5	-33.000	C	-0.30	-0.03	-7.60
12.5	-33.500	C	-0.40	-0.46	-7.40
12.5	-34.000	C	-0.04	-0.91	-7.50
15.0	-30.000	C	13.30	8.48	-21.10
15.0	-31.000	C	11.30	3.21	-13.70
15.0	-32.000	C	6.60	1.78	-8.70
15.0	-33.000	C	4.70	0.58	-7.20
15.0	-34.000	C	1.80	-0.60	-7.16
15.0	-34.500	C	1.30	-1.20	-7.30
15.0	-35.000	C	-0.20	-1.83	-6.52
15.0	-35.500	C	-0.60	-2.47	-7.30
17.5	-34.500	C	3.49	-0.91	-8.24
17.5	-35.000	C	0.20	-1.47	-8.30
17.5	-35.500	C	1.34	-2.48	-8.40
17.5	-36.000	C	-1.02	-3.23	-8.90
20.0	34.333	C	-0.06	-0.28	-10.10
20.0	-28.000	C	4.00	7.13	-32.60
20.0	-30.000	C	4.60	10.88	-30.30
20.0	-32.000	C	7.10	8.22	-26.90
20.0	-34.000	C	7.00	0.68	-13.80
20.0	-36.000	C	2.70	-3.32	-10.50
20.0	-37.000	C	-0.70	-9.17	-8.73
20.0	-38.000	C	-6.10	-7.08	-11.45
20.0	-39.000	C	-8.20	-9.07	-11.32
25.0	-26.000	C	4.30	3.38	-23.20
25.0	-28.000	C	3.70	4.08	-23.15
25.0	-30.000	C	3.50	6.14	-24.20
25.0	-34.000	C	4.90	1.36	-18.00
30.0	10.000	C	13.30	13.64	-7.90
30.0	20.000	C	7.60	7.16	-16.90
30.0	30.000	C	5.00	6.74	-20.90
30.0	32.000	C	6.30	6.20	-17.80
30.0	34.000	C	0.78	0.72	-15.80
30.0	36.000	C	-4.88	-4.69	-14.30
30.0	40.000	C	-14.90	-15.49	-19.80
30.0	50.000	C	-64.49	-64.99	-80.19
30.0	60.000	C	-72.89	-72.41	-101.70

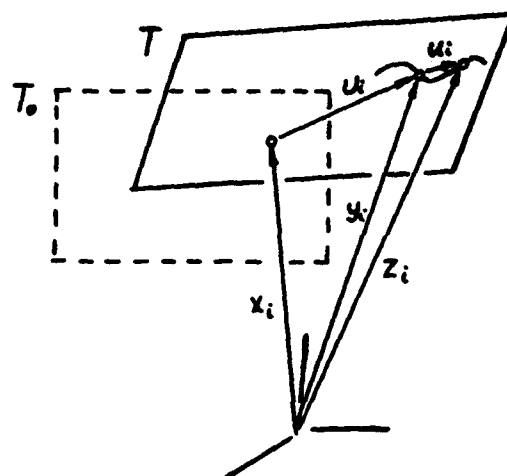
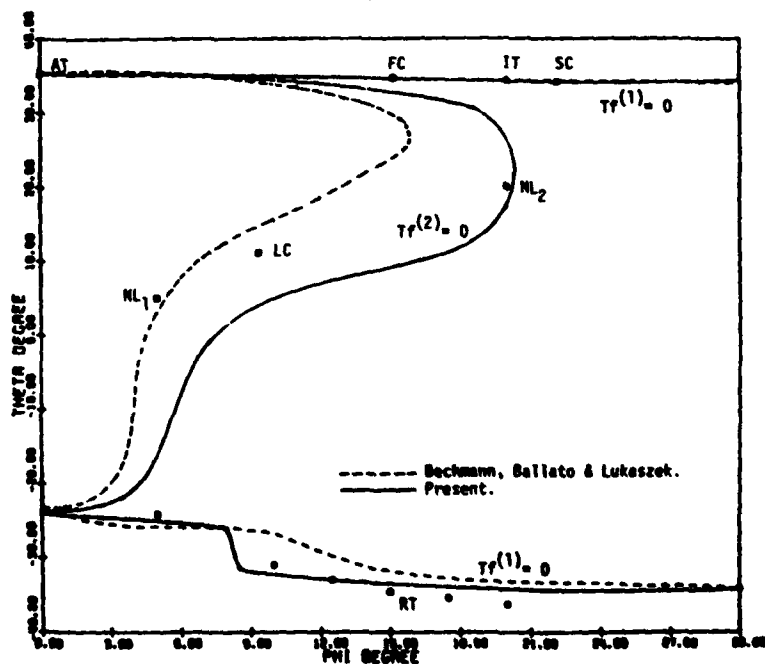
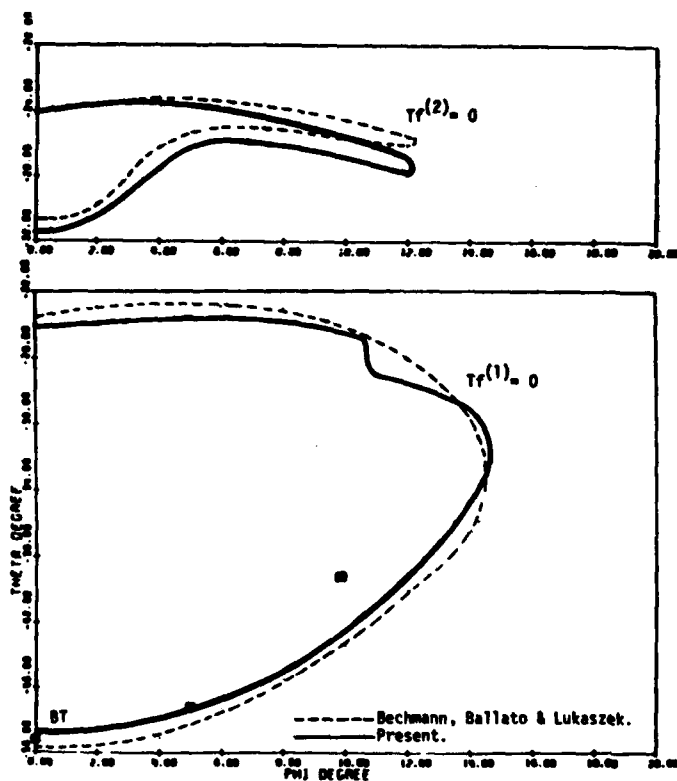


Fig. 1 Positions of a material point at the natural, initial, and final states.



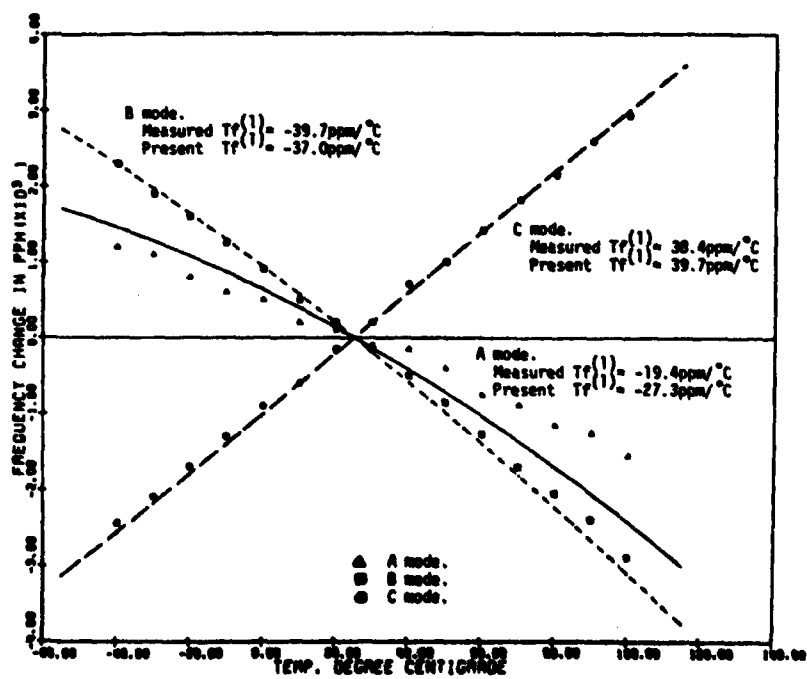
LOCI OF $Tf(1) = 0$ AND $Tf(2) = 0$ FOR THE THICKNESS MODE C

Fig. 2 Loci of $Tf(1) = 0$ and $Tf(2) = 0$ for the thickness mode C of quartz plates.



LOCI OF $T_f(1) = 0$ AND $T_f(2) = 0$ FOR THE THICKNESS MODE B

Fig. 3 Loci of $T_f(1) = 0$ and $T_f(2) = 0$ for the thickness mode B of quartz plates.



LC CUT. A, B & C THICKNESS MODES.

Fig. 4 Frequency-temperature characteristics for the LC-cut resonator operating on the A, B, and C modes.

TRANSIENT THERMALLY INDUCED FREQUENCY EXCURSIONS IN
AT- AND SC-CUT QUARTZ TRAPPED ENERGY RESONATORS

D.S. Stevens and H.F. Tiersten
Department of Mechanical Engineering,
Aeronautical Engineering & Mechanics
Rensselaer Polytechnic Institute
Troy, New York 12181

Abstract

In the AT- and SC-cuts of quartz the change in resonant frequency with temperature of a particular pure thickness mode (the C-mode) is essentially negligible in the absence of the electrodes, and the cut is said to be thermally compensated for that mode. Consequently, certain other effects, in particular the presence of the thin electrode films and the transverse trapping of energy, which is always present on account of the finite dimensions of the electrodes, have an important influence on the transient temperature dependence of the resonant frequencies of thermally compensated cuts. In this work the time-dependent frequency shifts due to rapid changes in the thermal environment at the surface of the crystal are calculated from the solution of the time-dependent heat conduction equation through the resulting thermally induced deformation state, the temperature derivatives of the fundamental elastic constants of quartz, the temperature dependence of the motional capacitive effect for the cuts considered and the now known trapped energy mode shapes with the aid of the equation for the perturbation in eigenfrequency of a piezoelectric solution due to a bias. Both the AT- and SC-cuts are treated and the results are compared. The influence of electrode size is automatically included in the calculation.

1. Introduction

A perturbation analysis of the linear electroelastic equations for small fields superposed on a bias has been performed¹. The change in resonant frequency due to any bias such as, e.g., a residual stress may readily be obtained from the resulting equation for the perturbation of the eigenfrequency if the bias is known. In addition, a system of approximate plate equations for the determination of thermal stresses in thin piezoelectric plates coated with large much thinner films was derived² from Mindlin's plate equations³⁻⁵ and subsequently modified⁶ somewhat. The resulting approximate equations⁶ simplify the treatment of many thermal stress problems considerably, and the three-dimensional detail not included in the approximate description is not deemed to be important for our purposes.

When a quartz resonator is subject to a change in the ambient temperature, it undergoes a time-dependent, inhomogeneous temperature distribution which causes the resonant frequency to drift in time^{7,8} until the new thermal equilibrium state is reached. In recent work⁹ on the influence of thermal transients on the resonant frequencies of pure thickness vibrations of quartz plates it was shown that for rapid temperature inputs in thermally compensated cuts the calculated frequency shifts beyond the final equilibrium resonant frequency were caused by the thermally induced biasing deformation state resulting from the presence of the thin electrode films. It was also shown that in a thermally uncompensated cut the influence of the electrode films is small and the frequency changes monotonically from one equilibrium resonant frequency to the other. Thus, in thermally compensated cuts the influence of small effects such as the presence of very thin electrodes is

significant because the change in frequency of pure thickness vibrations vanishes in the absence of the electrodes for the thermally compensated mode. Consequently, the presence of the thin electrode films and the transverse trapping of energy, which is always present on account of the finite dimensions of the electrodes, have an important influence on the transient temperature dependence of the resonant frequencies of thermally compensated cuts.

In this work the time-dependent, inhomogeneous temperature distribution in the quartz plate is obtained from the uncoupled heat conduction equation for thickness dependence only, subject to the appropriate initial and boundary conditions. Since the heat conduction is sufficiently slow compared to the speed of elastic waves, the mechanical inertia terms can be neglected in the stress equations of motion, thereby reducing them to the quasi-static stress equations of equilibrium. Accordingly, the time-dependent thermally induced biasing deformation state is determined from the exact equations of static linear thermoelasticity for zero thickness electrodes and from the aforementioned system of approximate static thermoelastic plate equations⁶ for finite thickness electrodes. Here, however, the approximate equations are extended to the case of inhomogeneous temperature distributions following earlier work⁹.

In the case of both zero and finite thickness electrodes the substitution of the quasi-static thermoelastic three-dimensional displacement field into the above-mentioned equation for the perturbation of the eigenfrequency along with the solutions for both AT- and SC-cut quartz trapped energy resonators^{10,11} and the first temperature derivatives of the fundamental elastic constants of quartz¹² enables the calculation of the temperature dependence of the resonant frequencies of AT- and SC-cut quartz trapped energy resonators due to the time-dependent thermally induced biasing deformation. The very important influence of the temperature dependence³⁻¹⁵ of the motional capacitive effect of the electrodes on the resonant frequency of thickness vibrations of the electroded portion of the plate, which is a result of the temperature dependence of the pertinent piezoelectric and dielectric constants for the thickness mode of interest in both the AT- and SC-cut quartz plates^{6,16} is included in the treatment.

Calculations are performed for both AT- and SC-cut quartz trapped energy resonators with rectangular gold electrodes for the fundamental and some harmonic overtones. For both cuts the calculations reveal that the magnitude of the transient frequency shift beyond the final equilibrium resonant frequency increases as the orientation of the quartz plate more closely approaches the exact (nominal) zero temperature cut for the particular transverse mode shape of the trapped energy resonator. Also, the more rapid the rate of change of the surface temperature the larger the transient frequency shift. This is in conformity with experimental observation^{7,8}. A realistic 8-shaped input temperature was assumed in all cases.

2. Perturbation Equations

The equation for the perturbation in eigen-frequency¹ mentioned in the Introduction may be written in the form

$$\Delta_\mu = H_\mu / 2\omega_\mu, \quad \omega = \omega_\mu - \Delta_\mu, \quad (2.1)$$

where ω_μ and ω are the unperturbed and perturbed eigen-frequencies, respectively, and

$$H_\mu = - \int [\tilde{K}_{LY}^{\mu} \tilde{g}_{Y,L}^{\mu} + \tilde{K}_L^{\mu} \tilde{F}_L^{\mu}] dV, \quad (2.2)$$

and V is the undeformed volume of the piezoelectric plate at the reference temperature T_0 . In (2.2) \tilde{g}_Y^{μ} and \tilde{F}_L^{μ} denote the normalized mechanical displacement vector and electric potential, respectively, of the μ th eigensolution, and are defined by

$$\tilde{g}_Y^{\mu} = \frac{u_Y^{\mu}}{N_\mu}, \quad \tilde{F}_L^{\mu} = \frac{\phi_L^{\mu}}{N_\mu}, \quad N_\mu^2 = \int \rho u_Y^{\mu} u_Y^{\mu} dV, \quad (2.3)$$

where u_Y^{μ} and ϕ_L^{μ} are the mechanical displacement and electric potential, respectively, which satisfy the equations of linear piezoelectricity

$$\begin{aligned} \tilde{K}_{LY}^{\mu} &= \tilde{g}_{LYM\alpha} u_{\alpha,M}^{\mu} + e_{MLY} \tilde{\phi}_{,M}^{\mu}, \\ \tilde{K}_L^{\mu} &= e_{LMY} u_{Y,M}^{\mu} - \epsilon_{LM} \tilde{\phi}_{,M}^{\mu}, \end{aligned} \quad (2.4)$$

$$\tilde{K}_{LY,L}^{\mu} = \rho \ddot{u}_Y^{\mu}, \quad \tilde{K}_{L,L}^{\mu} = 0, \quad (2.5)$$

subject to the appropriate boundary conditions. The quantities $\tilde{g}_{LYM\alpha}$, e_{MLY} and ϵ_{LM} denote the second-order elastic, piezoelectric and dielectric constants, respectively, and ρ denotes the mass density. Equations (2.4) are the linear piezoelectric constitutive relations and (2.5) 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.1 as is the remainder of the notation in this section. The variables \tilde{K}_{LY}^{μ} and \tilde{K}_L^{μ} in (2.2) denote the portion of the Piola-Kirchhoff stress tensor and reference electric displacement vector, respectively, resulting from the biasing state and a change in the fundamental material constants in the presence of the \tilde{g}_Y^{μ} and \tilde{F}_L^{μ} , and are given by

$$\begin{aligned} \tilde{K}_{LY}^{\mu} &= (\hat{c}_{LYM\alpha} + \Delta c_{LYM\alpha}) \tilde{g}_{\alpha,M}^{\mu} + (\hat{e}_{MLY} + \Delta e_{MLY}) \tilde{F}_{,M}^{\mu}, \\ \tilde{K}_L^{\mu} &= (\hat{e}_{LMY} + \Delta e_{LMY}) \tilde{g}_{Y,M}^{\mu} - (\hat{\epsilon}_{LM} + \Delta \epsilon_{LM}) \tilde{F}_{,M}^{\mu}, \end{aligned} \quad (2.6)$$

where $\hat{c}_{LYM\alpha}$, \hat{e}_{MLY} , and $\hat{\epsilon}_{LM}$ are effective constants that depend on the biasing state¹ and $\Delta c_{LYM\alpha}$, Δe_{MLY} and $\Delta \epsilon_{LM}$ denote small changes in the fundamental elastic, piezoelectric and dielectric constants, respectively, due to a change in temperature.

When nonlinearities due to biasing deformation only are included, we have

$$\begin{aligned} \hat{c}_{LM} &= b_{LMCD} E_{CD}^1 - 2e_o E_{LM}^1, \\ \hat{e}_{LMY} &= - \frac{1}{2} \tilde{g}_{LMYBC} E_{BC}^1 + e_{LMK} v_{Y,K}, \end{aligned}$$

$$\begin{aligned} \hat{c}_{LYM\alpha} &= T_{LM}^1 \delta_{Y\alpha} + \tilde{g}_{LYM\alpha B} E_{AB}^1 + \tilde{g}_{LYM\alpha K} v_{\alpha,K} \\ &+ \tilde{g}_{LM\alpha K} v_{Y,K}, \end{aligned} \quad (2.7)$$

where for a thermoelastic biasing state and relatively small changes in temperature T from the reference temperature T_0

$$\begin{aligned} T_{LM}^1 &= \tilde{g}_{LMKN} E_{KN}^1 - v_{LM} (T - T_0), \\ E_{KN}^1 &= \frac{1}{2} (w_{K,N} + w_{N,K}), \end{aligned} \quad (2.8)$$

$\tilde{g}_{LYM\alpha B}$, b_{LMCD} and $k_{LMY\alpha}$ denote the third-order elastic constants, the electrostrictive constants and the first-order electroelastic constants, respectively, ϵ_o denotes the permeability of free-space, w_K denotes the static biasing displacement field, and v_{LM} denotes the thermoelastic coupling coefficients. Thus, in this description the present position χ is related to the reference position $\tilde{\chi}$ by

$$\chi(\tilde{\chi}_L, t) = \tilde{\chi} + \tilde{w}(\tilde{\chi}_L) + \tilde{u}(\tilde{\chi}_L, t). \quad (2.9)$$

3. Temperature Induced Biasing State

A schematic diagram of the electroded crystal plate is shown in Fig.1 along with the associated coordinate system. The only difference between this section and the analysis presented in Ref.2 and modified somewhat in Ref.6 is that here we have a time-dependent inhomogeneous (in X_2) temperature change symmetric about the centerline of the plate, which we represent in the form

$$T - T_0 = (\bar{T} - T_0) [1 - f(X_2, t)], \quad (3.1)$$

where the temperature T changes relatively little from the homogeneous reference temperature T_0 and \bar{T} represents the final temperature of both surfaces of the electroded quartz plate, and in the previous references the temperature change is time-independent and homogeneous. Furthermore, as noted in the Introduction the mechanical inertia resulting from the time-dependence of the temperature field can be ignored so that the quasi-static stress equations of equilibrium can be employed. In view of the foregoing we simply present the appropriate results from Refs.2 and 6 while incorporating the changes resulting from the inhomogeneous temperature field. Accordingly, it has been shown² that referred to the coordinate system shown in Fig.1 the static (here quasi-static) purely extensional thermoelastic plate equations for the plated crystal plate may be written in the form

$$\chi_{AB,A}^{(0)} = 0, \quad \chi_{AB,A}^{(2)} = 0, \quad (3.2)$$

where A, B, C, D take the values 1 and 3 and skip 2 and $\chi_{AB}^{(0)}$ and $\chi_{AB}^{(2)}$ are the zero and second order plate stress resultants, respectively, for the electroded crystal plate. For symmetric inhomogeneous temperature states and identical electrodes on the upper and lower surfaces, by following the procedure in Secs.III and IV of Ref.2 we find that the extensional constitutive equations may be written in the form

$$\begin{aligned} \chi_{AB}^{(0)} &= 2h (v_{ABCD} + \frac{2h'}{h} v'_{ABCD}) E_{CD}^{(0)} + 2h^3 \left(\frac{1}{3} v_{ABCD} + \right. \\ &\left. \frac{2h'}{h} v'_{ABCD} \right) E_{CD}^{(2)} - 2h (s_{AB}^{(0)} + \frac{2h'}{h} s_{AB}^{(2)}) (\bar{T} - T_0), \end{aligned}$$

$$\chi_{AB}^{(2)} = \frac{2}{3} h^3 \left(\gamma_{ABCD} + \frac{6h'}{h} \gamma'_{ABCD} \right) E_{CD}^{(0)} + \frac{2}{5} h^5 \left(\gamma_{ABCD} + \frac{10h'}{h} \gamma'_{ABCD} \right) E_{CD}^{(2)} - \frac{2}{3} h^3 \left(\beta_{AB}^{(2)} + \frac{6h'}{h} \beta_{AB}^{*'} \right) (\bar{T} - T_0), \quad (3.3)$$

where \bar{T} is the final electrode temperature,

$$\beta_R^{(0)} = \beta_R (1 - \rho^{(0)} / 2h), \quad \beta_R^{(2)} = \beta_R (1 - 3\rho^{(2)} / 2h^3), \quad (3.4)$$

and

$$\beta_R = \nu_R - c_{RW} c_{WV}^{-1} \nu_V, \quad \gamma_{RS} = c_{RS} - c_{RW} c_{WV}^{-1} c_{VS},$$

$$\nu_R^{*'} = \nu_R' - \nu_2' c_{2R}' / c_{22}', \quad (3.5)$$

$$\rho^{(0)} = \int_{-h}^h f(x_2, t) dx_2, \quad \rho^{(2)} = \int_{-h}^h x_2^2 f(x_2, t) dx_2,$$

$$\hat{\nu}_R^{*'} = \nu_R^{*'} [1 - f(h, t)], \quad (3.6)$$

and we have introduced the compressed notation⁶ for tensor indices according to the scheme

$$R, S = 1, 3, 5; \quad W, V = 2, 4, 6. \quad (3.7)$$

The $\gamma_{\alpha\beta}$ are Voigt's anisotropic plate elastic constants, the β_{α} are the associated plate thermoelastic constants and the $\beta_{\alpha}^{*'}$ are the effective anisotropic plate thermoelastic constants. For the case of anisotropic extension considered here the three-dimensional strains $E_{\alpha\beta}$, which are needed in the perturbation equation, are related to the plate strains by^{2,6}

$$E_{KL} = \frac{1}{2} (w_{K,L} + w_{L,K}) = E_{KL}^{(0)} + x_2^2 E_{KL}^{(2)}. \quad (3.8)$$

The plate strains $E_{\alpha\beta}^{(n)}$ ($n=0, 2$), which occur in (3.3), are given by

$$E_{AB}^{(n)} = \frac{1}{2} (w_{A,B}^{(n)} + w_{B,A}^{(n)}), \quad (3.9)$$

and the remaining plate strains, which are needed in (3.8) as well, may be obtained from

$$E_W^{(0)} = -c_{WV}^{-1} c_{VS} E_S^{(0)} + c_{WV}^{-1} \nu_V \left(\frac{15}{8h^3} \rho^{(0)} - \frac{45}{5} \rho^{(2)} \right) (\bar{T} - T_0),$$

$$E_W^{(2)} = -c_{WV}^{-1} c_{VS} E_S^{(2)} + c_{WV}^{-1} \nu_V \left(1 - \frac{9\rho^{(0)}}{8h} + \frac{15\rho^{(2)}}{8h^3} \right) (\bar{T} - T_0), \quad (3.10)$$

which have also been used in obtaining Eqs.(3.3).

In Fig.1, the x_2 -coordinate axis is normal to the major surfaces of the plate at $T = T_0$. Since the outside edges of the plate are traction free, we have

$$N_A \chi_{AB}^{(0)} = 0, \quad N_A \chi_{AB}^{(2)} = 0 \quad \text{on outside edges}, \quad (3.11)$$

where N_A denotes the outwardly directed unit normal to the edge of the plate at $T = T_0$. The solution satisfying (3.2) and (3.11) takes the form

$$\chi_{AB}^{(0)} = 0, \quad \chi_{AB}^{(2)} = 0. \quad (3.12)$$

This solution is unique to within homogeneous plate rotations⁷ of zero and second order. Substituting from (3.3) into (3.12), we obtain

$$\left(\gamma_{RS} + \frac{2h'}{h} \gamma'_{RS} \right) E_S^{(0)} + h^2 \left(\frac{\gamma_{RS}}{3} + \frac{2h'}{h} \gamma'_{RS} \right) E_S^{(2)} = \left(\beta_R^{(0)} + \frac{2h'}{h} \beta_R^{*'} \right) (\bar{T} - T_0),$$

$$\left(\gamma_{RS} + \frac{6h'}{h} \gamma'_{RS} \right) E_S^{(0)} + h^2 \left(\frac{3}{5} \gamma_{RS} + \frac{6h'}{h} \gamma'_{RS} \right) E_S^{(2)} = \left(\beta_R^{(2)} + \frac{6h'}{h} \beta_R^{*'} \right) (\bar{T} - T_0). \quad (3.13)$$

Equations (3.13) constitute six inhomogeneous linear equations which may readily be solved for the six plate strains $E_{\alpha}^{(0)}$ and $E_{\alpha}^{(2)}$. When $E_{\alpha}^{(0)}$ and $E_{\alpha}^{(2)}$ have been determined from (3.13), $E_{\alpha\beta}^{(0)}$ and $E_{\alpha\beta}^{(2)}$ are readily determined from (3.10). Then the three-dimensional biasing strain can be obtained from (3.8), which is for anisotropic extension. There is no flexure because the resonator has identical electrodes which are subject to the same instantaneous temperature change.

It is clear from (2.7) that the biasing three-dimensional strain field which has just been determined is inadequate for the calculation of the change in frequency because the three-dimensional displacement gradients $w_{\alpha,\beta}$ are needed. It has been shown in Secs.III and IV of Ref.6 that the three-dimensional small rotation field

$$\Omega_{KL} = \frac{1}{2} (w_{L,K} - w_{K,L}) \quad (3.14)$$

may be obtained by requiring that the three-dimensional rotation gradient-strain gradient relations

$$\Omega_{KL,M} = E_{ML,K} - E_{MK,L} \quad (3.15)$$

be satisfied. In this way since the change in frequency due to homogeneous rigid rotations vanishes^{1,8} it has been shown⁶ that

$$w_{A,B} = E_{AB}^{(0)} + x_2^2 E_{AB}^{(2)}, \quad w_{2,A} = E_{2A}^{(0)},$$

$$w_{A,2} = E_{2A}^{(0)} + 2x_2 E_{2A}^{(2)}, \quad w_{2,2} = E_{22}^{(0)} + x_2^2 E_{22}^{(2)}. \quad (3.16)$$

Thus, once the plate strains $E_{\alpha\beta}^{(0)}$ and $E_{\alpha\beta}^{(2)}$ have been found from (3.13), the biasing displacement gradients $w_{\alpha,\beta}$ may be obtained as known linear functions of the change in surface temperature $(\bar{T} - T_0)$. Thus, we now obtain $\hat{c}_{\alpha\gamma\alpha\beta}$ in (2.7)₃ as a known linear function of $(\bar{T} - T_0)$ for the electroded region of the plate.

In the unelectroded region of the plate the situation is somewhat simpler and since, as noted in the Introduction, the mechanical inertia can be neglected in the stress equations of motion, we have the quasi-static stress equations of equilibrium

$$T_{ML,M}^1 = 0, \quad (3.17)$$

along with the constitutive equations (2.4)₁. Since the outside edges of the plate are traction free along with the major surfaces, we have

$$N_M T_{ML}^1 = 0, \quad (3.18)$$

where N_n denotes the unit normal to all surfaces of the plate at $T = T_0$. From (3.17) and (3.18) we have

$$T_{ML}^1 = 0, \quad (3.19)$$

for the unelectroded plate. Since the thermoelastic coupling constants v_{ML} are related to the coefficients of linear expansion α_{JK} by the usual relation

$$v_{ML} = c_{MLJK} \alpha_{JK}, \quad (3.20)$$

from (2.8)₁, (3.19), (3.20), and (3.1) we have

$$E_{JK}^1 = \alpha_{JK} (\bar{T} - T_0) [1 - f(X_2, t)]. \quad (3.21)$$

In the unelectroded region of the plate the time-dependent inhomogeneous strain state is given by (3.21) and we note that when the quasi-static stress equations of equilibrium, (3.17), are satisfied trivially, as is the case here, the vanishing of the stresses, (3.19), results in time-dependent inhomogeneous strains E_{KL}^1 and consistent rotations Ω_{KL}^1 , which are determined from (3.15). Since all variables are functions of the X_2 -coordinate only and it has been shown¹⁸ that the change in frequency due to a homogeneous infinitesimal rigid rotation vanishes, from (3.15) we obtain

$$\Omega_{13}^1 = 0, \quad \Omega_{21}^1 = E_{21}^1, \quad \Omega_{23}^1 = E_{23}^1. \quad (3.22)$$

From (3.8), (3.14) and (3.22), we obtain

$$w_{A,B} = E_{AB}^1, \quad w_{A,2} = 2E_{2A}^1, \quad w_{2,A} = 0, \quad w_{2,2} = E_{22}^1, \quad (3.23)$$

which with (3.21) yields the displacement gradients w_{KL} as known linear functions of the change in surface temperature of the unelectroded portion of the plate. The substitution of (3.23) with (3.21) in (2.7)₃ yields \hat{c}_{LYM} as a known linear function of $(T - T_0)$ for the unelectroded plate.

4. Temperature Dependence

The temperature behavior of the plate shown in Fig. 1 is obtained by satisfying the uncoupled one-dimensional heat conduction equation

$$k_{22} \frac{\partial^2 T}{\partial X_2^2} = c_s \frac{\partial T}{\partial t}, \quad (4.1)$$

subject to appropriate initial and boundary conditions at $X_2 = \pm h$. In (4.1) k_{22} is the thermal conductivity in the X_2 -direction and c_s is the specific heat. If the surfaces are subject to an S-shaped rise in temperature $(T - T_0)$, the initial and boundary conditions are

$$T - T_0 = 0 \quad \text{at } t = 0, \quad (4.2)$$

$$T - T_0 = (\bar{T} - T_0) [1 - (1 + \beta t) e^{-\beta t}] \quad \text{at } X_2 = \pm h. \quad (4.3)$$

The solution to (4.1) subject to conditions (4.2) and (4.3) can be written in the form¹⁹

$$T - T_0 = (\bar{T} - T_0) \left[1 - (1 + \beta t) e^{-\beta t} \frac{\cos \sqrt{\beta/k_{22}} X_2}{\cos \sqrt{\beta/k_{22}} h} - \frac{2\beta}{h} \sum_{n=0}^{\infty} \frac{(-1)^n \cos \zeta_n X_2}{(\kappa_{22} \zeta_n^2 - \beta)^2} \left(\frac{\beta}{\zeta_n^2} e^{-\zeta_n^2 \kappa_{22}^2 t} - \zeta_n \kappa_{22} e^{-\beta t} \right) \right], \quad (4.4)$$

where

$$\kappa_{22} = k_{22}/c_s, \quad \zeta_n = (2n+1)\pi/2h, \quad (4.5)$$

and we note that the value of β determines the sharpness of the temperature rise at the surfaces. In view of (3.1) and (3.6) it is clear that (4.4) enables the calculation²⁰ of $\hat{g}^{(0)}$, $\hat{g}^{(2)}$ and \hat{w}^* in the electroded region of the plate. In the unelectroded region of the plate, Eqs. (4.4), with (3.1) can readily be used in (3.21) and (3.23) to obtain the displacement gradients w_{KL} as known functions of X_2 and t and the given surface temperature data.

5. Eigenmodes in Trapped Energy Resonators

A plan view of one quadrant of the trapped energy resonator is shown in Fig. 2. It has been shown²⁰ that the solution functions for the trapped energy eigenmodes can be written in the form

$$\begin{aligned} \bar{u}_1 &= \bar{B} \sin \frac{n\pi X_2}{2h} \cos \bar{X}_1 \cos \bar{X}_3, \\ u_1^S &= \bar{B}^S \sin \frac{n\pi X_2}{2h} e^{-\xi^S(X_1-l)} \cos \bar{X}_3, \\ u_1^T &= \bar{B}^T \sin \frac{n\pi X_2}{2h} \cos \bar{X}_1 e^{-\nu^T(X_3-b)}, \\ u_1^C &= \bar{B}^C \sin \frac{n\pi X_2}{2h} e^{-\xi^S(X_1-l)} e^{-\nu^T(X_3-b)}, \end{aligned} \quad (5.1)$$

where

$$\bar{B}^S = \bar{B} \cos \bar{\xi} l, \quad \bar{B}^T = \bar{B} \cos \bar{\nu} b, \quad \bar{B}^C = \bar{B} \cos \bar{\xi} l \cos \bar{\nu} b, \quad (5.2)$$

and $\bar{\xi}$ and $\bar{\nu}$ are determined from the roots of the transcendental equations^{10,11,20}

$$\bar{\xi} \tan \bar{\xi} l = \left(\frac{\gamma_n}{M_n} - \bar{\xi}^2 \right)^{1/2}, \quad \bar{\nu} \tan \bar{\nu} b = \left(\frac{\gamma_n}{P_n} - \bar{\nu}^2 \right)^{1/2}. \quad (5.3)$$

The decay numbers ξ^S and ν^T are given by

$$\xi^S = \left(\frac{\gamma_n}{M_n} - \bar{\xi}^2 \right)^{1/2}, \quad \nu^T = \left(\frac{\gamma_n}{P_n} - \bar{\nu}^2 \right)^{1/2}, \quad (5.4)$$

and for SC-cut quartz the coefficients M_n , P_n and γ_n are given by¹¹

$$\begin{aligned} M_n &= c_{11} + \frac{2c_{16}^2}{\bar{c}(1)} + r_2(c_{12} + c_{66}) \\ &\quad + \frac{4(r_2 \bar{c}^{(2)} + c_{12})(r_2 \bar{c}^{(1)} - c_{66})}{n\pi \bar{c}^{(2)} \kappa_2} \cot \kappa_2 \frac{n\pi}{2}, \\ P_n &= c_{55} + r_3(c_{36} + c_{45}) \\ &\quad + \frac{4(r_3 \bar{c}^{(3)} + c_{45})(r_3 \bar{c}^{(1)} - c_{36})}{n\pi \bar{c}^{(3)} \kappa_3} \cot \kappa_3 \frac{n\pi}{2}, \end{aligned} \quad (5.5)$$

$$\gamma_n = \hat{c}^{(1)} \frac{n^2 \pi^2}{2h^2} \left[\frac{4k_{26}^2}{n^2 \pi^2} + R \right], \quad (5.6)$$

where

$$r_2 = \frac{c_{12} + c_{66}}{\bar{c}(1) - \bar{c}(2)}, \quad r_3 = \frac{c_{36} + c_{45}}{\bar{c}(1) - \bar{c}(3)},$$

$$\begin{aligned} \kappa_2 &= \sqrt{\frac{\bar{c}(1)}{\bar{c}(2)}}, \quad \kappa_3 = \sqrt{\frac{\bar{c}(1)}{\bar{c}(3)}}, \\ k_{26}^2 &= \frac{e_{26}^2}{\bar{c}(1) \epsilon_{22}}, \quad R = \frac{2\rho' h'}{\rho h}, \\ \hat{c}(1) &= \bar{c}(1) \left(1 - \frac{8k_{26}^2}{n\pi} - 2R\right). \end{aligned} \quad (5.7)$$

In (5.5) - (5.7) the material constants are referred to the transformed coordinates²⁰ associated with the eigenvector triad for waves in the thickness (X_3)-direction as discussed in Sec. 2 of Ref. 23 and the $\bar{c}^{(a)}$ are the eigenvalues for the three waves. The resonant frequency for the trapped energy mode in (5.1) is given by²⁰

$$\omega = \frac{n\pi}{2h} \sqrt{\frac{\hat{c}(1)}{\rho}} \left[1 + \frac{1}{2} \frac{1}{n\pi} \left(\frac{M}{\hat{c}(1)} 4h^2 \xi^2 + \frac{P}{\hat{c}(1)} 4h^2 \nu^2 \right)^{1/2} \right]. \quad (5.8)$$

Since the AT-cut can be considered to be a special case of the doubly-rotated SC-cut, the foregoing trapped energy eigensolution for the SC-cut can be used²⁰ for the simpler AT-cut. We now have the u_1 -displacement field for the trapped energy mode in either orientation.

As noted in Ref. 15 in addition to the $u_1(\bar{u}_1)$ displacement field there are accompanying $u_2(\bar{u}_2)$ and $u_3(\bar{u}_3)$ displacement fields, and to the same order in $\xi(\bar{\xi})$ there are accompanying corrections to the $u_1(\bar{u}_1)$ displacement fields, which are given by¹¹

$$\begin{aligned} \bar{u}_1 &= \bar{B} \left(1 + \frac{c_{16}}{\bar{c}(1)} \bar{\xi} X_2 \right) \sin \frac{n\pi X_2}{2h} \cos \bar{\xi} X_1 \cos \bar{\kappa} X_3, \\ \bar{u}_2 &= \left[\frac{r_2 \bar{\xi} 2h}{n\pi} \bar{B} \cos \frac{n\pi X_2}{2h} + \bar{C} \cos \frac{\kappa_2 n\pi X_2}{2h} \right] \sin \bar{\xi} X_1 \cos \bar{\kappa} X_3, \\ \bar{u}_3 &= \left[\frac{r_3 \bar{\nu} 2h}{n\pi} \bar{B} \cos \frac{n\pi X_2}{2h} + \bar{E} \cos \frac{\kappa_3 n\pi X_2}{2h} \right] \cos \bar{\xi} X_1 \sin \bar{\kappa} X_3, \end{aligned} \quad (5.9)$$

where

$$\begin{aligned} \bar{C} &= (-1)^{\frac{n+1}{2}} \frac{(r_2 \bar{\xi}^{(2)} + c_{12}) \bar{\xi} 2h \bar{B}}{\bar{c}^{(2)} \kappa_2 n\pi \sin \kappa_2 n\pi/2}, \\ \bar{E} &= (-1)^{\frac{n+1}{2}} \frac{(r_3 \bar{\nu}^{(3)} + c_{45}) \bar{\nu} 2h \bar{B}}{\bar{c}^{(3)} \kappa_3 n\pi \sin \kappa_3 n\pi/2}, \end{aligned} \quad (5.10)$$

and

$$\begin{aligned} u_1^S &= B^S \left(1 + \frac{c_{16}}{\bar{c}(1)} \bar{\xi}^S X_2 \right) \sin \frac{n\pi X_2}{2h} e^{-\bar{\xi}^S (X_1 - a_1)} \cos \bar{\kappa} X_3, \\ u_2^S &= \left[\frac{r_2 \bar{\xi}^S 2h}{n\pi} B^S \cos \frac{n\pi X_2}{2h} + C^S \cos \frac{\kappa_2 n\pi X_2}{2h} \right] e^{-\bar{\xi}^S (X_1 - a_1)} \cos \bar{\kappa} X_3, \end{aligned}$$

$$u_3^S = \left[\frac{r_3 \bar{\nu} 2h}{n\pi} B^S \cos \frac{n\pi X_2}{2h} + E^S \cos \frac{\kappa_3 n\pi X_2}{2h} \right] e^{-\bar{\xi}^S (X_1 - a_1)} \sin \bar{\kappa} X_3. \quad (5.11)$$

where

$$\begin{aligned} C^S &= (-1)^{\frac{n+1}{2}} \frac{(r_2 \bar{\xi}^{(2)} + c_{12}) \bar{\xi}^S 2h B^S}{\bar{c}^{(2)} \kappa_2 n\pi \sin \kappa_2 n\pi/2}, \\ E^S &= (-1)^{\frac{n+1}{2}} \frac{(r_3 \bar{\nu}^{(3)} + c_{45}) \bar{\nu} 2h B^S}{\bar{c}^{(3)} \kappa_3 n\pi \sin \kappa_3 n\pi/2}, \end{aligned} \quad (5.12)$$

with similar expressions²⁰ for u_2^S and u_3^S . As noted earlier in this section this solution is referred to the eigenvector triad of the pure thickness solution for the SC-cut¹¹. For purposes of calculation of the temperature dependence of the resonant frequency it is advisable²⁰ to transform back to the original conventional coordinate axes for the SC-cut, thus

$$\tilde{u}_\alpha = Q_{\alpha\beta} u_\beta, \quad (5.13)$$

where \tilde{u}_α denotes the components of the mechanical displacement in the original coordinate system for the SC-cut and $Q_{\alpha\beta}$ denotes the transformation from that system to the eigenvector triad¹¹. For the AT-cut this is not necessary.

6. Temperature Dependence of Resonant Frequency

The change in the resonant frequency with temperature of any electroded contoured SC-cut quartz plate resulting from the thermally induced biasing may now be determined from (2.1) and (2.2) with (2.6) and (2.7). However, Eq. (2.2) cannot be used for calculation as it appears because the temperature derivatives of the complete piezoelectric and dielectric tensors are not presently known, nor are all the fundamental coefficients appearing in (2.7) known. Nevertheless, as in earlier work⁶, since the piezoelectric coupling and wavenumbers along the plate are both small, the temperature dependence of only the transformed^{11,20} thickness piezoelectric and dielectric constants e_{26} and c_{22} need be retained in H_1 . Moreover, since $(1/e_{26})de_{26}/dT \gg (1/e_{22})de_{22}/dT$, we ignore $(1/e_{22})de_{22}/dT$. Furthermore $(1/e_{26})de_{26}/dT$ can and should be excluded from wave terms in H_1 because the existing temperature derivatives of the fundamental elastic constants of quartz effectively contain the small influence of the temperature dependence of the piezoelectric and dielectric constants, which results from the piezoelectric stiffening of the waves. In addition, we ignore the $\hat{e}_{n1\gamma}$ since they are not known. Then the $(1/e_{26})de_{26}/dT$ that we retain is not fundamental but effective⁸. In view of the foregoing, the general electroelastic perturbation integral in (2.2) with (2.6) and (2.7) may be written in the reduced form⁸

$$\begin{aligned} H_\mu &= - \int_V \left[(\hat{c}_{LYM\gamma} + \Delta c_{LYM\gamma}) \tilde{\kappa}_\alpha^\mu \tilde{\kappa}_\gamma^\mu L \right. \\ &\quad \left. - \frac{2e_{26}^2}{\epsilon_{22}} \frac{\Delta e_{26}}{e_{26}} \tilde{u}_{1,2} \frac{g_1(\omega)}{h} \right] dV, \end{aligned} \quad (6.1)$$

where

$$\Delta = (T - T_0) d/dT. \quad (6.2)$$

In (6.1) the first term under the integral sign is decomposed in the original conventional coordinate system for the SC-cut because the $\hat{c}_{LYM\alpha}$ and $\Delta\hat{c}_{LYM\alpha}$ are known in that coordinate system, while the second term is decomposed along the eigenvector triad¹¹ of the pure thickness solution²⁰. From (6.1) for the geometry shown in Fig.2, we obtain

$$R_{\mu} = -4 \int_{-h}^h dx_2 \left[\int_0^L dx_1 \left(\int_0^b (\bar{\theta} + \bar{c}) dx_3 + \int_b^{\infty} \bar{\theta} dx_3 \right) + \int_0^L dx_1 \left(\int_0^b \bar{\theta} dx_3 + \int_b^{\infty} \bar{\theta} dx_3 \right) \right], \quad (6.3)$$

where

$$\bar{\theta} = (\hat{c}_{LYM\alpha} + \Delta\hat{c}_{LYM\alpha}) \tilde{g}_{\alpha, M}^{\mu} \tilde{g}_{V, L}^{\mu},$$

$$\bar{c} = -2 \frac{e_{26}^2}{e_{22}} \frac{\Delta e_{26}}{e_{26}} \tilde{g}_{1, 2}^{\mu} \frac{\tilde{g}_1^{\mu}(h)}{h}. \quad (6.4)$$

The $\hat{c}_{LYM\alpha}$ in (6.4) are known as space and time dependent linear expressions in $(T - T_0)$ from the analyses in Secs.3 and 4 and the space and time dependent change in the elastic constants with temperature $\Delta\hat{c}_{LYM\alpha}$ are given by

$$\Delta\hat{c}_{LYM\alpha} = (dc_{LYM\alpha}/dT)(T - T_0), \quad (6.5)$$

where the $dc_{LYM\alpha}/dT$ are obtained from the first temperature derivatives of the fundamental elastic constants of quartz²³ dc_{ij}/dT referred to the principal axes by the tensor transformation relation

$$\frac{d}{dT} \hat{c}_{LYM\alpha} = a_{LD}^a a_{YE}^a a_{MF}^a a_{OG}^a \frac{d}{dT} \tilde{c}_{DEFG}, \quad (6.6)$$

where the a_{ij} are the matrix of direction cosines for the transformation from the principal axes to the coordinate system containing the axes referred to the electroded plate. When the conventional IEEE notation²¹ for doubly-rotated plates is written in the form $(Y, X, w, L)\phi, \theta$, where $\phi = 0$, the rotation angles ϕ and θ are the first two Euler angles, from which the a_{ij} can be determined. Clearly, the transformation relations for the second and third order elastic, piezoelectric and dielectric constants, and coefficients of linear expansion may be written in the respective forms

$$\hat{\xi}_{KLMN} = a_{KD}^a a_{LE}^a a_{MF}^a a_{NG}^a \tilde{\xi}_{DEFG},$$

$$\hat{\xi}_{KLMNAB} = a_{KD}^a a_{LE}^a a_{MF}^a a_{NG}^a a_{AH}^a a_{BI}^a \tilde{\xi}_{DEFGHI},$$

$$e_{KLM} = a_{KD}^a a_{LE}^a a_{MF}^a \tilde{e}_{DEF},$$

$$e_{KL} = a_{KD}^a a_{LE}^a \tilde{e}_{MN}, \quad \alpha_{KL} = a_{KD}^a a_{LE}^a \tilde{\alpha}_{MN}, \quad (6.7)$$

where the tensor quantities with the upper cycle are referred to the principal axes of the crystal.

The temperature dependence of the effective piezoelectric constants for the thickness mode of interest in both the AT and SC cuts of quartz have been estimated^{2,15} from data provided by Lukaszek and are given by

$$(1/e_{26}) de_{26}/dT = 1.56 \times 10^{-4}/^{\circ}K, \text{ AT},$$

$$-4.8 \times 10^{-4}/^{\circ}K, \text{ SC}. \quad (6.8)$$

Calculations have been performed using the known values of the second order elastic, piezoelectric and dielectric constants of quartz²³, the third order elastic²⁴ and thermoelastic²⁴ constants of quartz and the recently obtained first temperature derivatives of the fundamental elastic constants of quartz²³ along with the estimates in (6.8). The results of the calculations are presented in Figs.3-10, each of which shows the relative change in frequency with time for a temperature input value of β of 10 sec^{-1} except for Fig.10 which is for $\beta = 1/100 \text{ sec}^{-1}$. All calculations were for a trapped energy resonator 1.7 mm thick with square gold electrodes 2000 Å thick (although sometimes of zero thickness) and an L/h ratio of 20. Figure 3 shows the calculated change in frequency with temperature of an AT-cut (nominal angle $\theta = -35.25^{\circ}$) compensated to only 1 PPM for the first, third and fifth harmonic trapped energy modes. The figure shows that the change in frequency is considerably smaller for the harmonics than the fundamental, which exhibits a slight initial transient shift in the opposite direction before beginning its monotonic change to the final equilibrium value. The behavior of the harmonics is monotonic throughout. Figure 4 is for an SC-cut which is also compensated to only 1 PPM. The figure shows that the harmonics exhibit a slight initial transient shift in the opposite direction while the fundamental changes monotonically. Figure 5 shows a comparison of the behavior of the fundamental trapped energy mode for both the AT and SC cuts shown in Figs.3 and 4, respectively. The figure shows that the changes in frequency to the new equilibrium values are opposite in sign and not that different in magnitude and that the AT-cut exhibits more severe transient behavior. Figure 6 shows the change in frequency with temperature of an AT cut compensated to 10 PPT (parts per trillion) for the fundamental trapped energy mode. The figure shows rather severe transient shifts in frequency in each direction for both the 2000 Å thick gold electrodes, for which the compensation was achieved and electrodes of zero thickness. Note that the transient frequency excursions are larger for the more compensated case. Figure 7 is for an SC-cut which is also compensated to 10 PPT. The figure shows rather severe transient shifts in frequency as in the case of the AT cut, but now in only one direction, for both electrode thicknesses considered. Note that again the transient frequency excursion is larger for the more compensated case. Figure 8 shows the change in frequency due to the presence of 2000 Å thick gold electrodes²⁵ for which the actual changes for the AT cut are shown in Fig.3. Note that the change in frequency due to the films is about two and three orders of magnitude smaller than the actual change shown in Fig.3 and that for the AT cut the changes due to the electrode films are monotonic. Since a portion of the aging rate is a result of the relaxation of residual stress in the electrodes, the long time asymptotic limit of the change in frequency shown in Fig.8 is the portion of the actual change in frequency that contributes to the aging rate. Figure 9 shows the change in frequency due to the presence of 2000 Å thick gold electrodes for which the actual changes for the SC-cut are shown in Fig.4. Note that the change in frequency due to the films is about five and six orders of magnitude smaller than the actual change shown in Fig.4 and that for the SC cut they are not monotonic. Since the long time asymptotic limit is associated with the aging rate, Figs.8 and 9 indicate that the aging due to electrode stress should be between two and three orders of magnitude smaller for the SC cut than for the AT cut. Figure 10 shows the actual change in frequency with temperature of the AT and SC cuts, both of which are compensated to 10 PPT, for a relatively slow rate of change of surface temperature corresponding to $\beta = 1/100 \text{ sec}^{-1}$. The figure shows that the transient frequency excursions are about two orders

of magnitude smaller than in the case of the rapid temperature input corresponding to $\beta = 10 \text{ sec}^{-1}$, which are shown in Figs. 6 and 7. Figure 10 also shows that the transient frequency shifts for the AT and SC cuts are of the same order of magnitude, although the shift for the AT cut is about twice as large as for the SC cut. The results of the calculations presented here clearly indicate that the change in frequency of the trapped energy modes due to electrode stress is more than two orders of magnitude smaller in the SC than in the AT cut and that the transient frequency excursions are of the same order of magnitude in both cuts.

Acknowledgements

This work was supported in part by the Army Research Office under Contract No. DAAG 29-82-K-0130 and the National Science Foundation under Grant No. MEA-8115340.

References

1. H.F. Tiersten, "Perturbation Theory for Linear Electroelastic Equations for Small Fields Superposed on a Bias," *J. Acoust. Soc. Am.*, **64**, 832 (1978).
2. 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).
3. 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.
4. R.D. Mindlin, "High Frequency Vibrations of Crystal Plates," *Quart. Appl. Math.*, **19**, 51 (1961).
5. H.F. Tiersten, *Linear Piezoelectric Plate Vibrations* (Plenum, New York, 1969), Chap. 13.
6. 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**, 1704 (1983).
7. A.W. Warner, "Use of Parallel-Field Excitation in the Design of Quartz Crystal Units," *Proceedings of the 17th Annual Frequency Control Symposium*, U.S. Army Electronics Research and Development Laboratory, Fort Monmouth, New Jersey, 248 (1963).
8. A. Ballato and J. Vig, "Static and Dynamic Frequency-Temperature Behavior of Singly and Doubly-Rotated, Oven-Controlled Quartz Resonators," *Proceedings of the 32nd Annual Symposium on Frequency Control*, U.S. Army Electronics Research and Development Command, Fort Monmouth, New Jersey, 180 (1978).
9. B.K. Sinha and H.F. Tiersten, "Transient Thermally Induced Frequency Excursions in Doubly-Rotated Quartz Thickness-Mode Resonators," *Proceedings of the 34th Annual Symposium on Frequency Control*, U.S. Army Electronics Research and Development Command, Fort Monmouth, New Jersey, 393 (1980).
10. H.F. Tiersten, "Analysis of Trapped Energy Resonators Operating in Overtones of Coupled Thickness-Shear and Thickness-Twist," *J. Acoust. Soc. Am.*, **52**, 879 (1976).
11. H.F. Tiersten and D.S. Stevens, "An Analysis of Contoured SC-Cut Quartz Crystal Resonators," *Proceedings of the 36th Annual Symposium on Frequency Control*, U.S. Army Electronics Research and Development Command, Fort Monmouth, New Jersey, 37 (1982).
12. B.K. Sinha and H.F. Tiersten, "First Temperature Derivatives of the Fundamental Elastic Constants of Quartz," *J. Appl. Phys.*, **50**, 2732 (1979).
13. M. Onoe, "Relationship Between Temperature Behavior of Resonant and Antiresonant Frequencies and Electromechanical Coupling Factors of Piezoelectric Resonators," *Proc. IEEE*, **57**, 702 (1969).
14. A. Ballato, "Apparent Orientation Shifts of Mass Loaded Plate Vibrators," *Proc. IEEE*, **64**, 1449 (1976).
15. D.S. Stevens and H.F. Tiersten, "Temperature-Induced Frequency Changes in Electroded Contoured SC-Cut Quartz Crystal Resonators," *Proceedings of the 36th Annual Symposium on Frequency Control*, U.S. Army Electronics Research and Development Command, Fort Monmouth, New Jersey, 46 (1982).
16. Ref. 5, Chap. 7, Sec. 1.
17. This is shown by the static version of the Uniqueness Theorem obtained in Ref. 4, which applies in this quasi-static situation being treated here.
18. This is shown in the Appendix of Ref. 2, which applies to the quasi-static case considered here.
19. H.S. Carslaw and J.C. Jaeger, *Conduction of Heat in Solids* (Oxford University Press, 1959), p. 104.
20. For more detail see D.S. Stevens and H.F. Tiersten, "Thermally Generated Transient Frequency Excursions in AT- and SC-Cut Quartz Trapped Energy Resonators," to be issued as a technical report, Rensselaer Polytechnic Institute, Troy, New York.
21. IEEE Standard on Piezoelectricity - IEEE STD. 176-1978.
22. R. Bechmann, "Elastic and Piezoelectric Constants of Alpha-Quartz," *Phys. Rev.*, **110**, 1060 (1958).
23. R.N. Thurston, H.J. McSkimin and P. Andreatch, Jr., "Third Order Elastic Constants of Quartz," *J. Appl. Phys.*, **37**, 267 (1966).
24. F. Kohlrausch, *Lehrbuch der prakt. Physik*, 16. Aufl. 5. 158 (1930). Constants employed in Ref. 25.
25. R. Bechmann, A.D. Ballato and T.N. Lukaszek, "Higher Order Temperature Coefficients of the Elastic Stiffnesses and Compliances of Alpha-Quartz," *Proc. IRE*, **50**, 1812 (1962).
26. The change in frequency due to the presence of the electrodes is obtained by subtracting the calculated change without the electrodes from the calculated change with the electrodes.

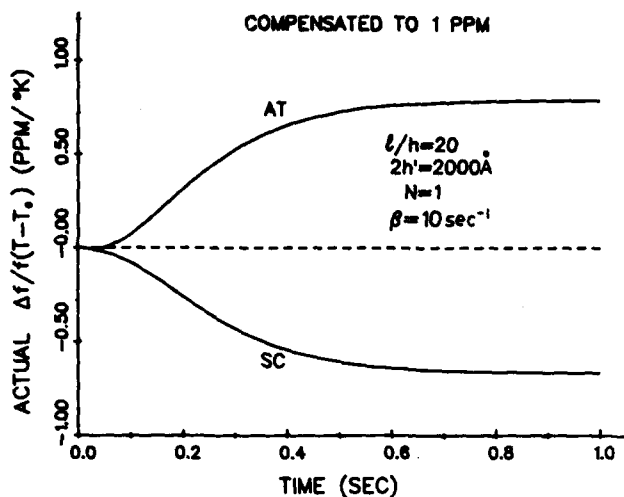


Figure 5 Comparison of Actual Time-Dependent Frequency Shifts per $^{\circ}\text{K}$ for the Fundamental Mode of AT- and SC-Cut Quartz Trapped Energy Resonators Compensated to 1 PPM

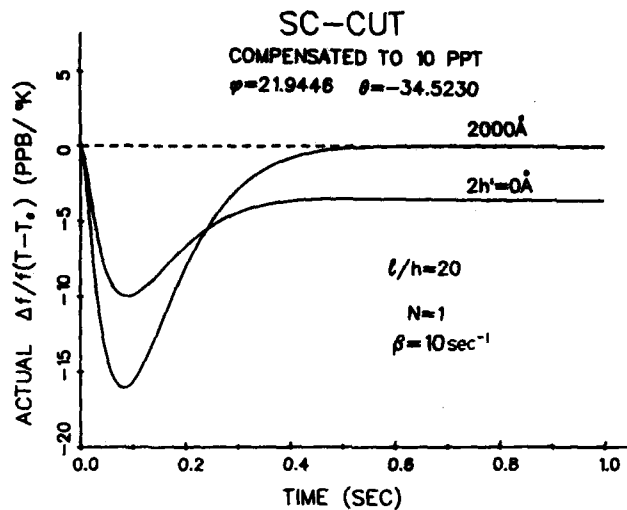


Figure 7 Actual Time-Dependent Frequency Shift per $^{\circ}\text{K}$ for SC-Cut Quartz Trapped Energy Resonator Compensated to 10 PPT

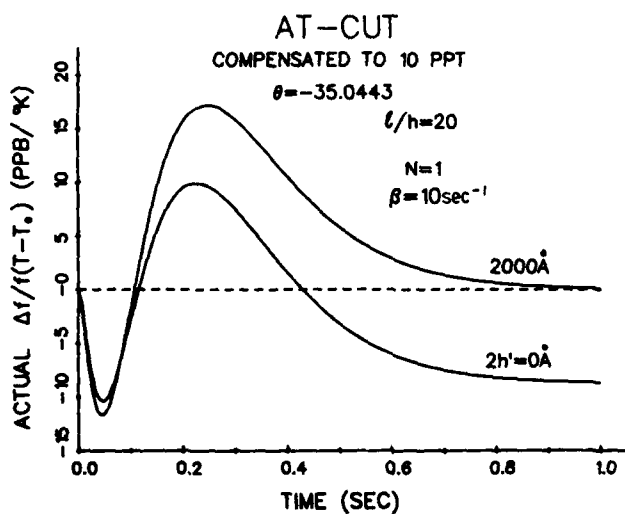


Figure 6 Actual Time-Dependent Frequency Shift per $^{\circ}\text{K}$ for AT-Cut Quartz Trapped Energy Resonator Compensated to 10 PPT

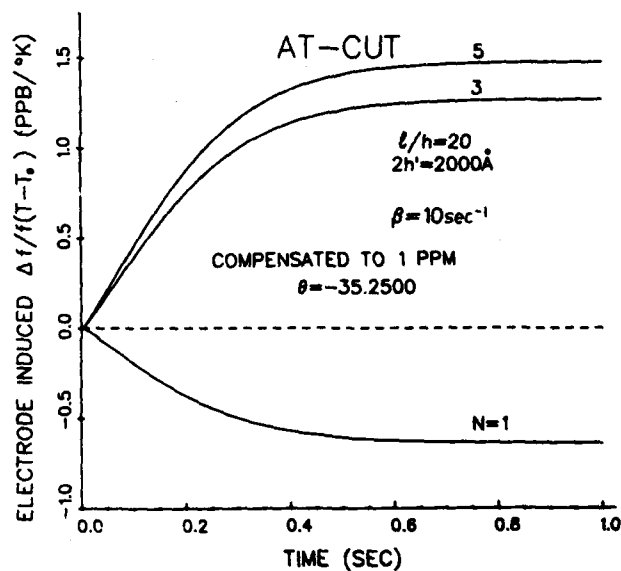


Figure 8 Time-Dependent Frequency Shift per $^{\circ}\text{K}$ due to 2000 Å Thick Gold Electrodes for AT-Cut Quartz Trapped Energy Resonator Compensated to 1 PPM

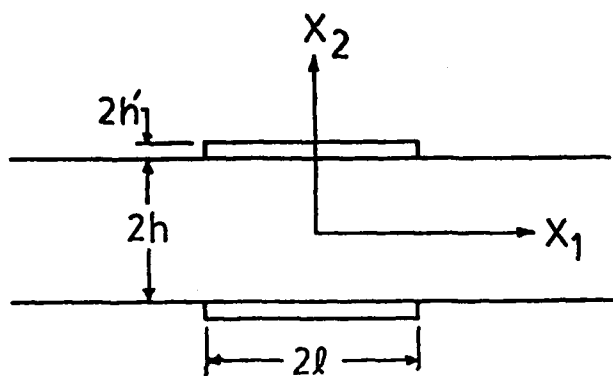


Figure 1 Cross-section of the Trapped Energy Resonator

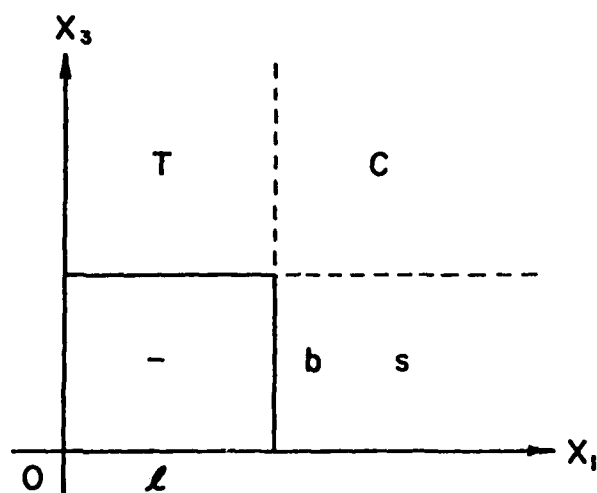


Figure 2 Plan View Showing One Quadrant of the Trapped Energy Resonator

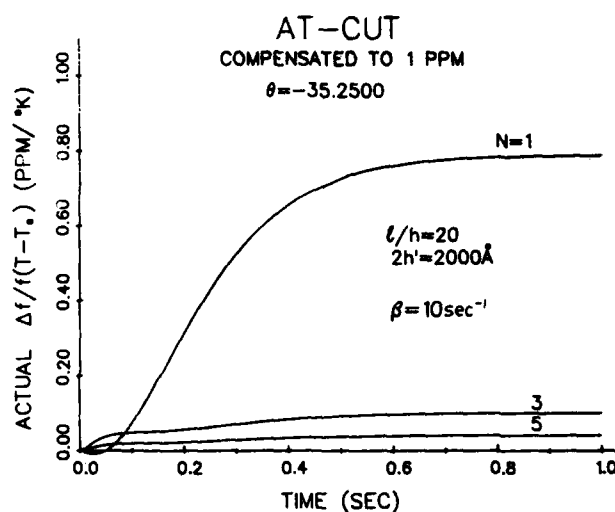


Figure 3 Actual Time-Dependent Frequency Shift per °K for AT-Cut Quartz Trapped Energy Resonator Compensated to 1 PPM. Quartz plates are 1.7 mm thick with square gold electrodes either 2000 Å thick or of zero thickness and 34 mm on a side for each figure. Figures 3-9 are for a rapid temperature input to the surface.

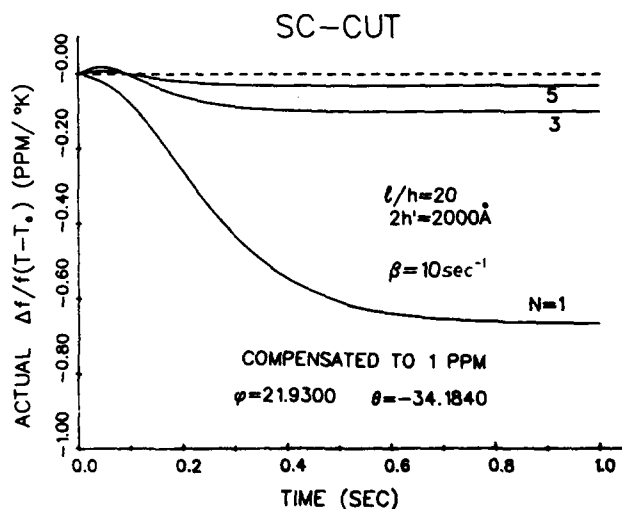


Figure 4 Actual Time-Dependent Frequency Shift per °K for SC-Cut Quartz Trapped Energy Resonator Compensated to 1 PPM

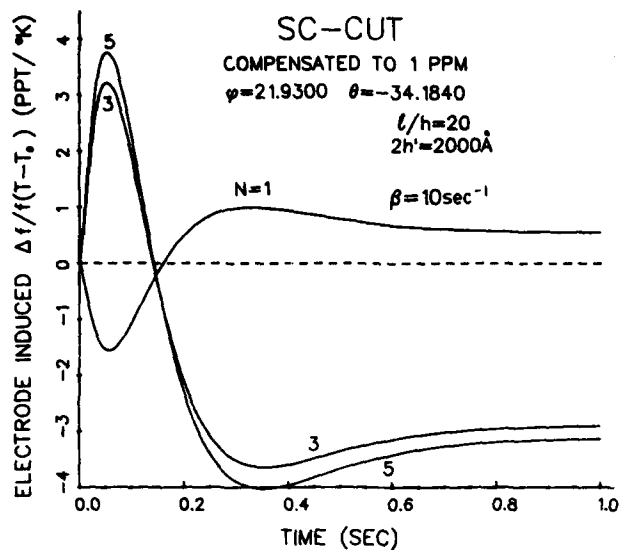


Figure 9 Time-Dependent Frequency Shift per $^{\circ}\text{K}$ due to 2000 \AA Thick Gold Electrodes for SC-Cut Quartz Trapped Energy Resonator Compensated to 1 PPM

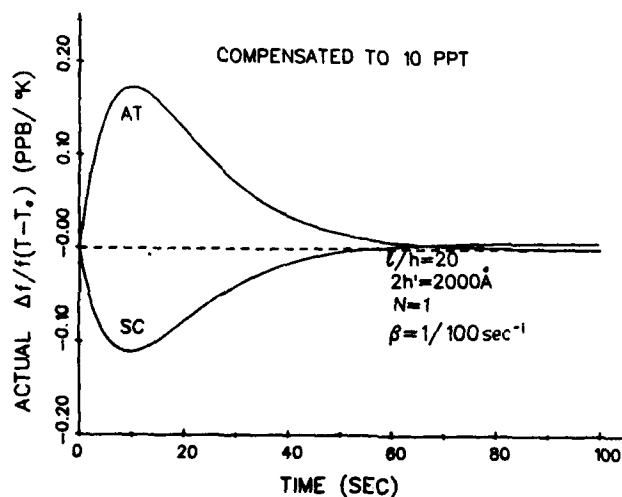


Figure 10 Comparison of Actual Time-Dependent Frequency Shifts per $^{\circ}\text{K}$ for the Fundamental Mode of AT- and SC-Cut Quartz Trapped Energy Resonators Compensated to 10 PPT for a Relatively Slow Temperature Input to the Surface

AD P 002479

EXCESS NOISE IN QUARTZ CRYSTAL RESONATORS

J. J. Gagnepain and M. Olivier

Laboratoire de Physique et Métrologie des Oscillateurs du CNRS
associé à l'Université de Franche-Comté Besançon
32, Avenue de l'Observatoire - 25000 Besançon - France

and

F. L. Walls

National Bureau of Standards - Time and Frequency Division
Boulder, Colorado 80303

Summary

Frequency and phase noise in quartz crystal resonators are studied as a function of the driving power. At low power, where the crystal behaves linearly, $1/f$ fluctuations of the resonance frequency are observed. At medium power the nonlinearities of the crystal significantly increase the phase fluctuations at low Fourier frequencies. At high power, thermal instabilities and chaotic behavior occur characterized by the generation of high level white noise. ←

Introduction

The short and medium term frequency stabilities of quartz crystal oscillators are limited by additive and parametric noises of the electronics and by parametric noise of the quartz resonator.¹ In an oscillator it is not always possible to distinguish between these different contributions and thereby determine the level of frequency fluctuations of the resonator. A significant advance in solving this problem came with the introduction of a new technique to test quartz crystals independently of the oscillator electronics.² This enabled measurements of flicker and random walk frequency noises of various pairs of crystals at different frequencies and of different origins.^{2,3} A second advance on the same problem came with the intro-

duction of the concept of a dual crystal system, in which a quartz oscillator with good spectral purity is phase locked to a passive quartz resonator.⁴ In this approach, the frequency stability in medium and long term is dominated by the passive quartz resonator and hence its stability becomes observable. Such measurement systems can be used for comparing quartz crystal instabilities and for trying to understand the origins and the mechanisms of these frequency fluctuations.

The previous measurements^{2,3} have shown that the resonance frequency of quartz resonators exhibits flicker ($1/f$) fluctuations and random walk ($1/f^2$) fluctuations. Random walk frequency noise has been shown to be correlated to temperature fluctuations.^{3,5} The mechanisms contributing to the generation of $1/f$ noise are not well understood yet.

When the resonator is excited at⁹ a moderately high drive level a large increase of the noise level is observed at low Fourier frequencies. It will be shown that this excess noise is due to the nonlinearities of the crystal. At very high drive levels, the resonator can exhibit chaotic behavior producing high level white noise.

Measurements at Low Power

The measurement system, presented for the first time in ref. 2, will be briefly described. As schematically shown in Fig. 1, two identical resonators (or two that adjusted to be nearly identical) are driven at their resonance frequency by a frequency source with high spectral purity. When the two resonators are at the same frequency (frequencies are tuned with adjustable capacitors) and with equal Q-factors (Q-factors are balanced with additional resistors), the frequency fluctuations of the source are rejected by 50 to 60 dB. The double balanced mixer (DBM) operated at its quadrature point by means of the 90° phase shifter detects the phase fluctuations $d\phi$ induced by the frequency fluctuations of each resonator, following the relation

$$d\phi = -2Q \frac{d\omega_1 - d\omega_2}{\omega_0} \quad (1)$$

where Q is the resonator loaded Q-factor and $d\omega_1(d\omega_2)$ is the instantaneous frequency difference of the resonance frequency of resonator 1 (2) from the driving frequency ω_0 . Equation (1) is valid for Fourier frequencies within the resonator's bandwidth and for low drive powers (typically 10 μ W were used). Outside this bandwidth it must be corrected to take into account the filtering effect of the resonators. The mixer output is amplified and processed by a spectrum analyzer, yielding the power spectral density $S_v(f)$ of the output voltage noise. This is related to the frequency spectrum of the resonator $S_{y_0}(f)$,

$$S_v(f) = (G\mu)^2 S_{y_0}(f) = (2G\mu Q)^2 S_{y_0}(f), \quad (2)$$

where μ is the mixer sensitivity, and G is the low noise amplifier gain. The corresponding resonance frequency fluctuations of the resonators are deduced from relation (1) (again for Fourier frequencies within the resonator's bandwidth). Each resonator contributes to the total noise, which is measured. If the two resonators are identical the noise amplitude of each of them is obtained by dividing by $\sqrt{2}$.

Great care must be taken to minimize the spurious influence of temperature changes. The measurements reported below were obtained using a digitally controlled oven. Temperature was measured with a LC-cut quartz sensor. The temperature control loop had a 10^{-4} K resolution. Additional thermal filtering, with copper mass and thermal discontinuities, reduced the residual temperature fluctuations of the crystals under test down to a few μ K over a few sec.⁶ The two resonators were in the same oven and thermally coupled in order to reduce random frequency walk noise. This does not perturb 1/f noise measurements, because 1/f noise is not correlated with temperature fluctuations.^{3,5} Over a day, temperature stability was of the order of 5×10^{-5} K. The resonators generally were operated at their turn-over temperature. This measurement system was used to test 5 MHz, 10 MHz, and 2.5 MHz resonators.

1) 5 MHz resonators. Six AT-cut crystals (fifth overtone) were studied. Two were commercial BVA⁷ resonators and four were standard commercial high quality resonators. A typical spectrum (pair of BVA crystals) is shown in Fig. 2. The noise is essentially 1/f frequency noise. The $1/f^3$ noise is due to the filtering effect of the resonator; this occurs at its half bandwidth. For three pairs of resonators the frequency noise, $S_{y_0}(f)$, measured 1 Hz from the carrier was equal to $2.5 \times 10^{-13}/\sqrt{\text{Hz}}$, $2.5 \times 10^{-13}/\sqrt{\text{Hz}}$ and $3.10^{-13}/\sqrt{\text{Hz}}$. The raw data were divided by $\sqrt{2}$ under the assumption that the noise of each resonator in a given pair was identical. Obviously by doing all possible pairs one can obtain a more accurate value for each resonator independent of this assumption. All these resonators had unloaded Q-factors in the range $2.5 \times 10^6 < Q_0 < 2.7 \times 10^6$.

2) 10 MHz resonators. Two kinds were measured: AT-cut and BT-cut crystals. Both were third overtone resonators, but with very different Q-factors, typically 3.10^5 for AT-cut and 1.7×10^6 for BT-cut. BT-cut crystals have much higher Q-factors because this crystallographic orientation has lower internal losses for the thickness-shear mode.⁸ Noise spectra are shown

in Fig. 2. A large difference in the noise levels can be observed between the AT and BT cuts. At 1 Hz from the carrier these levels are of the order of 7×10^{-13} /Hz for the BT resonators to 3.2×10^{-11} /Hz for the AT resonators. These results show a strong dependence between the $1/f$ noise level and Q-factor.

3) 2.5 MHz resonators. Two pairs of fifth overtone, AT-cut resonators were tested with unloaded Q-factors close to 4×10^6 . Although these pairs had almost the same Q-factors, the noise levels differed by almost one order of magnitude. The presence of $1/f^2$ noise in the spectra of $S_{y_0}(f)$ indicates that the sensitivity to temperature fluctuations are very large for this type of resonator. The $1/f$ noise power 1 Hz from the carrier, evaluated for each individual crystal, was plotted as a function of the unloaded Q-factors in Fig. 3. A linear regression among these experimental points gives

$$S_{y_0}(1 \text{ Hz}) \approx \frac{2}{Q} \quad (3)$$

These results are also summarized in Table I.

The data show a clear dependence of $1/f$ noise on the resonator's unloaded Q-factors, following a $1/Q^4$ law. The only exception occurs with the 2.5 MHz (#2) crystals, which show excessive noise most likely due to thermal transient effects.

If these crystals are used in an oscillator, the $1/f$ spectrum will give a flicker floor in time domain, whose corresponding values are given in Table I, and which corresponds to the best stability achievable in that case with an oscillator.

Measurements at Medium Power

When the drive level of the crystal is increased it exhibits nonlinear effects due mainly to the higher order elastic constants. This is the well known amplitude frequency effect where distortions appear in the amplitude and phase resonance curves and even hysteresis

can be developed.⁹

The nonlinear behavior of a resonator, driven in transmission can be represented by the phenomenological relation,

$$\frac{d^2 i}{dt^2} + \frac{\omega_0}{Q} \frac{di}{dt} + \omega_0^2 [1 - 2\epsilon(\Omega) \cos \Omega t] i (1 + k i^2) = F \cos \omega t, \quad (4)$$

where i is the current through the crystal, Q the loaded Q-factor, and F the amplitude of the driving force. k is the nonlinear coefficient (related to the next higher order elastic constants). $\epsilon(\Omega) \cos \Omega t$ introduces a modulation of the resonance angular frequency ω_0 , which represents the frequency noise of the resonator. Solving this equation with a perturbation method gives the phase noise of the output signal.

Let $S_{y_0}(\Omega)$ be the frequency noise spectrum of the crystal. When driving it at low power, in the linear range, the corresponding phase spectrum is given by

$$S_\phi(\Omega) = S_{y_0}(\Omega) \frac{\omega_0^2}{\Omega^2 + \omega_0^2/4Q^2} \quad (5)$$

At medium levels, when the resonator is driven near the jump frequency, the phase spectrum becomes

$$S_\phi(\Omega) = S_{y_0}(\Omega) \frac{\omega_0^2(\Omega^2 + \omega_0^2/4Q^2)}{\Omega^2(\Omega^2 + \omega_0^2/Q^2)} \quad (6)$$

Thus the ratio between the phase noises at medium and low powers is

$$\frac{S_\phi(\text{medium power})}{S_\phi(\text{low power})} = \frac{\Omega^2 + \omega_0^2/4Q^2}{\Omega^2(\Omega^2 + \omega_0^2/Q^2)} \quad (7)$$

This ratio goes to unity for $\Omega \gg \omega_0/Q$ and is equal to $\omega_0^2/16Q^2\Omega^2$ for $\Omega \ll \omega_0/4Q$. Therefore the induced phase noise for the lower Fourier frequency components can be greatly increased by the crystal nonlinearities. Such a noise was

experimentally observed on a 5 MHz (5th overtone AT-cut) resonator driven at 2.5 mW, as shown on Fig. 5.

It should be noted that higher drive levels were once considered as a means to improve the (white phase noise) signal/noise ratio and therefore the short term stability in oscillators. However, our results indicate this is not in general a practical solution since the nonlinear response of the resonator increases the close in noise at the same time.

Measurements at High Power

For still higher levels, quartz resonators exhibit large instabilities. Such phenomena are known in nonlinear systems as chaotic behavior. Chaos has been observed in many different systems¹⁰, for example phase locked loops¹¹ and Josephson junctions^{12,13} which can be considered as low Q-factor resonators. The main difference between these systems and quartz crystal resonators stem from the high Q-factor and thermal effects, which strongly modify the behavior of the crystal at high power.

The dissipated power induces large temperature rises. Positive or negative frequency shifts follow according to the sign of temperature coefficient of the resonator, i.e. to the location of the operating temperature on the frequency-temperature characteristics.

Near room temperature, the temperature coefficient is negative (of the order - 4 Hz/K). A temperature rise will induce a negative frequency shift, which can be large enough to pull the crystal to the frequency where the down jump phenomenon occurs (Fig. 4). Thus the amplitude becomes much smaller decreasing the dissipated power. Temperature therefore decreases causing the frequency to increase until it reaches the second jump and so on. This gives a cycling with large amplitude and phase perturbations, as shown in Fig. 6. This phenomenon occurs when the amplitude-frequency effect and

the temperature coefficient have opposite signs.

Above the turn-over temperature the sign of the temperature coefficient becomes positive. In this case the temperature rises, inducing a positive frequency change, which corresponds to an additional amplitude-frequency effect. This is equivalent to increasing the nonlinearities of the crystal. Fig. 7 illustrates the amplitude resonance curve of a 5 MHz (AT-cut, fifth overtone) crystal excited with a 7 V rms driving signal. As the drive frequency increases the crystal goes from stable state to chaotic states. The transition is sudden, and no set of cascading subharmonic bifurcations is observed, as would be the case in low-Q resonators. The absence of subharmonics is explained by the high Q of the quartz crystal, which filters all the subharmonic frequencies, which are out of its linewidth.

The sideband frequency noise of the transmitted signal was measured in the same conditions. In Fig. 8 the power spectral density of frequency fluctuations shows a f^2 dependence versus Fourier frequency. This corresponds to a white phase spectrum, characteristic of the generation of white noise in the crystal resonator by the chaotic behavior (white noise was also observed in other chaotic systems.¹⁰⁻¹³ The level of this phase noise is $S_{\phi} = -34$ dB (rad^2/Hz), which is about 80 dB larger than the phase noise levels of Fig. 5.

Conclusion

The noise behavior of a quartz crystal resonator is strongly influenced by the amplitude of the oscillation and the corresponding nonlinearities. At low power $1/f$ frequency noise is observed. At medium power the nonlinear response of the crystal can amplify this noise and give larger phase fluctuations. At high power, thermal effects become an additional source of instability. At sufficiently high drive power chaotic states take place which

generate very large white phase noise. The problem of chaos is of most interest however, from the basic physics point of view. The theory of cascading bifurcations is a good explanation of how a deterministic system can have stochastic solutions although it does not directly apply in the case of quartz crystal, because of the high Q-factors.

Acknowledgements

The authors wish to thank L. Bidart from CEPE (France), C. Stone from Brightline Corp. and J. Vig from US Army ERADCOM for providing some of the crystals used in this work. They are also grateful to G. Marianneau and J. Gros Lambert from LPMO for their help in the noise measurements.

References

1. L. S. Cutler and C. L. Searle, "Some aspects of the theory and measurement of frequency fluctuations in frequency standards", Proc. IEEE, 54, 2, p. 136 (1966).
2. F. L. Walls and A. E. Wainwright, "Measurement of the short-term stability of quartz crystal resonators and the implications for quartz oscillator design and applications, IEEE Trans. Instrum. Meas., IM-24, p. 15 (1975).
3. J. J. Gagnepain, "Fundamental noise studies of quartz crystal resonators", Proc. 30th Ann. Freq. Cont. Symp., p. 84 (1976).
4. S. R. Stein, C. M. Maney, Jr., F. L. Walls, and J. E. Gray, "A systems approach to high performance oscillators", Proc. 32nd Ann. Freq. Cont. Symp., p. 484 (1981).
5. Y. Noguchi, Y. Teramachi, F. Musha, "1/f fluctuations of a quartz crystal oscillator", Proc. 35th Ann. Freq. Cont. Symp., p. 484 (1981).
6. G. Marianneau, J. J. Gagnepain, "Digital temperature control for ultrastable quartz oscillators", Proc. 34th Ann. Freq. Cont. Symp., p. 52 (1980).
7. R. J. Besson, "A new piezoelectric resonator design", Proc. 30th Ann. Freq. Cont. Symp., p. 78 (1976) "A new electrodeless resonator design", Proc. 31st Ann. Freq. Cont. Symp., p. 147 (1977).
8. J. Lamb and J. Richter, "Anisotropic acoustic attenuation with new measurements for quartz at room temperatures", Proc. Roy. Soc. A, vol. 293, p. 479 (1966).
9. J. J. Gagnepain and R. Besson, "Nonlinear effects in piezoelectric quartz crystals", Physical Acoustics, vol. XI, p. 245, W. P. Mason, R. N. Thurston Eds, Academic Press (1975).
10. "Evolution of order and chaos", Proc. Int. Symp. on Synergetics, H. Haken Ed. Springer-Verlag (1982).
11. D. D'Humieres, M. Bearley, B. Huberman, A. Libchaber, "Chaotic states and routes to chaos in the forced pendulum", Ginzton Lab. Report 33429, Stanford University (1982).
12. B. A. Hubertson, J. P. Cratchfield, "Chaotic states of anharmonic systems in periodic fields", Phys. Rev. Letters, 43, n° 23, p. 1743 (1979).
13. N. Petersen, A. Davidson, "Chaos and noise rise in Josephson junctions", Appl. Phys. Lett., 39, n° 10, p. 830 (1981).

resonator	2.5 MHz ● 1	2.5 MHz ● 2	5 MHz BVA	5 MHz	5 MHz	10 MHz	10 MHz	10 MHz	10 MHz	10 MHz
crystal cut	AT	AT	AT	AT	AT	AT	AT	BT	BT	BT
Q_0	$4 \cdot 10^6$	$4 \cdot 10^6$	$2.6 \cdot 10^6$	$2.7 \cdot 10^6$	$2.4 \cdot 10^6$	$2.5 \cdot 10^6$	$3 \cdot 10^5$	$1.5 \cdot 10^6$	$1.8 \cdot 10^6$	$1.9 \cdot 10^6$
S_y (1 Hz)	$2.5 \cdot 10^{-26}$	$2.4 \cdot 10^{-26}$	$6.2 \cdot 10^{-26}$	$5.8 \cdot 10^{-26}$	$9 \cdot 10^{-26}$	$1 \cdot 10^{-21}$	$6.5 \cdot 10^{-22}$	$2.6 \cdot 10^{-24}$	$2 \cdot 10^{-24}$	$5 \cdot 10^{-28}$
flicker floor	$1.9 \cdot 10^{-13}$	$1.5 \cdot 10^{-12}$	$2.9 \cdot 10^{-13}$	$2.8 \cdot 10^{-13}$	$3.5 \cdot 10^{-13}$	$3.7 \cdot 10^{-11}$	$3 \cdot 10^{-11}$	$1.9 \cdot 10^{-12}$	$1.7 \cdot 10^{-12}$	$6.3 \cdot 10^{-13}$

Table I

1/f noise level at 1 Hz from the carrier, and corresponding flicker floor, for the different samples (2.5 MHz ● 2 crystals were not taken into account in the linear regression of Fig. 3)

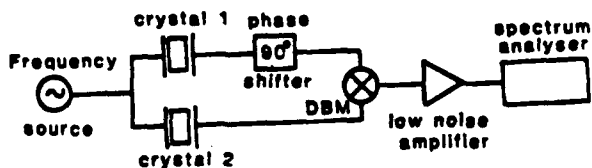


Fig. 1 : Passive frequency noise measurement system :

Relative phase fluctuations induced by the two resonator's resonance frequency fluctuations are measured with a phase bridge composed of a 90° phase shifter and double balanced mixer (DBM)

Mixer sensitivity : 0.15 V/rad at 5 MHz

Amplifier gain : 500

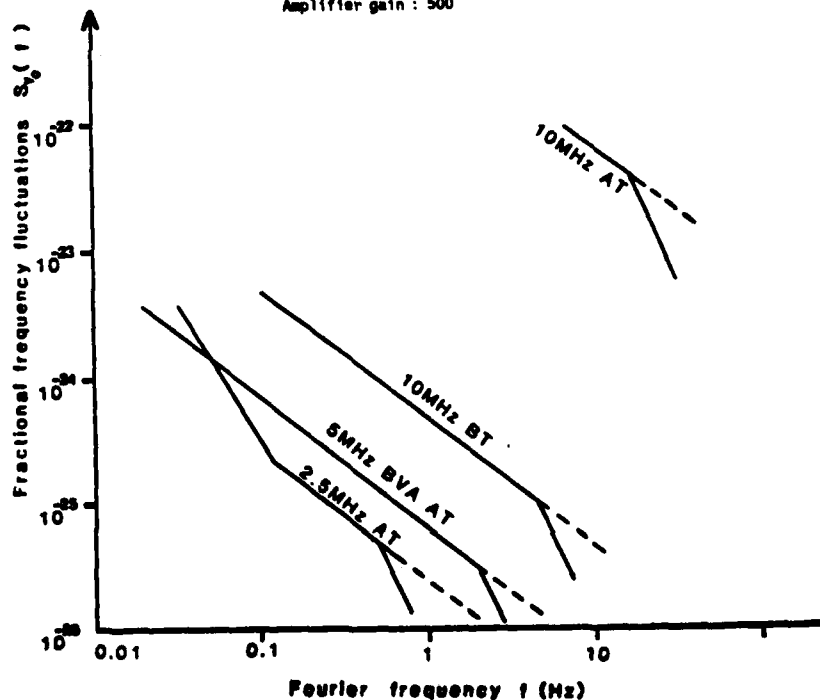


Fig. 2 The solid curve shows the apparent frequency noise spectrum of 2.5 MHz, 5 MHz, and 10 MHz resonators. At Fourier frequencies above the half bandwidth, the spectrum must be corrected for the filtering effect of the resonators yielding the dashed lines.

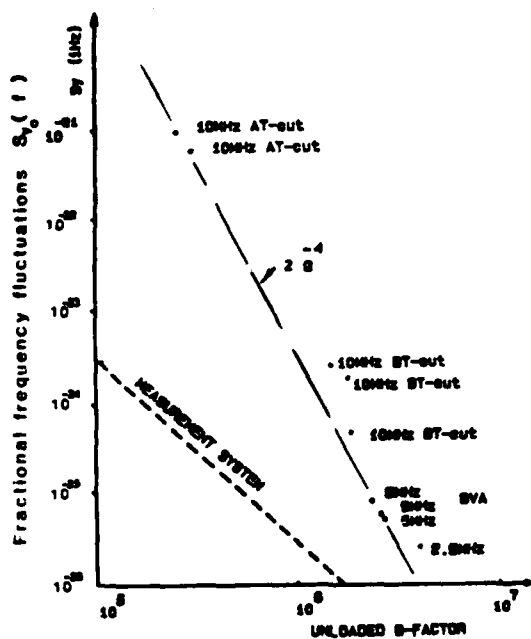


Fig. 3. $1/f$ noise level, measured at 1 Hz from the carrier as a function of the unloaded Q-factor for the different resonators tested. The measurement system noise is indicated by ----.

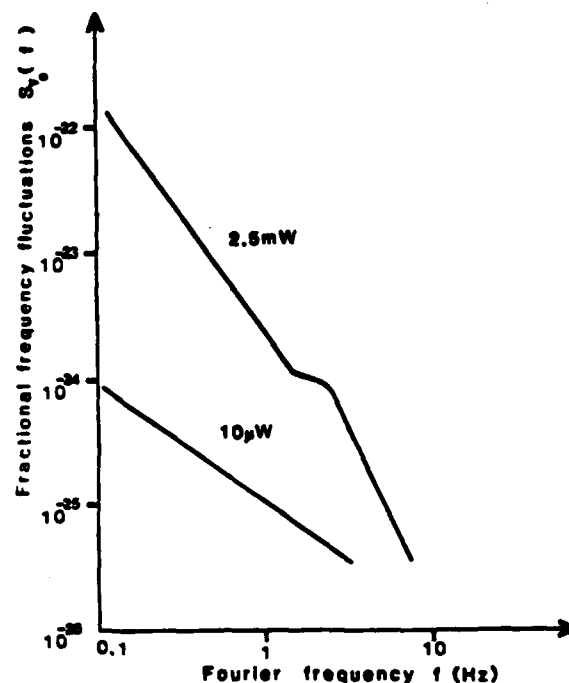


Fig. 5. Frequency noise spectrum of a 5 MHz driven at low and medium power.

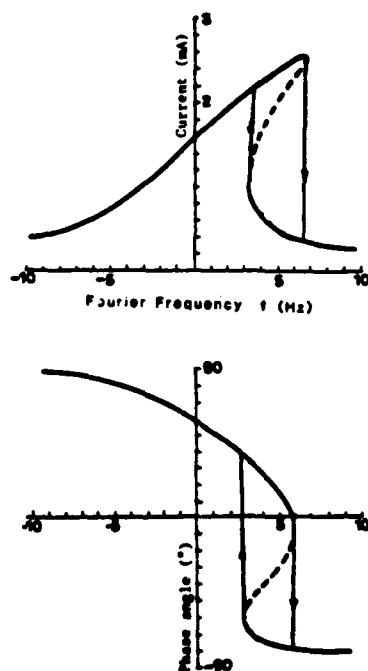


Fig. 4. Amplitude and phase resonance curves of quartz resonators

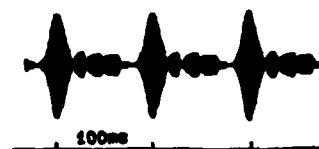


Fig. 6. Amplitude perturbation due to thermal effects at high power.

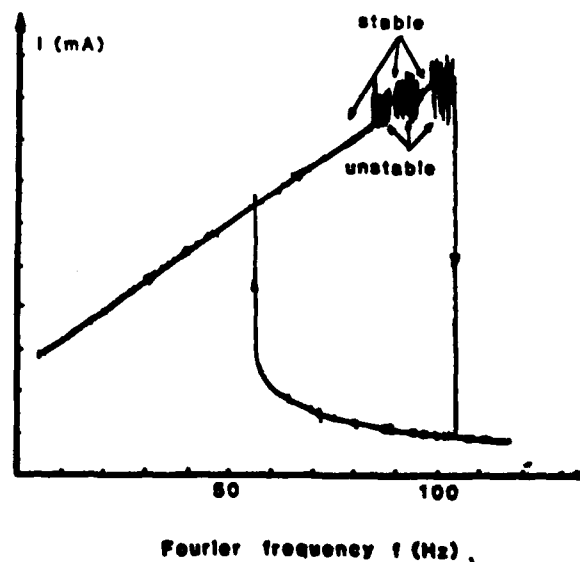


Fig. 7. Stable and unstable states of a 5 MHz resonator driven at high power.

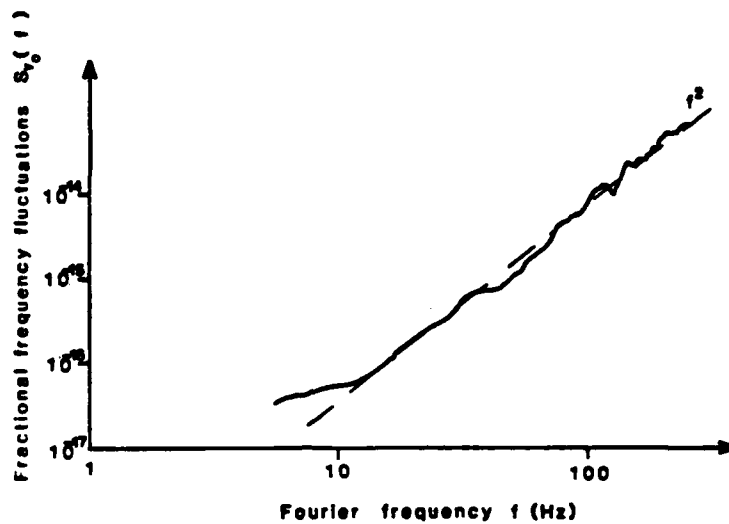
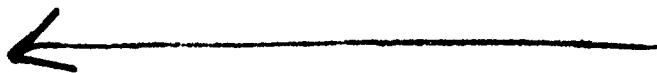


Fig. 8. Frequency fluctuations of a resonator driven at 7 VRMS illustrating the chaotic state of the crystal.



THICKNESS MODES IN CIRCULAR AT-CUT
QUARTZ PLATES WITH CIRCULAR ELECTRODES

Dr. W. D. Beaver
COMTEC ECONOMATION INC
3300 E Sparrow Avenue, Flagstaff, Arizona 86001

AD P 002480
ABSTRACT

Over the past years, several papers have been presented which contain theoretical analyses of the coupled thickness-shear and thickness-twist mode resonances in rectangular and circular quartz plates. These studies involved the solution of the differential equations with certain assumed boundary conditions and the determination of the eigenfrequencies related to the thickness modes using computers.

However, to this date, the complexities of the mathematics have prevented a similar analysis of the commonly manufactured resonator consisting of a circular AT-cut quartz blank with circular electrode on its major surfaces. The purpose of this study was to derive useful equations which can be employed in designing such AT-cut quartz resonators.

This paper describes the results of an empirical analysis to determine how the frequency and the relative activity of the thickness modes in AT-cut quartz resonators depends on the diameter of the circular electrodes and the mass loading.

Specifically provided in the paper are:

1. an equation between the dimensions of the effective motional volume of a quartz plate and the electrode diameter and mass loading;
2. an equation for the frequencies of the inharmonic thickness shear and twist modes as a function of the mass loading and resonator dimensions;
3. a graphical presentation of the interrelationship between the relative amplitude of the thickness shear and thickness twist modes and the product of the frequency and the motional inductance of quartz resonators.

From analysis of the empirical data, the lengths of the major, λ_x , and minor axes, λ_z , are given respectively by equations which express them as functions of the diameter of the electrode, the thickness of the quartz plate, and the mass loading.

INTRODUCTION

The purpose of this paper is to share the results of an experimental analysis of thickness shear modes in AT-cut quartz plates. The objective of the analysis was to explain the feasibility of using different zones inside the boundary of an electroded region as separate controllable resonances. In the process sufficient data was derived to establish an equation for the frequencies of inharmonic thickness shear modes in resonators having circular quartz plates with circular electrodes and to gain an understanding of their relative activity levels. Most narrow-band filter requirements can be satisfied utilizing monolithic crystal filter configurations. However, intermediate band width and wide-band width filter requirements cannot be satisfied using the monolithic crystal filter approach and must be realized using

conventional filter configurations which employ discrete quartz crystal resonators. It is because of these types of filters that an understanding of the nature of unwanted inharmonic thickness shear modes is vitally important. It is important to be able to reasonably accurately predict the frequencies and relative amplitude of these modes.

What this paper will provide is an equation which will allow the user to predict the inharmonic mode frequencies as a function of the critical dimensions of the crystal resonator. The designer may then design the resonator to stagger the frequencies and minimize the effects of coincidence. Also provided is experimental data related to the relative amplitude of the inharmonic modes in resonators designed to meet various inductance requirements. The designer may utilize this data to determine the trade off of impedance levels and mode suppression. The paper will also introduce what we call the Shared Electrode Resonator SER and describe some of its characteristics.

INHARMONIC MODE FREQUENCIES

The equation¹ which we will employ to compute the inharmonic mode frequencies is shown in Figure 1. We assume that essentially all of the acoustic energy is trapped within the bounds of the resonator. This equation gives the frequencies of a rectangular resonator where F_{mnp} is the resonant frequency of the mode having $.5m$ wavelengths along the dimension T , $.5n$ wavelengths along the dimension X , and $.5p$ wavelengths Z with m , n and p being integers. We will modify the equation to fit the empirical data, so that we can apply it to resonator designs with acceptable accuracy. For an AT-cut quartz plate the equation shown in Figure 1 involves ρ , the density of quartz, C_{66} , C_{11} , and C_{55} which are the elastic constants, T the quartz plate thickness and λ_x and λ_z which are the effective lengths of the motional region of the quartz plate in the X and Z' directions respectively.

What is left to be determined by our empirical analysis is how λ_x and λ_z depend on the dimensions of the electrode and the quartz plate. It is easily demonstrated that if we employ only the dimensions of the electrode in the X and Z' direction that the calculated frequencies will not correlate with the measured frequencies. This is because the dimensions which define the motional region of the quartz plate are not bound by the dimensions of the electrode. We have, therefore set up several experiments to determine the dependency of λ_x and λ_z on the dimensions of the electrode and quartz plate.

DETERMINATION OF λ_x AND λ_z

In the first set of experiments we employed quartz resonators of the kind which are illustrated in Figure 2. We used quartz blanks which were fairly large in diameter, in this case 100 plate thicknesses. Large electrodes were initially vacuum deposited on the

two major surfaces of the quartz plate. The diameter of these electrodes was 71 plate thicknesses and they were located directly opposite each other. The resonators were mounted and bonded onto a simple base which would allow for the deposition of a small circular frequency adjustment spot directly opposite on both surfaces of the quartz plate. The experimental set-up is diagramed in Figure 3. This shows a small vacuum chamber containing two evaporation sources located on each side of the quartz resonator but shielded from the quartz resonator by a shutter. The crystal resonator is held in a fixture which contains PI-circuits in accordance with the recommended test method in IEC Publication 444, which was employed in these experiments.

A mask was fitted over the quartz resonator and was designed to very accurately align the adjustment deposition pattern on both sides to the center of the quartz plate. The experimental procedure consisted of sequentially making small frequency adjustments and then measuring the motional parameters, primarily the motional inductance. Comtec's Model 770204 Test System² was employed to make these measurements. An example of the resulting data is illustrated in Figure 4. For the resonator shown in Figure 2, the initial motional inductance was measured to be 1.9 milli-henries. After an initial small frequency adjustment is made, two resonant frequencies very close to each other are generated. One frequency was slightly higher than the original frequency, and the other slightly lower. The inductance of the lower of these two frequencies was measured to be approximately 3 milli-henries and the inductance of the higher of these two frequencies was approximately 5.6 milli-henries. As additional mass was deposited onto the center of the electrode and the frequency was sequentially lowered, the inductance associated with the lower of the two frequencies continued to increase, whereas the inductance associated with the higher of the two frequencies decreased and became asymptotic to the inductance value related to the original frequency of the resonator.

Clearly, what takes place, is the constriction of the effective motional volume as a function of increased mass loading associated with the central region of the quartz plate and the expansion of the effective motional volume associated with the peripheral region around the lower frequency central region. The inductance associated with the central region increases and becomes asymptotic to the inductance value which is associated with the area of the adjustment spot, which in the example discussed here is approximately 43 milli-henries. It is also important to note that the product of the two inductances divided by their sum equals 1.9 milli-henries, the initial value of the inductance.

The conductive electrode which extends over the entire motional area, functions as a collector of the surface charge which is established during the excitation of the central region at the lower frequency. As we know the motional inductance is inversely proportional to the area of the region in resonance. Therefore, by measuring the motional inductance with the lower frequency central region, we can directly obtain a measurement of the effective motional volume of the quartz plate under resonance.

This condition is illustrated in Figure 5. The effective motional volume of the quartz plate is extended beyond the area of the small electrode by an amount σ_x in the X-crystallographic direction and σ_z in the Z' crystallographic direction. Since the fre-

quency of the fundamental thickness shear mode of motion in the region under the small electrode is lower than that of the peripheral region around it, the mode cannot propagate into that region, but is acoustic displacement exponentially decays as the wave "tunnels" outward from the extremities of the smaller electrode. The magnitude of the exponential decay depends upon the relative difference between ω_{e2} and ω_{e1} . As the difference between these two frequencies increases, the rate of the exponential decay in the displacement increases; therefore, the effective motional volume constricts to the dimensions defined by the adjustment spot. If we assume that at the defining edge of this motional volume, the effective amplitude of the wave in the X direction is equal to that of the wave in the Z' direction, then we have σ_x is approximately equal to $1.26 (\sigma_z)$.

It is important to point out that in this configuration we have a single resonator with two 'fundamental' frequencies whose inductances may be controlled within certain limits. By using this configuration we can achieve the complete half lattice area of the filter. This will be discussed in detail later.

Using the approximation for the ratio of $\sigma_x:\sigma_z$ and assuming the motional region to be elliptic in shape we compute from the inductance measurements the related areas, and in turn we compute σ_x and σ_z . Figure 6 shows a plot of the magnitude of σ_x and σ_z as a function of the differential mass loading, Δ given by the equation

$$\Delta = \frac{\omega_{e1} - \omega_{e2}}{\omega_{e1}}$$

As can be seen from this data, the plot appears to be well mannered when plotted on a logarithmic scale. The ordinate in this case is σ_x and σ_z and the abscissa is the differential mass loading. Utilizing this data we can determine the value for the effective motional length in the λ_x and λ_z direction as a function of the diameter of the electrode and the mass loading.

ALTERNATIVE DETERMINATION OF λ_x AND λ_z

The next series of experiments involved in the fabrication of a large quantity of resonators covering the frequency range from 6 to 20MHz. The frequency spectrum of the resonators were scanned to determine the frequencies and relative magnitudes of the inharmonic thickness shear modes.

The method employed in the scanning of the frequency spectrum associated with the quartz resonators is that defined in the IEC Publication 283. The crystals are placed in a hybrid balanced bridge circuit as shown in Figure 7. The experimental procedure involved sequentially scanning the frequency spectrum to determine the inharmonic mode frequencies then increasing of the mass loading by electroplating the entire electrode and repeating the scanning step. By identifying the 112 and the 121 modes when detected, but mainly utilizing the 113 and 131 modes, the effective motional lengths λ_x and λ_z were computed using the equation given in Figure 1 and substituting into it the frequency difference between the fundamental F_{111} mode and the F_{113} and F_{131} modes respectively.

The results from a large body of experimental data which actually encompassed many thousands of frequency measurements, is shown in Figure 8. Figure 8 shows

σ_x and σ_z , the extensions of the motional volume beyond the dimensions of the electrode in the X and Z directions respectively, plotted as a function of the mass loading, Δ .

A relatively simple equation can be fit into these curves. The base equation which fits this type of curve is that the natural log of sigma divided by T is equal to a constant K1 plus a constant K2 times the natural log of delta, the mass loading. From the data plotted in Figure 8 we determine that for the Z direction K1 is equal to -1.292, K2 is equal to -0.602. We can write the equation for σ_z as

$$\frac{\sigma_z}{T} = 0.275 \Delta^{-0.602}$$

Therefore, λ_z is given by the equation

$$\lambda_z = DE + 0.55 T \Delta^{-0.602}$$

For the X direction extension K1 equals -1.411, and K2 equals -1.67.

σ_x is given by the equation

$$\frac{\sigma_x}{T} = 0.244 \Delta^{-0.67}$$

And the equation for λ_x is given by

$$\lambda_x = DE + 0.488 T \Delta^{-0.67}$$

If we employ λ_x and λ_z as defined by these equations, in the equation for Fmnp we can get reasonably accurate results, which can be used to determine suitability for use in filter applications.

Two examples are illustrated in the data given below. In this figure we have two sets of data involving resonators in the 7.1MHz frequency region. The first case is for a resonator having an electrode diameter of 3.22mm and a mass loading of 1.0%. As can be seen from the tabulation, the deviation between the computed and measured value is less than 2KHz for modes having break-ups only along the X or Z' crystallographic axis. In the second case, the electrode diameter is substantially larger in size and the mass loading is 1.185%. In this case the deviation between the computed and measured values is approximately 2KHz.

COMPARISON WITH MEASURED DATA

CASE 1: Electrode Diameter 3.22mm
Mass Loading 1.0%

MODE	FREQUENCY (KHZ)		DEVIATION (KHZ)
	COMPUTED	MEASURED	
F111	7107	7107	
F112	7120	7119	-1
F121	7123	7122	-1
F122	7173	7169	-4
F113	7203	7204	1
F131	7213	7211	-2

CASE 2: Electrode Diameter 4.88mm
Mass Loading 1.185%

MODE	FREQUENCY (KHZ)		DEVIATION (KHZ)
	COMPUTED	MEASURED	
F111	7144	7144	
F112	7176	7178	2
F121	7180	7182	2
F113	7230	7129	-1
F131	7241	7245	3

During the acquisition of the experimental data, the activity of the unwanted modes was recorded and can be shown as a function of mass loading.

Figure 9 illustrates a summary of this kind of data. In the figure we have shown the suppression of the strongest mode as a function of mass loading for various types of resonator designs.

We have defined the different resonator designs by it's B-factor. The B-factor is defined as the product of the resonators frequency times the motional inductance. The frequency being measured in Megahertz and the inductance in milli-henries. Therefore, if the B-factor were 100 and the frequency of the resonator was 10MHz, then the inductance associated with that resonator would be 10 milli-henries. Shown in Figure 9 are curves which range from a B-factor of 105 to 700. The curves illustrated in Figure 11 are very typical for resonators of the values shown. As can be seen, the value of the mass loading for which the suppression of all the modes is a maximum, varies with the magnitude of the B-factor. In the case of a B-factor of 105, the maximum suppression is obtained with a mass loading of approximately 0.6% as shown here. For a B-factor of 140, the maximum suppression is obtained with a mass loading of approximately 1.01%. The same is true for a B-factor of approximately 186, but as the B factor increases to 203, a higher value of mass loading is required to achieve the maximum suppression of the inharmonic modes. This data can be summarized in a slightly different manner and plotted in Figure 10. What we have done here, is plot the maximum mode suppression db versus the B-factor. As shown here, if the resonator design required a B-factor of approximately 100, then the maximum unwanted mode suppression that one could expect would be approximately 17db. This may or may not satisfy the requirements of the filter design. If the designer was required to have a greater suppression of the inharmonic modes then resonator designs in which the B-factor exceeded approximately 190 would be necessary. The mode suppression is not significantly improved for resonators having B-factors of 200 or greater.

THE SHARED ELECTRODE RESONATOR (SER)

The resonator construction illustrated in Figure 2 has the equivalent circuit shown in Figure 11. As previously discussed, the motional inductance and capacitance of the series resonant arms are functions of the diameter of the large electrode and the adjustment spot and the relative mass loading of the two regions. Within the limitations prescribed by these parameters, the device may be produced to have a specified inductance ratio and frequency separation. For intermediate bandwidth and wide bandwidth filters

the inductance ratio is important, but the frequency accuracy is of less importance. It is in these applications that we feel the SER will find application.

Further studies of SER devices will be reported in the near future. Devices have been fabricated with up to three controlled resonant regions. Utilization of this type of device can reduce the number of quartz devices by the number of controlled resonant regions in the SER and thereby reduce the cost and size of the filter network.

$$F_{MNP} = \frac{0.5}{\pi \sqrt{\rho}} \left[\frac{C_{66} M^2}{T^2} + \frac{C_{11} N^2}{\lambda_x^2} + \frac{C_{55} P^2}{\lambda_z^2} \right]^{\frac{1}{2}}$$

Where:

F_{MNP} is frequency of the MNP mode

C_{66}, C_{11} & C_{55} are elastic constants

M, N & P are integers denoting the modes

λ_x & λ_z are lengths of the motional region of the quartz plate in X and Z' axis respectively

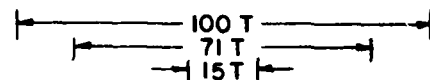
T is the thickness of the plate

DE is the diameter of the electrode

ML is the mass loading of the electrode

ρ is the density of quartz

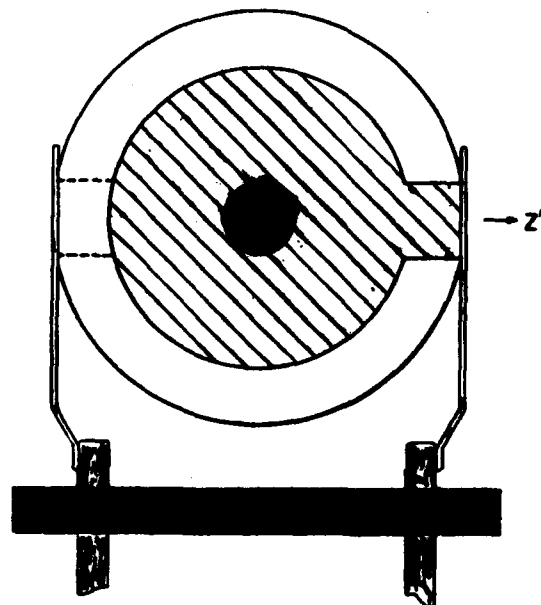
FIGURE 1



$F = 11.6 \text{ MHz}$

$T = 0.14 \text{ MM}$

FIGURE 2



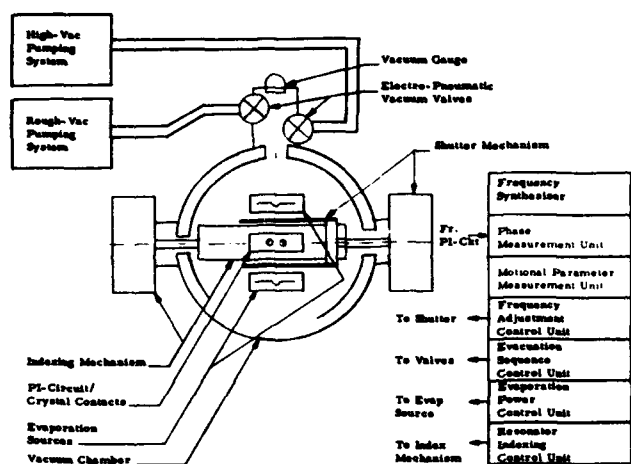
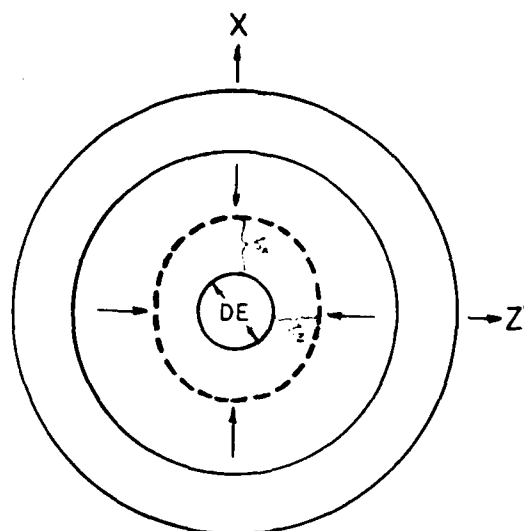


FIGURE 3



$$\omega_{e2} < \omega_{e1} < \omega_u$$

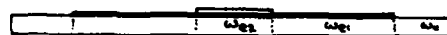


FIGURE 5

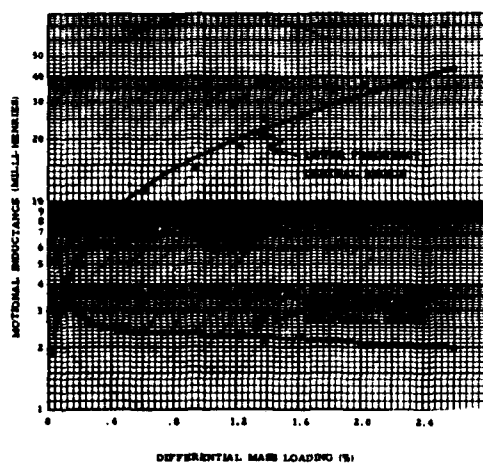


FIGURE 4

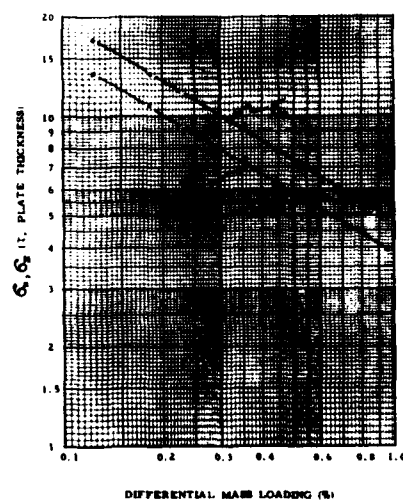


FIGURE 4

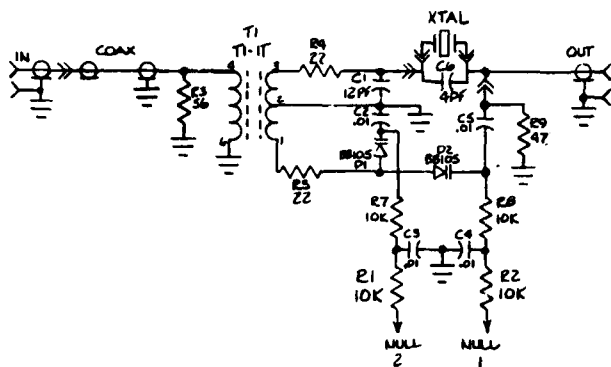


FIGURE 7

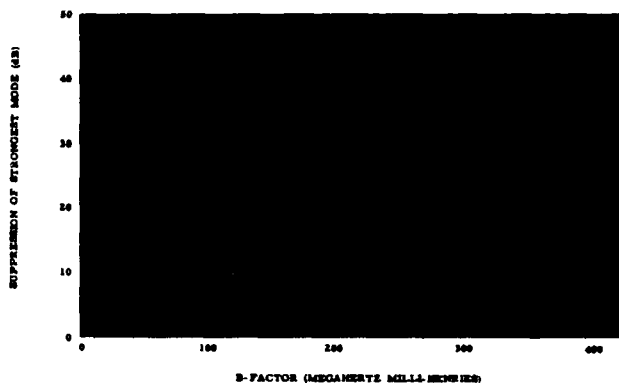


FIGURE 10

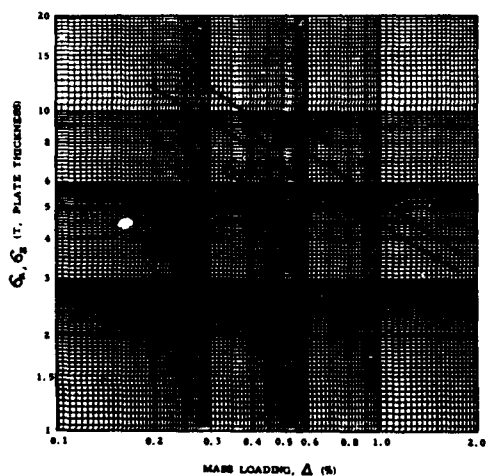


FIGURE 1

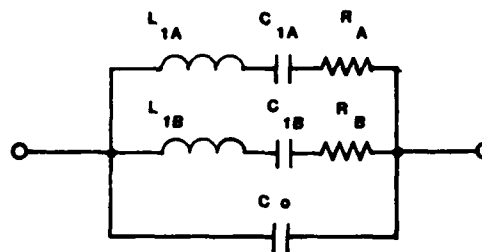


FIGURE 11

Equivalent circuit of a SHARED ELECTRODE RESONATOR with two resonances

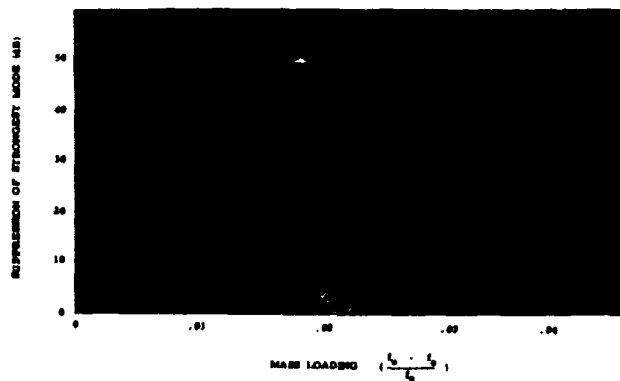


FIGURE 2



PROPERTIES OF AT QUARTZ RESONATORS ON WEDGY PLATES

Lawrence N. Dworsky

Motorola, Inc.
Schaumburg, Ill.

ABSTRACT

It is well known in the quartz literature that non-parallel plate surfaces will cause deterioration in the performance of AT resonators. This deterioration consists of a decrease in the motional capacitance of the resonator's desired operating mode, often without a corresponding decrease in the motional capacitances of observed spurious modes.

A model was developed for thickness-shear vibrations of a heavily electroded rectangular AT plate, assuming a small linear thickness variation (wedginess) across the plate. The model predicts that the standing waves corresponding to the different anharmonic overtone modes of vibration will "bunch up" in the thicker region of the plate. Higher order anharmonic modes will be less affected by the wedginess than will lower order anharmonics. The observed consequence of this behavior is that the motional capacitance of the lowest mode (the desired operating mode) will deteriorate, while the motional capacitances of the higher order modes (spurious modes) may not deteriorate at all.

INTRODUCTION

In 1976, Tiersten presented an accurate model for the rectangular trapped energy resonator on an AT quartz plate.¹ This model allowed for calculation of the resonant frequencies and motional parameters of all trapped modes. In 1979, Tiersten and Smythe presented a model for the trapped energy resonator on a slightly convex plate.² Their model was derived by assuming a small quadratic thickness variation over the plate, and then assuming an approximation based upon Tiersten's earlier work. In this paper, we use this same technique to derive a model for the trapped energy resonator on a plate with a slight linear thickness variation, or "wedginess" over the plate. Simple approximate relationships are then derived which show how the observed motional capacitances of the resonator vary with the magnitude of the wedginess.

THE WEDGY PLATE

Although energy trapping theory predicts that essentially spurious mode (spur) free resonators are possible, often practical considerations preclude this possibility. These considerations typically include process related minimum electrode thickness

and motional parameter related minimum electrode area. If the electrode thickness and/or the electrode area is sufficient for several anharmonic modes to be trapped - in both X and Z - we consider the resonator to be "heavily electroded" and note that the X thickness shear displacement amplitude may be approximated by

$$u_x \approx \sin\left(\frac{N_y \pi y}{2h}\right) \cos\left(\frac{N_x \pi x}{2L_x}\right) \cos\left(\frac{N_z \pi z}{2L_z}\right) \quad (1)$$

where u_x is the X thickness shear displacement amplitude
y is the thickness direction
x and z are the lateral directions
 N_y are the number of $\frac{1}{2}$ wavelengths in thickness
 N_x and N_z are the number of $\frac{1}{2}$ wavelengths in x and z, respectively
 $2L_x$ and $2L_z$ are the electrode dimensions

Piezoelectric excitation of this resonator occurs only for odd values of N_x , N_y and N_z . The "heavily electroded" approximation is only valid for low values of N_x and N_z . However, $N_x = N_z = 1$ is typically the operating mode, in which case $N_x = 1$, $N_z = 3$ and $N_y = 1$ are the most troublesome spurious modes.

Equation (1) is the solution of

$$M_h \frac{\partial^2 u_x}{\partial x^2} + C_{55} \frac{\partial^2 u_x}{\partial z^2} + \overline{C_{66}} \frac{\partial^2 u_x}{\partial y^2} + \rho \omega^2 u_x = 0 \quad (2)$$

where C_{55} and C_{66} are elastic constants
 ρ is the mass density
 ω is the (radian) frequency
 M_h is Tiersten's correction to C_{11}

subject to the boundary conditions $u_x = 0$ at the edges of the electrodes and that the surfaces of the electrodes are free.

For small variations of h with x and z, assume that

$$u_x = U(x, z) \sin\left(\frac{N_y \pi y}{2h}\right) \quad (3)$$

This reduces (2) to

$$C' \left(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial z^2} \right) + U [\rho \omega^2 - \overline{C_{66}} (N_y \pi / 2h)] = 0 \quad (4)$$

$$\text{where } C' = (M_n + C_{55})/2, \quad (5)$$

an approximation too poor for accurate frequency calculations, but adequate for this discussion. Accuracy of a particular example may be improved, if so desired, by replacing C' with the exact values for that example.

Now, let

$$h = h_0 + \alpha_x x + \alpha_z z \quad (6)$$

For small variations in h about h_0 ,

$$\frac{1}{h^2} \approx \frac{1}{h_0^2} \left[1 - \frac{2}{h_0} (\alpha_x x + \alpha_z z) \right] \quad (7)$$

Substituting (7) into (4),

$$C' \left(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial z^2} \right) + [\rho \omega^2 - \overline{C_{66}} (N_y \pi / 2h_0)^2 - \gamma^2] U = 0 \quad (8)$$

Letting $U = F(x)G(z)$ and following a standard separation of variables procedure results in two ordinary differential equations, identical in form:

$$d^2 F/dx^2 + K_x^2 F = 0, \quad F(\pm L_x) = 0 \quad (9)$$

$$d^2 G/dz^2 + K_z^2 G = 0, \quad G(\pm L_z) = 0 \quad (10)$$

where

$$K_x^2 = [\rho \omega^2 - \overline{C_{66}} (N_y \pi / 2h_0)^2 - \gamma^2] / C' \quad (11)$$

$$K_z^2 = \gamma^2 / C' \quad (12)$$

$$\beta = \overline{C_{66}} N_y^2 \pi^2 \alpha / 2 C' h_0^3 \quad (\text{for both } x \text{ and } z) \quad (13)$$

Note that, for a given design, i.e. a given choice of N_y and h_0 , β is proportional to α . Also, from (11) and (12),

$$\omega^2 = [\overline{C_{66}} (N_y \pi / 2h_0)^2 + C' (K_x^2 + K_z^2)] / \rho \quad (14)$$

The solutions F and G , to (11) and (12), describe the x and z dependencies of the standing wave $U = FG$ on the wedgy plate. Once F and G are known it is possible to calculate the motional parameters of the resonator. Equations (11) and (12) are eigenvalue equations, and each has an infinite number of solutions. Only the lowest values of K_x and K_z are needed in order to calculate the motional parameters of the operating mode of a resonator. The next-to-lowest values of K_x and K_z are needed in order to calculate the motional parameters of the dominant spurs.

WAVE FUNCTIONS ON THE PLATE

Since (9) and (10) are identical in form, it is only necessary to solve either of them. Consider a finite difference approximation to (9):

$$\text{Let } \Delta = 2L_x / (N + 1) \quad (15)$$

be the separation between N equally spaced points

within $-L_x < x < L_x$.

For the i th point,

$$x_i = \Delta(i - L_x/\Delta), \quad 0 \leq i \leq N+1 \quad (16)$$

The second derivative of F is approximated by

$$\frac{d^2 F}{dx^2} \approx \frac{F_{i+1} - 2F_i + F_{i-1}}{\Delta^2} \quad (17)$$

and the ordinary differential equation (9) is replaced by the set of linear algebraic equations

$$F_{i+1} + F_{i-1} - 2F_i + (K_x \Delta)^2 F_i + (1 - L_x/\Delta) \Delta^2 F_i = 0 \quad (18)$$

with the boundary conditions

$$F_0 = F_{N+1} = 0 \quad (19)$$

In matrix form, (18) appears as

$$\begin{bmatrix} A_1 & 1 & 0 & 0 & 0 & - \\ 1 & A_2 & 1 & 0 & 0 & - \\ 0 & 1 & A_2 & 1 & 0 & - \\ - & - & - & - & - & - \\ - & - & - & - & - & - \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ - \\ - \\ - \end{bmatrix} = 0 \quad (20)$$

where

$$A_i = -2 + (K_x \Delta)^2 + \beta(1 - L_x/\Delta) \Delta^2 \quad (21)$$

and the resonances are found by setting the determinant of the coefficient matrix = 0.

The determinant of the coefficient matrix is easily evaluated by the repeated transformation

$$\begin{bmatrix} A_1 & - & - & - & - \\ 1 & A_2 & - & - & - \\ 0 & 1 & A_3 & - & - \\ 0 & 0 & 1 & A_4 & - \\ - & - & - & - & - \end{bmatrix} =$$

$$\begin{bmatrix} A_1 & - & - & - & - \\ 0 & (A_2 - 1/A_1) & - & - & - \\ 0 & 1 & A_3 & - & - \\ 0 & 0 & 1 & A_4 & - \\ - & - & - & - & - \end{bmatrix} =$$

$$\begin{vmatrix}
 A_1 & - & & & \\
 0 & (A_2 - 1/A_1) & - & & \\
 0 & 0 & (A_3 - \frac{1}{A_2 - 1/A_1}) & - & \\
 0 & 0 & 1 & A_4 & - \\
 - & - & - & - & -
 \end{vmatrix} \quad (22)$$

etc.

This procedure transforms the determinant to an upper diagonal determinant, which is evaluated by simply taking the product of the diagonal terms. This procedure is efficient enough that relatively large order determinants may be quickly evaluated on a small personal computer.

To find the value(s) of K_x by determinant evaluation it is necessary to "guess" at an initial value of K_x , evaluate the determinant, and then to iteratively improve upon the guess until a value of K_x is found which causes the determinant to be equal to zero. Once a value for K_x is found, the values of F_i are found by setting F_1 (say) = 1, and then solving the matrix equation (20) for the remaining F_i .

Figure 1 shows K_x vs β for the parameters shown in the figure. Note that there is a region of β ($0 \leq \beta \leq 2 \times 10^{12}$) where K_x hardly varies from its $\beta = 0$ (no wedginess) value, and then for higher values of β , K_x falls rapidly. From (14) we may interpret the information in figure 1 as saying that for $\beta < 2 \times 10^{12}$ the resonator frequency essentially stays constant, and then as β gets large the frequency falls rapidly.

Figure 2 shows the function $F(x)$ for the same parameters as figure 1, for several values of β . As may be seen, for $\beta = 0$, $F(x)$ is a $\frac{1}{2}$ wave cosinusoid. As β increases, $F(x)$ appears to get squeezed towards the thicker region of the plate. This agrees physically with the concepts of energy trapping in AT plates, and also explains why the resonant frequency decreases as β increases.

Figure 3 shows $F(x)$ for a higher order solution, ($N_x = 5$), for the same parameters as used in figures 1 and 2. As may be seen, the effects of wedginess appear the same in figure 3 as in figure 2, except that it takes much larger values of β for the case of figure 3 to produce noticeable squeezing of $F(x)$. This means that, in actual devices, we would expect the operating mode, ($N_x = 1$) to suffer the effects of wedginess much more severely than the spur modes ($N_x > 1$). In other words, we would expect the motional capacitance to degrade much more severely for the operating mode than for the spur modes.

A SIMPLIFIED MODEL

In order to develop simple closed form equations relating observed motional capacitance to wedgi-

ness, it is necessary to assume a simple form for $F(x)$. Inspection of figures 2 and 3 shows that $F(x)$ may be approximated by

$$F(x) \approx \begin{cases} \sin \frac{N_x \pi (x-a)}{L_x - a} , & a \leq x \leq L_x \\ 0 & -L_x \leq x \leq a \end{cases} \quad (23)$$

where a is a parameter which is in some way dependent upon β , and $-L_x \leq a \leq L_x$.

To find the dependence of a upon β , let (23) be used as a trial function in a variational formulation of the (eigenvalue) problem:

$$K_x^2 \leq \frac{\int_{-L_x}^{L_x} (dF/dx)^2 dx - \int_{-L_x}^{L_x} \beta x F^2 dx}{\int_{-L_x}^{L_x} F^2 dx} \quad (24)$$

noting that the above form of $F(x)$ satisfies the original boundary conditions.

Substituting (23) into (24),

$$K_x^2 \leq \left(\frac{N_x \pi}{L_x - a} \right)^2 - \beta_x (L_x + a)/2 \quad (25)$$

Equation (24) states that the correct value of K_x is always equal to or less than the approximate expression shown. If the actual function $F(x)$ were available and were used in (24), the exact value of K_x would result. This means that an expression such as (23), with an adjustable parameter, is most accurate when that parameter is set to give the minimum value of K_x in (24). In order to find the value of a which minimizes K_x , we take the derivative of (25) with respect to a and set it to 0:

$$\begin{aligned} d(K_x^2)/dx &= 0, \text{ or,} \\ a &= L_x - (2N_x \pi)^2 / \beta_x \quad (26) \end{aligned}$$

Equation (26) is valid only for $-L_x \leq a \leq L_x$. In other words, (26) is valid for

$$\beta_x \geq (N_x \pi)^2 / 2L_x^2 \quad (27)$$

$$\text{otherwise } a = -L_x \quad (28)$$

Figures 4, 5, and 6 are repeats of figures 1, 2, and 3, but using (23), (27), and (28) to generate the figures. As may be seen, the simplified model reproduces the results of the finite difference solution quite well. Naturally the corners which appear in the curves of figures 4, 5, and 6 are artifacts of the simplified model, and are not at

all physical.

MOTIONAL CAPACITANCE

The calculation of motional capacitance is well treated in the literature. For example, reference 1 has detailed expressions for the motional capacitance of trapped energy resonators. For the purposes of this treatment, we are interested only in how the motional capacitance of a resonator changes as a function of wedginess. For that reason it is perfectly reasonable to set all elastic, piezoelectric, and dielectric coefficients = 1, and consider only the functional form of an expression for motional capacitance:

$$C_m = \frac{\left[\iint_s U(x,z) ds \right]^2}{\iint_s U^2(x,z) ds} \quad (29)$$

where s is the electrode area.

Evaluating (29) for the simplified F and G forms (Eqs. (23), (26), and (28)) yields

$$C_m = \frac{(L_x - a_x)(L_z - a_z)}{N_x^2 N_z^2} \quad (30)$$

For a given N_x and N_z , normalizing to the $\alpha_x = \alpha_z = 0$ case, [i.e. letting $a_x = -L_x$ and $a_z = -L_z$], (30) becomes

$$C_{mn} = \frac{(L_x - a_x)(L_z - a_z)}{4L_x L_z} \quad (31)$$

As an example, consider the case of a square resonator with equal amounts of wedginess along both x and z axes. In this case, (31) becomes

$$C_{mn} = \left(\frac{L_x - a_x}{2L_x} \right)^2 \quad (32)$$

Figures 7, 8, and 9 show C_{mn} vs. electrode size, for several values of thickness change over the electrode. All 3 figures are similar in appearance, with only different parameters to distinguish them. In each figure, for a given thickness change over the electrode, there is a maximum size electrode which will show essentially no motional capacitance deterioration due to wedginess. Once that size is exceeded, the motional capacitance falls off rapidly. Alternatively, the figures show what thickness change over the electrode may be tolerated for a given size electrode without appreciable motional capacitance deterioration.

DESIGN GUIDELINES

The maximum amount of wedginess which can be tolerated without a noticeable deterioration in resonator motional capacitance is that value of α , (or β), at which appreciable squeezing of the standing wave $F(x)$ and/or $G(z)$ occurs. From (26), therefore,

$$2L_x < (2\pi)^{2/3} / \beta^{1/3} \approx 3.4 / \beta^{1/3} \quad (33)$$

A reasonable approximation for the frequency of a resonator, for the accuracy required here, is the frequency of an infinite plate,

$$F_0 = \frac{N_x}{2h} \left(\frac{C_{66}}{\rho} \right)^{1/2} \quad (34)$$

Combining (34) and (12),

$$\beta \approx 3.5 \times 10^9 F_{MHz}^2 \alpha / N_y \quad (35)$$

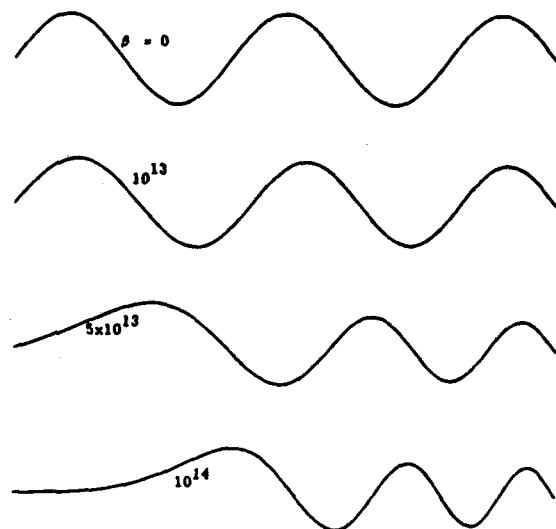
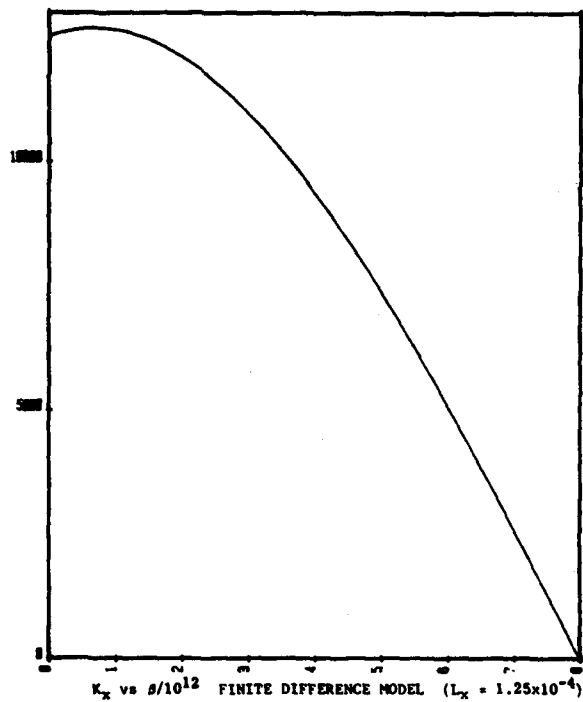
Equations (33) and (35) form an approximate plate parallelism design guideline: For a given design, knowing the electrode size required, (33) specifies the maximum allowed value of β . Knowing the overtone of operation and the operating frequency, (35) then specifies the maximum value of α . The maximum allowed thickness variation over the electrode is then simply $2\alpha L_x$.

SUMMARY

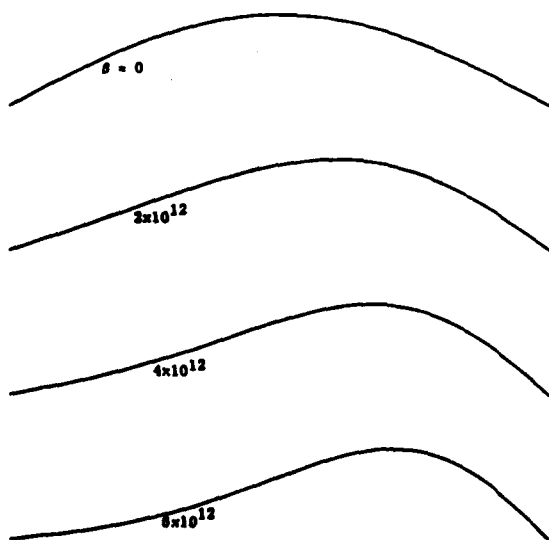
An approximate model has been developed for the heavily electroded wedgy AT quartz plate. Based upon a simplified approximation of this model, closed form expressions were developed which relate the deterioration of the motional capacitance of the resonator to the plate wedginess. This in turn led to the specification of a maximum allowed thickness variation over the electroded area of a plate which could be tolerated without significant deterioration of the motional capacitance.

REFERENCES

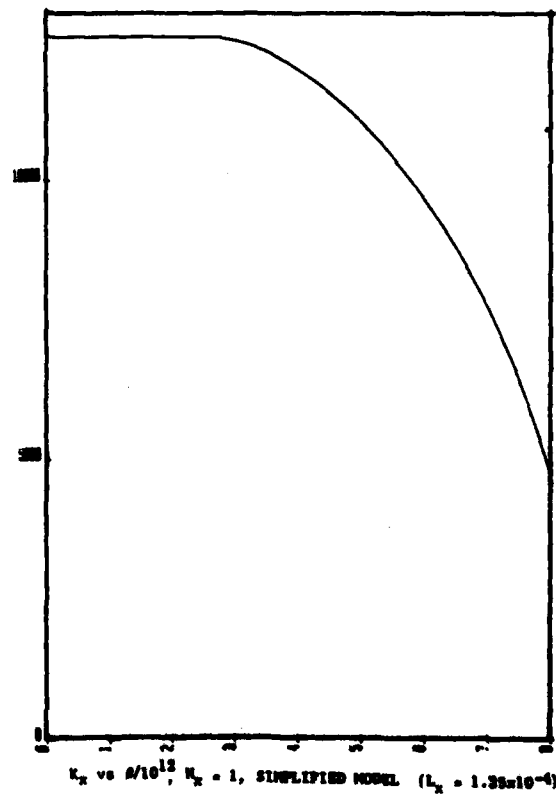
1. H. F. Tiersten, "Analysis of Trapped Energy Resonators Operating in Overtones of Coupled Thickness-Shear and Thickness-Twist." J. Acoust. Soc. Am., Vol 59, p 879, (1976).
2. H. F. Tiersten and R. C. Smythe, "An Analysis of Contoured Crystal Resonators Operating in Overtones of Coupled Thickness Shear and Thickness Twist." J. Acoust. Soc. Am., Vol 65, p 1455, (1979).

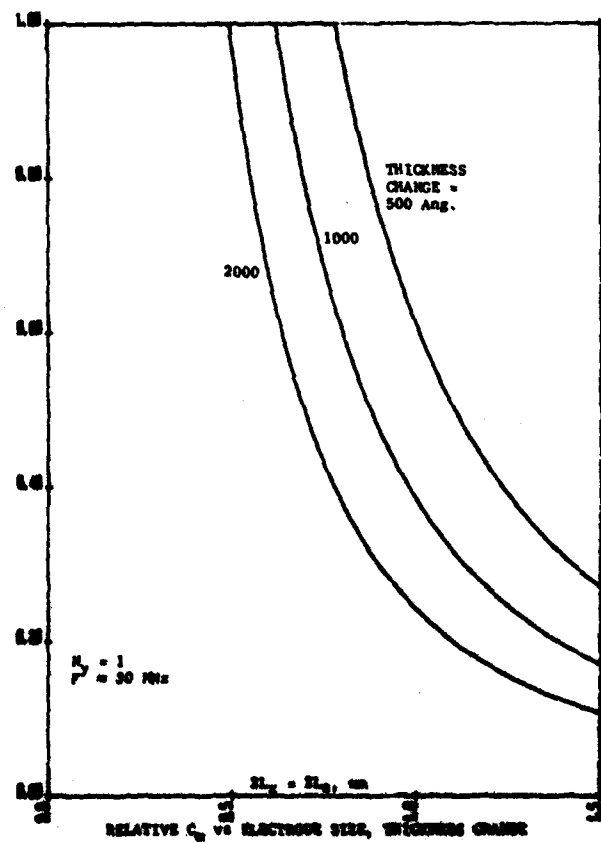
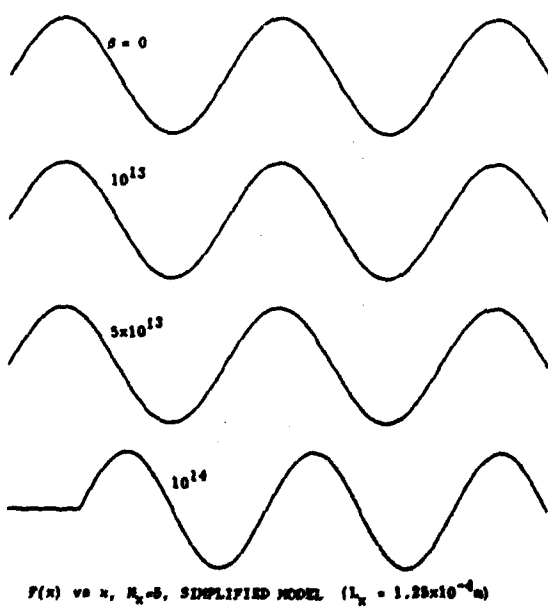
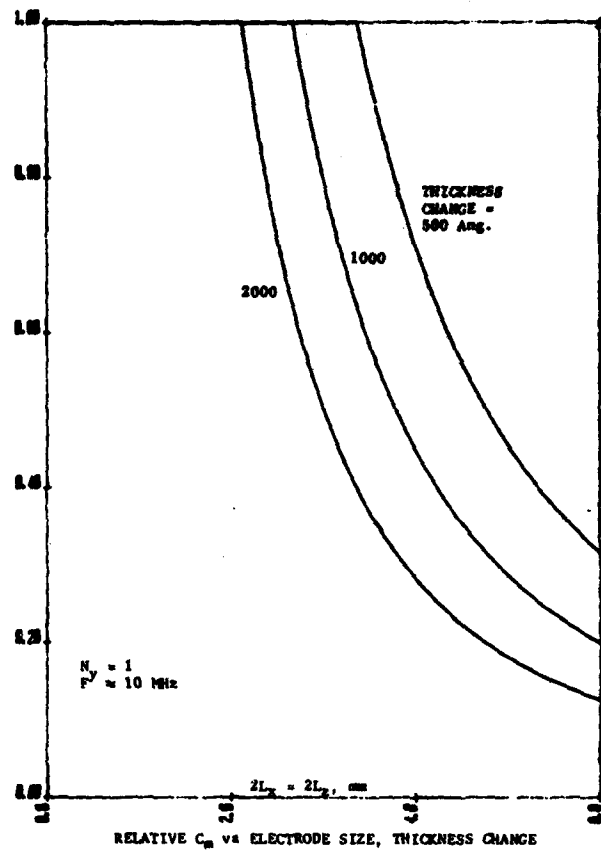
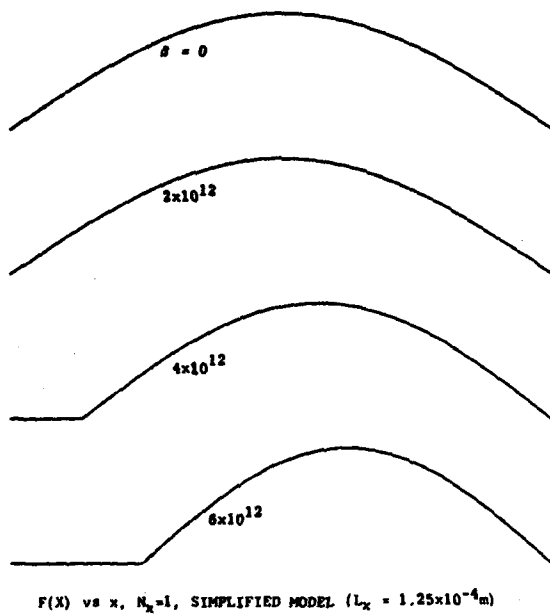


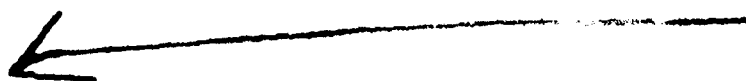
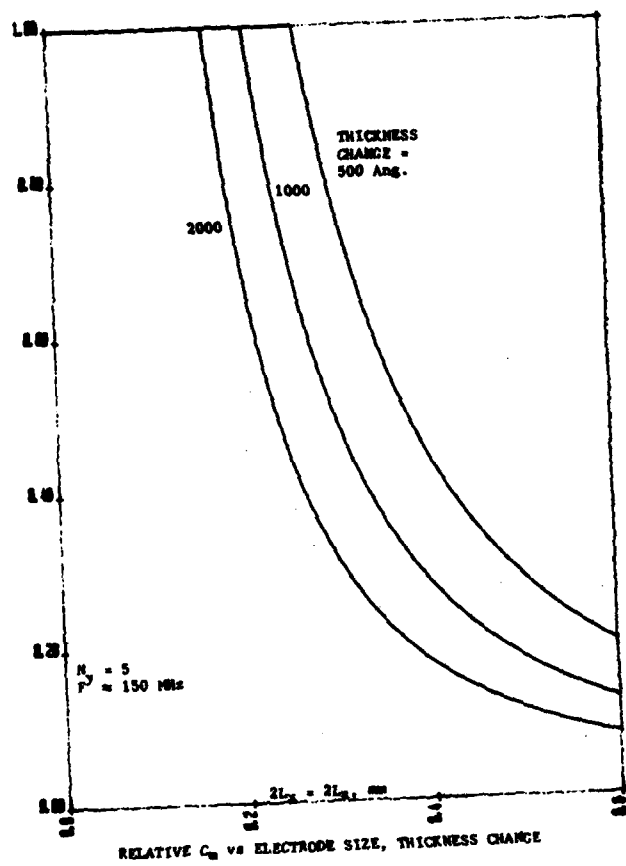
$F(x) \text{ vs } x$, $N_x=5$, FINITE DIFFERENCE MODEL ($L_x = 1.25 \times 10^{-4}$ m)



$F(x) \text{ vs } x$, $N_x=1$, FINITE DIFFERENCE MODEL ($L_x = 1.25 \times 10^{-4}$ m)







J. DETAINT, H. CARRU, P. AMSTUTZ, J. SCHWARTZEL

CENTRE NATIONAL D'ETUDES DES TELECOMMUNICATIONS
196, rue de Paris Bagneux 92220 FRANCEABSTRACT

Acoustical radiation is of particular intensity for resonators having large normal components of displacement. This can be used as a simple and efficient way to transfer energy between two resonators in order to constitute a filter structure. A one dimensional modelisation of gas coupled extensional composite resonators is described. This model indicates a simple dependance of the response on several parameters of the gas gap. A theoretical investigation of vibrational and electrical behaviour of composite resonators and of acoustically coupled biresonator filters made by a bi or tridimensional finite element analysis is reported. This analysis was performed for two different types of structures : one of cylindrical geometry, the other of parallelepipedic geometry.

Air coupled filters : Experimental work reported concerns composite flexural discs coupled by fixing them at a short distance. The influence of several parameters such as coupling distance and gas pressure is given.

In order to obtain devices with low losses, low aging and good temperature behaviour, composite flexural bars made of Durinval alloys and Y-45 Lithium Niobate and X + 5° flexural quartz bars were used. On the whole such 2 poles filters exhibit good performances and may be a simpler and smaller valid alternative to classical LC, crystal and mechanical filters for low frequencies.

Pressure transducers : The pressure dependance of insertion loss of gas coupled filters is quasi-linear over three decades. A response is obtained across at least 5 decades of pressure.

I - INTRODUCTION

Acoustical radiation is known to be one important energy loss mechanism of piezoelectric devices contained in pressurized enclosures. This phenomenon is of particular intensity for resonators having vibration modes with a large normal component of displacement, on a boundary surface of large area. This condition occurs for many types of low frequency resonators. Acoustical radiation can transfer energy between resonators [1]. A question arises then : Is it possible to constitute efficient filters structures using acoustical coupling ? Our purpose in the present paper is to examine the principal properties of gas closely coupled resonators through theoretical and experimental investigations. This research was conducted in view to obtain simple filters in the supravocal frequency domain for telephony applications in particular for an anti-larsen system for a free hand set. In the course of this work it was shown that this type of structure can

have also interesting application as pressure transducer.

II - THEORETICAL INVESTIGATIONS

The Figure 1 represents the different structures we have considered in this study, most use composite resonators which permit to obtain larger coupling area than classical ones and permit an easy fixation of two resonators at a short distance. The composite flexural bars that have a length-thickness motion can be made with very low resonance frequency together with small dimensions. As we will see latter composite flexural bars can be easily temperature compensated and their coupling coefficients can be adjusted between 0 and about 25 % (with zero F.T.C).

II.1 1 - D model

We consider two resonators of the symmetrical Langevin type coupled by a small gas gap (Figure 2). The end parts of resonators are made of a mechanical filter metallic alloy. The piezoelements are either Y + 36° LiNbO₃ or a high Q, low FTC piezo-ceramic.

The 1-D model considers out-side radiation i.e. loading of the end faces by the mechanical impedance of a semi infinite column of gas. The model uses the formalism of 4 x 4 electro-mechanical chain matrix with parallel electrical connections, introduced by Sittig [2] for the piezoelements, and classical mechanical 2 x 2 chain matrix for the elastic layer including gas. Connection, at the inner side of piezoelement, between the 4 x 4 and 2 x 2 matrix, taking account of the electrical open circuit condition is accomplished by algebraical manipulations equivalent to those needed when a serial combination of a nullator and a norator is considered. Mechanical losses in solids are introduced by imaginary parts of material constants [3], compressional viscosity is considered in gases that are regarded otherwise as ideal. Temperature and pressure effects are taken into account for the gas. For the piezoelectric layer, stiffened bar constants are used according to the global vibration mode of the resonators. The computer program set up for 1 D modelisation can analyse either gas coupled filter or classical 1 D electromechanical filters composite resonators. It considers any type or 1-D modes (thickness modes, length extensional modes, torsional...) of stiffened or unstiffened nature for the piezoelectric layers [2], in the non-quasi mono mode approximation. Results are obtained in the form of total chain matrix, S₂₁ and insertion loss for quadrupoles and impedance for dipoles.

Before presenting computed results, let us examine specific properties of small distance gas coupling and establish, with a very simple analytic approximation,

AD P 002482

some characteristics of short distance gas coupled resonators. The 1-D chain matrix of a gas layer is [4] :

$$\begin{pmatrix} F \\ u_1 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} F_2 \\ u_2 \end{pmatrix}$$

where $\begin{cases} \alpha = \delta = \cos \theta_g & \text{with } \theta_g = 2\pi f l_g / v_g \\ \beta = j Z_g \sin \theta_g \\ \gamma = \frac{j}{Z_g} \sin \theta_g \end{cases} \quad \begin{cases} Z_g = 2\pi f \rho_g / v_g \\ \rho_g = \rho_g \end{cases}$

since l_g is very small as compared to wave-length in gas we can approximate $\sin \theta_g$ by θ_g and $\cos \theta_g$ by 1. In the hypothesis of an adiabatic ideal gas v_g and Z_g depend as follow of Pressure and Temperature :

$$\begin{aligned} v_g(P,T) &= v_0 \sqrt{\frac{273+T}{273}} \quad (T \text{ in } ^\circ\text{C}) \\ Z_g(P,T) &= \rho_0 \frac{P}{760} \frac{(273)}{(273+T)} v_g(P,T) \quad (P \text{ in mm Hg}) \quad (2) \\ \text{reference state } T_0 &= 0^\circ\text{C}, P_0 = 760 \text{ mm Hg} \end{aligned}$$

If we consider 2 purely mechanical resonators, with a resonant frequency $f_r = v_m/2.l_m$, coupled by the gas gap of length l_g , the total mechanical chain matrix is

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} D & B \\ C & A \end{pmatrix} = \begin{pmatrix} 1+B\gamma & B^2\gamma+2B+\beta \\ \gamma & 1+B\gamma \end{pmatrix} \quad (3)$$

since $\begin{cases} A = D = -1 \\ B = -jZ_m = \frac{f-f_r}{f_r} \text{ and } \\ C = 0 \end{cases} \quad \begin{cases} \alpha = \delta = 1 \\ \beta = jZ_g \theta_g \\ \gamma = \frac{j}{Z_g} \theta_g \end{cases}$

(Approximation $\theta_m \approx \pi$); [5]

By the electromechanical equivalence ($F \rightarrow V, u \rightarrow i$) we have :

$$\frac{e}{v_2} = K \left[(1+B\gamma) + \frac{1}{(R_1+R_2)} (2B+B^2\gamma+R_1R_2\gamma+\beta) \right] \quad (4)$$

for $l_g \ll \lambda_g$; $\frac{l_g}{\lambda_g} \ll \frac{l_m}{\lambda_m}$ and for f near f_r the predominant terms of $\left| \frac{e}{v_2} \right|$ contain all the factor γ so :

$$\left| \frac{e}{v_2} \right| = K' \gamma = K' \frac{l_g}{Z_g v_g}$$

or in fonction of P, T, S_g, l_g

$$\left| \frac{e}{v_2} \right| = K'' \rho_0 v_0^2 \frac{P}{760} \frac{S_g}{l_g} \quad (5)$$

independent of T .

In the electromechanical case this relation can be verified numerically.

Figure 3 represents the computed insertion loss for the composite structure whose dimensions are given on the figure. It was computed with $Q = 4000$ for all solid layers. A flat response occurs for $R_1 = 5,5 \text{ k}\Omega$ and $R_2 = 50 \text{ k}\Omega$. With equal resistances peak responses occur near resonant frequency with low values and near antiresonant frequency for high values. We must emphasize, that, without losses, two "true" band-pass (i.e. with zeros of A) can exist in these conditions. For the geometry chosen, they are very narrow.

In this study we consider the "false" band-pass already mentioned that is much larger (a large fraction of the fa-fr interval) and tends to become a zero attenuation bandpass as either surface S_g , or pressure is increased or l_g is decreased (5). A more detailed examination of the purely mechanical approximation would have also revealed the interest of decreasing the mass of resonators (the B term of (3) contains this mass). With the chosen geometry (constant section) it is not possible to vary significantly the ratio S_g/M_r . On Figure 4 is represented the experimental response for the same structure. Agreement is good in view of the uncertainty in material constants. Please note a different definition of A (here A is the effective attenuation) that explains the difference of about 5dB for flat response with computed results.

Figure 5 gives the computed pressure dependance of amplitude at the center of a flat response in linear coordinates. The linearity of the A(P) relation is well verified. It was also numerically verified that no temperature dependance occurs due to the gas (less than $1/100^{\text{th}}$ of dB for 50°C variations). Computation have shown that, in this case, external radiation of the resonators has only a very very small influence on insertion loss and that viscosity of gas has a negligible influence. On Figure 6 are represented the variation of insertion loss (flat transmission) when the coupling distance l_g is varied. The $1/l_g$ law is verified for small distances. Figure 7 indicates the computed sensitivity of the frequency response (flat transmission) when the resonant frequency of only one resonator is changed (by varying the dimension of one layer of one resonator).

II.2 - Two dimensional analysis

The results given by the 1-D model are thought to be of general significance. So to obtain with gas coupling a wide flat response with low insertion loss we have to take devices with large coupling area, low mass and low coupling distance. So an attracting solution is to use thin flexural discs. The principal drawback of the 1-D model was to consider (plane wave hypothesis) no possibility for the gas to escape the coupling space. This and the composite nature of resonators have led to consider a 2-D modelisation for the structure using flexural disc. (Figure 1).

The detailed geometry of a disc resonator is given on Figure 8 (Half disc). Only modes of axial symmetry are considered. The equations in metal are those of isotropic elasticity, in ceramic those of piezoelectricity and in gas the wave equation. Boundary conditions are $T_n = 0$ for external surfaces (no external radiation) and continuity of u_r and u_z at the interfaces. This leads to the condition $\frac{\partial u_r}{\partial r} = \omega^2 u_r$ [6] at the (inner) gas interface. The solution was obtained by Finite Element Method [6]. The spatial discretisation of 2 coupled composite discs uses 62 elements of rectangular cross section (4 nodes elements with constant shear to enhance the precision for flexural motion), Eigen frequencies (f_r and f_a) and eigen-modes of uncoupled resonators at $V = 0$ or $I = 0$ were obtained by the conventional method [7] [8]. The electrical response is

obtained by condensation of the internal (electro-mechanical) degrees of freedom [9].

Figure 8 displays for a half-disc the first flexural mode of the composite disc as computed by FEM. The associated resonance frequency of 13272 Hz is in good agreement with the experimental value of 12750 Hz according to uncertainty in material constants and to the fact that in experimental resonators the fixation adds some mass that lower the resonance frequency.

On Figure 9 are shown for different R_1 and R_2 the computed effective attenuation of two air coupled discs. Agreement with corresponding experiment is good as can be seen on Figure 10 except for absolute value of attenuation; the difference can be attributed to the modest Q value of resonators 80 to 100.

II.3 - Three-dimensional model

In order to obtain devices with higher Q factors and with easier fixation, we have then considered composite length-thickness flexural bars vibrating in their first free-free mode. The 3-D model considers essentially the same equations and boundary conditions as the cylindrical 2 D-one. The solution (electrical response) is also obtained as a function of geometrical parameters by the finite element method. The discretization uses the twenty modes (2nd order) brick element, 16 elements are used for the filter structure (Figure 11). The same method was also used to obtain the electrical response of resonators [10]. The computed electrical response is somewhat larger than the experimental one which is given on figure 12a. However the model has permitted to demonstrate that without losses a bandpass between low and equal resistances can be obtained. This fact was confirmed only very recently by the experiment (figure 12b).

III - EXPERIMENTAL RESULTS

Several mechanical filter alloys from Acieries d'Imphy (Creusot-Loire) were used in this study. Most important properties of these alloys are given in Table 1

TABLE 1

Alloy	$v(\text{ms}^{-1})$ (ext. bar.)	$\rho(\text{gcm}^{-3})$	FTC $10^{-6}\text{ }^{\circ}\text{C}^{-1}$	Q
Durival C	4900.	8.04	0.±2 Adjustable	$2.5 \cdot 10^4$ (Air)
F.T.C Compensating Durival Alloys	4800. to 4900.	8.04	-5 to +30	up to $8 \cdot 10^4$ vacuum
Elinvar	4815.	8.11	0.±1	$1.7 \cdot 10^4$ (Air)

the piezoelement of the composites were the P X 5 ceramic for the discs (Philips); two developmental types of low F.T.C., high Q ceramics were used for flexural bar: The PX 652 from Philips and the 9281 from Thomson-CSF. For high Q flexural bars, Y-47° lithium niobate was used. The Q value of composite flexural bars can be in the range 2000 to 20 000 according to the piezoelement chosen and to the bonding technology (conductive epoxy bonding or thin film tin soldering) [11].

III.1 - Disc geometry

Further experiments with discs were made with thinner resonators coupled at shorter distances (.5 to .8 mm); the minimal effective attenuation was in the range of 10 to 12 dB. An important part of attenuation was attributed to the low Q factor of resonators (about 100) that results from the use of a PZT 5 like ceramic. On figure 13 a typical example is given. A 8 kHz experimental filter structure is displayed on figure 14. The effect of coupling distance was measured with the same kind of structure. The attenuation is a $1/\lambda$ function only for short distances, then solid angle effects or radiation diagram effects modify the law (Figure 15). The attenuation is a quasi-linear function of pressure down to about 50 mmHg, below the response becomes non linear. The structure can detect pressure well below 10^{-3} torr (Figure 16).

III.2 - Flexural bars

Composite flexural bars permit to achieve good Q factors. The bars can be temperature compensated by the choice of the alloy and heat treatment to better than $\pm 5 \cdot 10^{-6}$ for 10°C . The figure 17 represents typical response of such a resonator. The coupling coefficient can be adjusted by choosing the dimension of the piezo-element between 0 and about 25 %. Figure 18 represents the frequency response of argon coupled bars. The effective attenuation is as low as 4 dB, the 3 dB bandwidth is 175 Hz, that is 3 %. The structure is much smaller than a classical mechanical filter and, a fortiori, than a LC filter. Use of different termination resistances (the filter acts as an impedance transformer) is not a drawback in this frequency range. Attenuations of some dB are generally not constraining in this frequency range since an extra gain is generally available at low cost in the systems. However if less losses are needed or higher order filters required, the same type of resonators can be used to obtain lattice or ladder "classical" filters [12]. The Figure 19 represents the frequency response of a 3 resonators Ladder filter.

X + 5° flexural bars can be used in air coupled structures as can be seen on Figure 20 but with little advantage. It is difficult to fix them rigidly at a short distance and to lower the input-output parasitic capacitance. A preliminary investigation was made using two AT quartz plates ($f_r = 2.6 \text{ MHz}$) at a distance of about .5 mm; surprisingly there was a noticeable response as shown on figure 21. It seems that the coupling effect occurs by the small normal component of vibration that exist, particularly in contoured resonators, rather than by a shear viscosity effect. Further investigations are needed on this point.

IV - GAS COUPLED RESONATORS AS PRESSURE TRANSDUCER

The model have pointed out 2 important features of gas coupled resonators:

- 1) the linearity of the transmission - pressure relation
- 2) the absence of temperature effect due to the gas on response.

Figure 22 represents in linear coordinate the V_8/e response of two composite flexural bars coupled by air; we can see that the relation between attenuation and pressure is quasi-linear down to about 50 mmHg; then, as in figure 16, the relation becomes non linear. Gas coupled flexural discs can also be used as pressure transducers: the figure 23 represents the $V_8(P)$ relation for such a transducer. A polynomial least square fitting of the experimental data was obtained and is

displayed on the figure.

Although not strictly necessary, the flat response is also an advantage for pressure transducer use, it avoids temperature effect due to the resonators which translate the curves in frequency. However the composite resonators can easily be compensated by choice of the alloy and heat treatment, down to a value less than 1 ppm/°C. The main practical concern is to obtain resonators with sufficiently temperature independent Q factors and to fix them at an invariable distance ; this can be solved easily with a good technology. The attracting features of that type of transducer are :

- the high value of electrical signal obtained
- no need of reference pressure
- a large pressure range (virtually no limit for high pressure, lower limit in the order of 10^{-4} torr).

V - CONCLUSION

It has been shown that, at least at low frequencies, gas coupled mechanical two poles filters can be obtained with good performances low volume and low cost. Higher order filter can also be obtained by including intermediate air coupled mechanical resonators between the two end transducers.

The dependance of filter characteristics on constitutional parameters and the specific feature of gas coupling were obtained by modelisation. The case of gas coupled AT quartz needs to be investigated in depth. The methodology used in this paper can also be applied to study the energy transfer in gas between a quartz plate and its pressurized crystal holder.

Gas coupled resonators appears also to be an interesting absolute pressure transducer.

Acknowledgments

The authors want to acknowledge the skilful assistance of D. Servajean for experimental work and the artful aid of Mrs Jouanin and MM. Gaillard and Chenebault for cutting and polishing the many samples used in this study. Special thanks are due to Mrs M.P. Louis for preparing this manuscript.

References

- [1] M. GREENSPAN
Sound wave in Gas at Low Pressures-Physical Acoustics vol 2A; W.P. Mason ed. Academic Press 1965 (p 32)
- [2] E.K. SITTIG
Design and Technology of Piezoelectric Transducers for Frequencies above 100 MHz-Physical Acoustics vol 9 W.P. Mason ed. Academic Press 1972
- [3] R. HOLLAND, E.P. EER NISSE
Design of Resonant Piezoelectric Devices MIT Press 1969
- [4] W.P. MASON
Piezoelectric crystals and their application to ultrasonics VAN Nostrand 1950
- [5] D.F. SHEAHAN, R.A. JOHNSON
Crystal and Mechanical Filters IEEE CAS 22 p 69-89 Feb 1975
- [6] O.C. ZIENKIEWICZ
The Finite Element Method Mc Graw Hill London 1977

- [7] H. ALLIK, T.J.R HUGHES
Finite Element Method for Piezoelectric Vibrations
Int. J. Num Meth. vol 2 p 151 1970
- [8] Y. KAGAWA ; G.M.L GLADWELL
Finite Element Analysis of Flexure type Vibrators With Electrostrictive Transducers
IEEE Su 17 n° 1 p 41 1970
- [9] R.G. ANDERSON, G.M. IRONS, O.C. ZIENKIEWICZ
Vibration and Stability of plates Using Finite Element
Int. J. Solids Structures vol 4 p 1045 1968
- [10] H. CARRU
Three Dimensional 'Sandwich' Transducer Modelling by finite Element method
Electronics Letters vol 15 n° 24 p 805 1979
- [11] C. DUCHET ; G. VILLELA
Length expander transducers for mechanical filters of composite metal-lithium niobate structure
Proc. of 1979 ISCAS p 900 IEEE cat n° 79 CN 1421-7CAS
- [12] G. FERRIEU ; P. AMSTUTZ
A New Principle to Avoid undesirable Oscillations in Electro acoustic loop : Application to the design of a Hands-free Telephon set without Voice Switching
Proc. of ICASSP p 1448 IEEE cat. 82 CH 1746-7(1982)

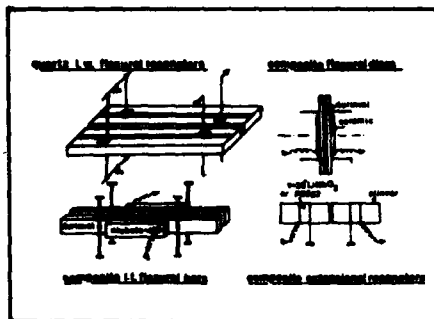


Figure 1 : Structures for gas coupled filters

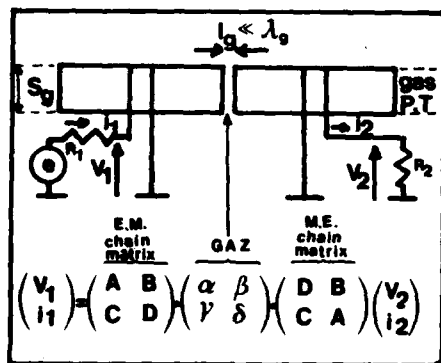


Figure 2 : Air coupled composite extensional turn

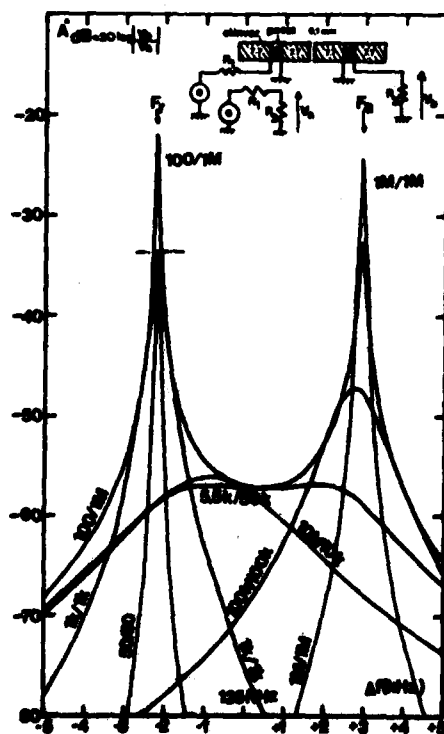


Figure 3 : Computed insertion loss (Q=4000)

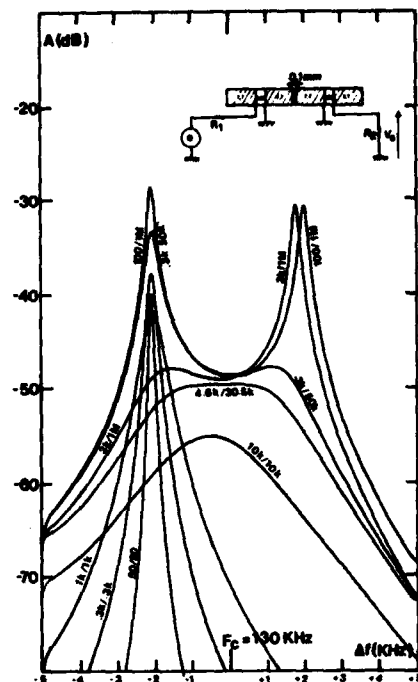


Figure 4 : Experimental effective attenuation

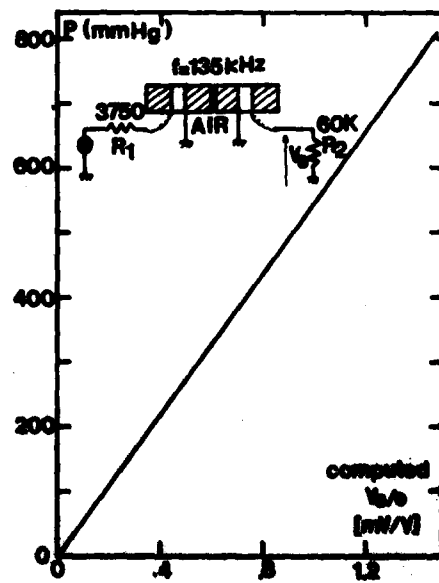


Figure 5 : Computed A (P) relation

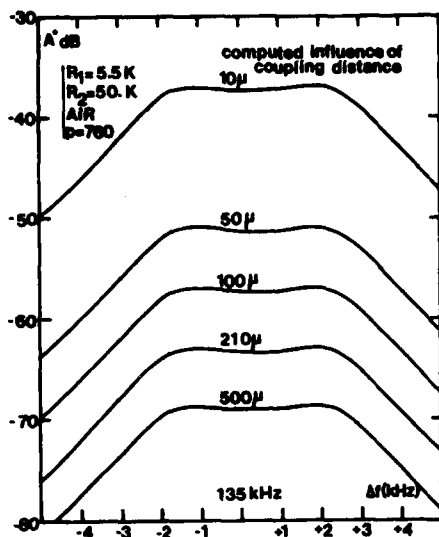


Figure 6 : Computed influence of Δr

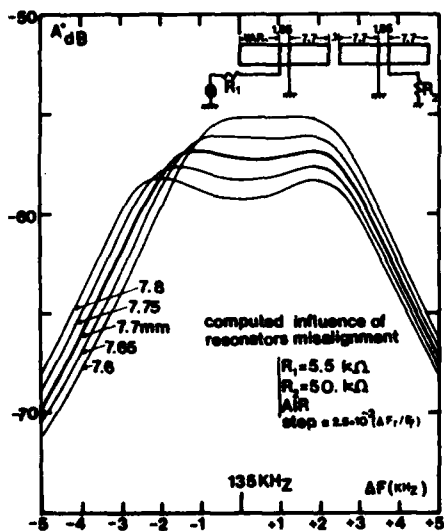


Figure 7 : Effect of frequency misalignment

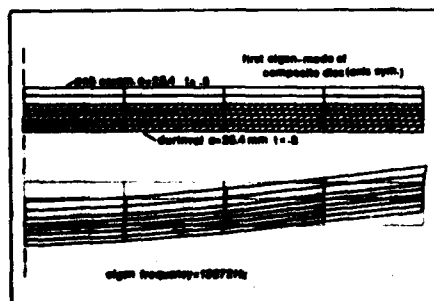


Figure 8 : Computed 1st mode of composite disc

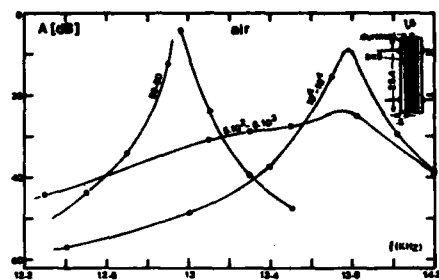


Figure 9 : Computed effective attenuation of air-coupled composite flexural discs

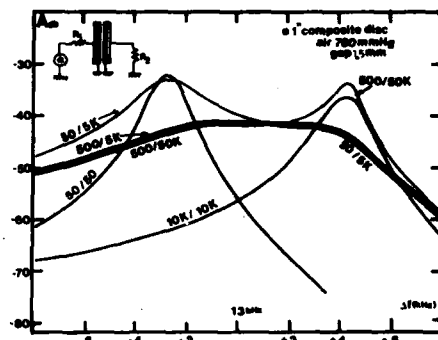


Figure 10 : Experimental attenuation of air-coupled composite discs

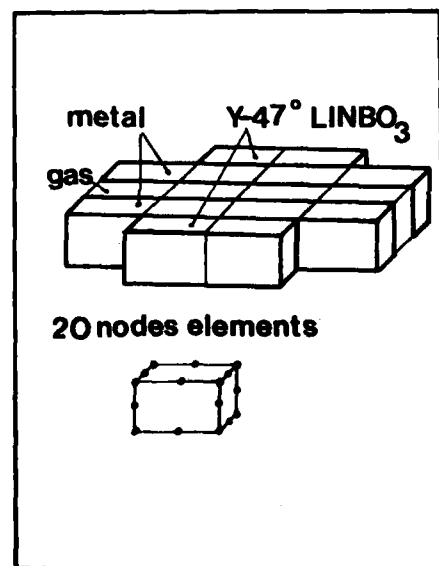


Figure 11 : Discretization for the 3.D. Finite Elements Analysis

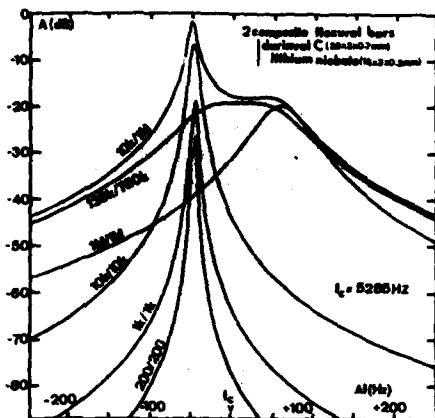


Figure 12a : Experimental attenuation of gas coupled flexural bars



Figure 14 : Experimental air coupled discs

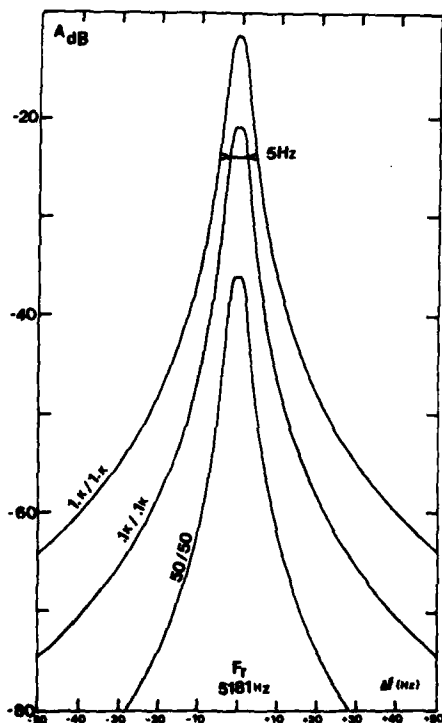


Figure 12b : Equal resistance bandpass
(@ factor of resonators is 5000)

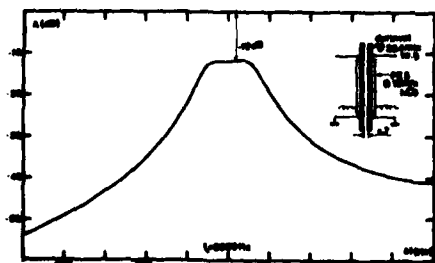


Figure 13 : Effective attenuation of thin air coupled flexural discs (P=760 mm)

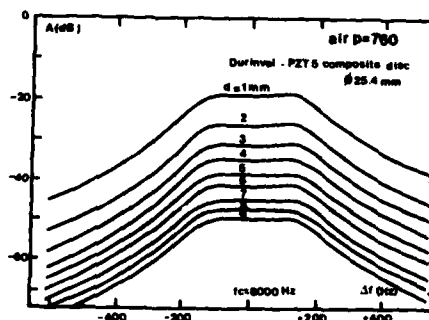


Figure 15 : Effect of distance on A

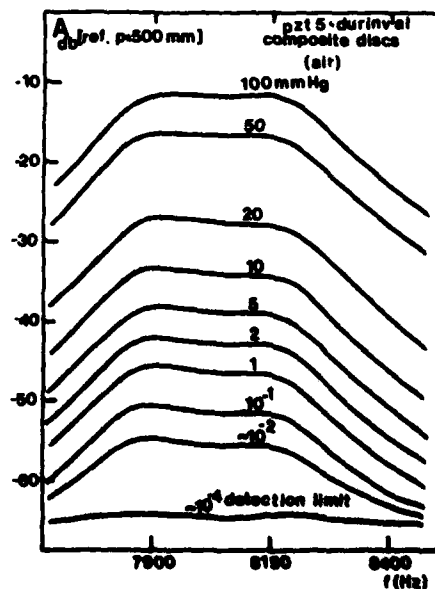


Figure 16 : Effect of Pressure on A
(Low Pressures)

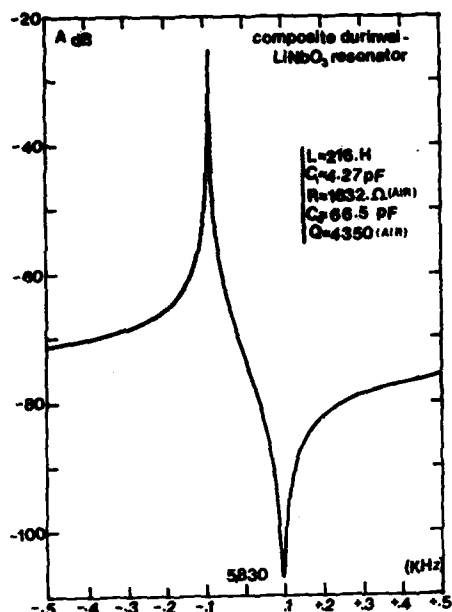


Figure 17 : Composite d.t. flexural resonator

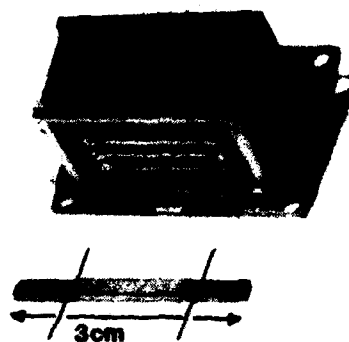


Figure 19b : 3 resonators ladder filter
(length of resonators is 3 cm.)

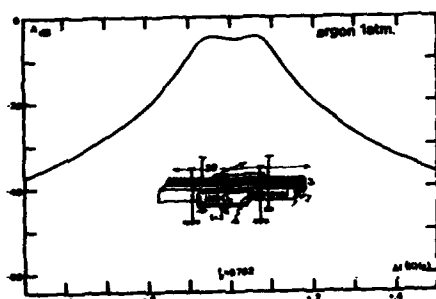


Figure 18 : Best results for Argon coupled flexural bars

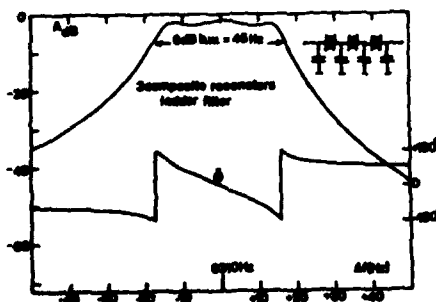


Figure 19a : "Classical" ladder filter using composite resonators

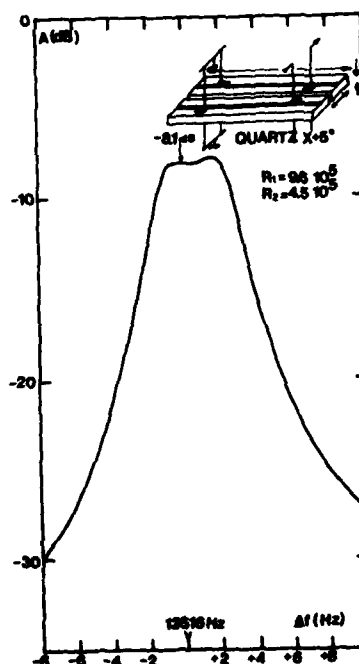


Figure 20 : Air coupled X+5° flexural quartz

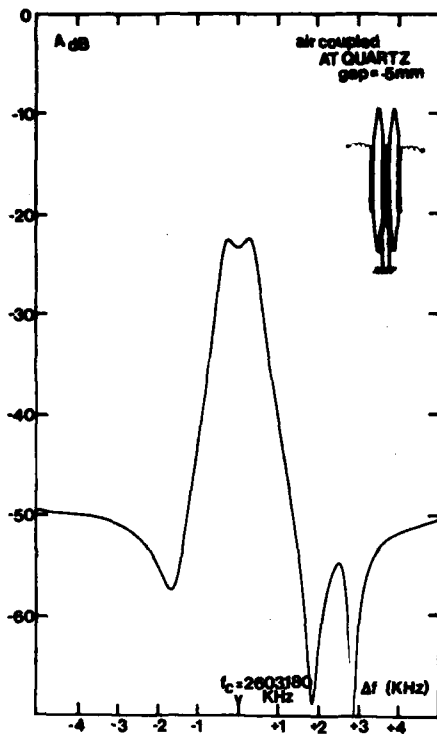


Figure 21 : Response of air coupled AT quartz

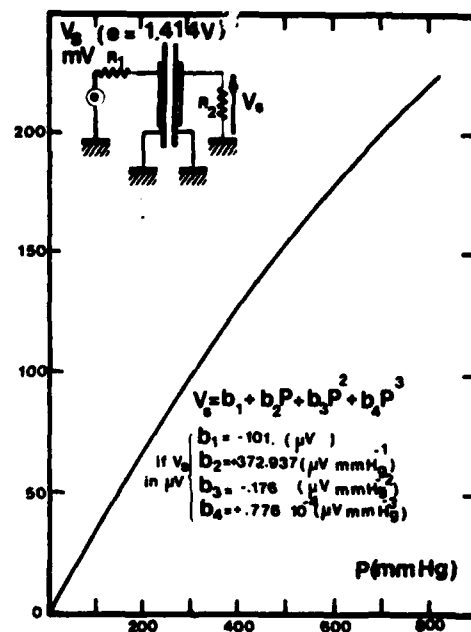


Figure 23 : Pressure-Attenuation relation for air

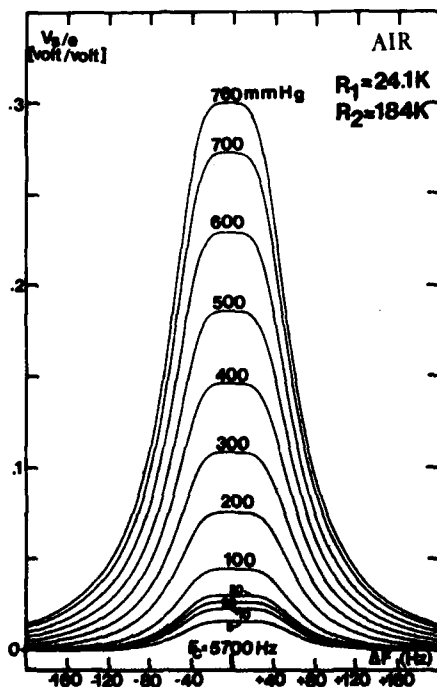


Figure 22 : Attenuation-Pressure relation for air coupled flexural bars

SHIH S. CHUANG

STATEK CORPORATION
Orange, CA

Abstract

The purpose of this paper is to present designs and experimental test results of double-ended tuning fork quartz crystals vibrating in the fundamental and overtone flexure modes. These double-ended tuning fork quartz crystals were produced using photolithographic, chemical etching and laser frequency adjustment manufacturing techniques.

The motions of the fundamental and overtone flexure modes of a double-ended tuning fork are described. Electrode designs which depend on the crystal vibrational mode are used to optimize crystal parameters.

The first order frequency-force coefficient of a double-ended tuning fork quartz crystal vibrating in the fundamental flexure mode is higher than that of the crystal vibrating in the higher order flexure modes. Experimental results show the ratio of these coefficients is approximately 2.3 and 3.2 respectively for the ratio of the coefficient of the fundamental mode to that of the first and second overtone modes respectively. These ratios are smaller than the frequency ratios between the overtone modes and the fundamental mode.

The frequency temperature characteristic of the double-ended tuning fork quartz crystal vibrating in flexure modes is very similar to that of the single-ended tuning fork quartz crystal. The turnover temperature of the higher order flexure mode is higher than that of the lower order flexure mode.

The experimental results of the crystal parameters, including the motional resistance, Q , the frequency temperature characteristic, and the frequency-force characteristic of double-ended tuning fork quartz crystals are shown and discussed.

Introduction

A double-ended tuning fork quartz crystal has been known as a force sensor^{1,2,3}, and is potentially useful in many sensor applications, there is no published detailed analysis or description of the crystal designs and parameters. Using photolithographic and chemical etching manufacturing techniques⁴, a better control in the precision of crystal geometrical dimensions is provided. These techniques also create more flexibility for a crystal designer to use more complex crystal geometry to optimize the double-ended tuning fork quartz crystal parameters.

The purpose of this paper is to present designs and experimental test results of a double-ended tuning fork quartz crystal vibrating in the fundamental and overtone flexure modes. It will show that this crystal is not only useful as a force sensor, but also useful in frequency control and filter applications.

In the following sections, the motions of the fundamental and overtone flexure modes of a double-ended tuning fork will be described. Electrode designs which depend on the crystal vibrational mode are used to optimize crystal parameters. The first order frequency-force constant relating to the crystal design will be shown. The experimental results of the crystal parameters will be presented and discussed.

Flexure Motion of Double-Ended Tuning Fork

To analyze and obtain a complete analytic function to describe the flexure motion of a double-ended tuning fork is not simple. In practice, the result of the analysis of the flexure motion of a beam clamped on both ends can provide essential important information on the motion of the tines and the frequency of the double-ended tuning fork for a first approximation.

Figure 1 shows a typical double-ended tuning fork structure. The relative displacement of the tine as a function of the position along the tine length direction can approximately be described by the following equation⁵:

$$U(X_1) = P \left\{ \left(\cosh m \frac{X_1}{L} - \cos m \frac{X_1}{L} \right) - \left(\sinh m \frac{X_1}{L} - \sin m \frac{X_1}{L} \right) \right\} \quad (1)$$

$$\text{where } P = \frac{\sinh m - \sin m}{\cosh m - \cos m}$$

and m is a constant which is determined by the equation:

$$\cosh m \cdot \cos m = 1 \quad (2)$$

For different flexure modes, the eigenvalue m has the following values:

flexure mode	m
fundamental	4.73
first overtone	7.86
second overtone	10.99

Figure 2 shows the lateral displacement of $U(X_1)$ of the fundamental and overtone flexure modes.

The frequency of the flexural motion is approximately:

$$f = f_m (1 - \alpha_m T)^{1/2} \quad (3)$$

$$\text{where } f_m = \frac{1}{2\pi} \sqrt{\frac{1}{12\rho S} \frac{m^2 w}{L^2} (1 + K_{cm} (\frac{w}{L})^2)} \quad (4)$$

is the frequency of the flexure motion without external force,

α_m = frequency force constant

ρ = mass density

S = elastic compliance constant

T = external applied force on the crystal along the length direction

w = tine width

L = tine length from crotch-to-crotch

t = tine thickness

K_{cm} = a constant

When no external force is applied to the double-ended tuning fork, and w/L is very small, the frequency ratios of the overtone flexure modes, with respect to

AD P 002483

the fundamental flexure mode, are 2.8 and 5.4 respectively for first and second overtone flexure modes.

Equation (3) has shown that the frequency of a double-ended tuning fork crystal can be changed by the external force T . If we expand Equation (3) to:

$$f = f_m(1 + \alpha_{m1} T + \alpha_{m2} T^2 \dots) \quad (5)$$

the first order frequency-force constant α_{m1} is approximately⁶:

$$\alpha_{m1} = h_{m1} \frac{l^2 S}{m_1 w^3 t} \quad (6)$$

This constant is proportional to the square of the tine length, and inversely proportional to the thickness and third power of width. The orientation of the quartz crystal will also effect the α_{m1} through the elastic compliance constant, S ; h_{m1} is approximately equal to 0.0735, 0.036 and 0.022 m_1 for fundamental, first and second overtone modes respectively.

The first order frequency-force constant (α_{m1}) for the lower order flexural mode will be larger than the higher order flexural mode. The ratio of these constants is approximately 2.0 and 3.34 respectively.

Electrode Design

Electrodes are placed on the piezoelectric resonator to obtain the maximum electrical mechanical coupling. In the case of a double-ended tuning fork quartz crystal the electrode design will depend on its geometry and the crystal orientation.

If the quartz crystal orientation is (ZYw) θ , the tine length is almost in the y direction, width is in the X direction. Electrodes are arranged so that the electric field, which will contribute to the piezoelectric effect, is mainly from the field in the X direction. It can be shown that the stress related to the piezoelectric effect is approximately proportional to the second derivative of the lateral displacement $U(Y)$ of the tine with respect to Y.

Figure 3 shows the relative stress as a function of the position along the tine length direction for the fundamental and overtone flexure modes. The nodal points shown in the figure will be the approximate locations in a practical double-ended tuning fork. The actual locations will depend on the detailed geometry of the crystal. However, the locations shown in Figure 3 provide the initial positions for the design.

Figure 4 illustrates examples of the electrode designs of the fundamental and overtone flexure modes of a double-ended tuning fork. Since there is more than one reverse, the stress along the length of the tine, the electrical polarity change of the electrodes along the tines is necessary. The positions where the polarity of the electrodes change are at the minimum, or zero stress points, along the length of the tines.

Experimental Results

The double-ended tuning fork quartz crystals were produced by using photolithographic, chemical etching manufacturing techniques. In the following experiments the crystal electrode design is described in the previous section. The crystal orientation is (ZYw) θ .

Figure 5 shows an example of the double-ended tuning fork quartz crystal.

Crystal Parameters of Fundamental and Overtone Flexure Modes

To demonstrate the crystal parameters of fundamental and overtone flexure modes of a double-ended tuning fork quartz crystal, the following crystal orientation and dimensions were chosen:

Crystal orientation:	(ZYw) 2°
Tine width:	10 mil
Tine length:	227 mil
Thickness:	5 mil

The electrode designs used for these crystals are as shown in Figure 4.

Table 1 shows the list of typical crystal parameters of the crystals.

The frequency ratios of the first and second overtone flexure modes, with respect to the fundamental flexure mode of the crystal, are approximately 2.67 and 5.13 respectively. The values are slightly smaller than the calculated values based on simplified analysis as described in the previous section.

The quality factor of the crystal operating at higher overtone flexure mode is higher than that operating at fundamental mode. The quality factor of the double-ended tuning fork quartz crystal is much lower than that of the single-ended tuning fork operated in flexure mode at the identical frequency.

The motional capacitance of the double-ended tuning fork quartz crystal is much higher than the single-ended tuning fork, since it requires approximately 2.5 times more tine length to obtain the same frequency as with the single-ended tuning fork.

The first order frequency force constant of the experimental value is very close to the calculated value based on Equation 6. As predicted by the theoretical calculations, this constant for the crystal vibrating at the fundamental flexure mode is higher than that at the higher overtone flexure mode. However, the ratio of the constant is approximately 2.3 and 3.2 respectively.

The frequency temperature characteristics of a double-ended tuning fork quartz crystal are very similar to those of the single-ended tuning fork quartz crystal⁶ operating at the flexure mode. The frequency-temperature relation approximately follows a parabolic function. The turnover temperature of the higher order flexure mode is higher than the lower order flexure mode.

Effect of Crystal Geometry on the Frequency-Force Constant

As shown in the previous section, the first order frequency-force constant α_{m1} of a double-ended tuning fork crystal is approximately proportional to $l^2/(w^3 t)$, where l , w , t are length, width and thickness of the crystal tine respectively. The following experiments will show this relation.

Figure 6 shows the experimental result of the first order frequency-force constant, α_1 , as a function of the crystal thickness. In this experiment the tine width and tine length are 10 mil and 227 mil respectively. The constant, α_1 , decreases from 223 ppm/gram to 71 ppm/gram as the crystal thickness is increased from 3 mil to 10 mil. The curve in Figure 6 very closely follows Equation 6.

The α_1 , as a function of the tine length, was performed with identical crystal designs at 5 mil thickness. As the tine length is changed from 227 mil to 88 mil, α_1 decreases from 135 ppm/gram to 22 ppm/gram which closely follows the square of the length relation.

The α_1 , as a function of the tine width, was determined with two different tine widths, 5.6 and 8.8 mil. The crystal tine length is 187 mil and the thickness is 5 mil. The frequency-force constants, α_1 , are 160 ppm/gram and 500 ppm/gram for width 8.8 mil and 5.6 mil respectively. The ratio of those two α_1 's is approximately inversely proportional to the cube of the width ratio.

The Effect of Crystal Orientation

The first order frequency-force constant of a double-ended tuning fork quartz crystal will depend on the crystal orientation. Equation 6 shows the constant is linearly proportional to the elastic compliance constants. The elastic compliance constant in the length direction will be changed when the orientation of the crystal is changed.

Figure 7 shows an example of the experimental result of the first order frequency-force constant as a function of the crystal orientation angle θ , where θ is the angle designated by $(ZYW)\theta$; i.e., the angle of rotation around the X-axis of the quartz crystal. The crystal thickness, tine length, and tine width are 5, 227, and 10 mil respectively. All the crystals were chemically etched in the same process and at the same quartz etch time. There are some variations in the shape of the crystal tine cross-sections due to the anisotropic nature of the quartz. α_1 decreases as θ goes to positive angle, and increases as θ goes to negative angle. The overall change of α_1 is more than 30% when θ changes from -20° to 120° .

Figure 8 shows the crystal frequency as a function of the angle θ . The frequency decreases as angle θ is increased positively. This relation is very close to the result of the theoretical calculation.

Figure 9 shows the turning point temperature of the crystal as a function of the crystal orientation angle θ . This relation is very close to the conventional tuning fork quartz crystal.

The Effect of Base Configuration

The base configuration of a double-ended tuning fork quartz crystal will have some effect on the crystal parameters because of the change of the boundary condition between the motion of the tines and the coupling of the base.

The following experiment was performed to investigate the effect of the base shoulder width d_1 and d_2 near the crystal crotch areas. The crystal tine length, tine width and thickness are 227 mil, 10.2 mil and 5 mil respectively. The crotch width $2d_c$ is 5.6 mil.

Figure 10 shows the first order frequency-force constant as a function of d_1/d_c .

There are two cases: $d_1=d_2$ and $d_2=0$. Curve A shows when $d_1=d_2$; Curve B shows $d_2=0$. When comparing the value of α_1 of Curve B to that of Curve A, the value of α_1 is smaller when $d_2=0$. For both curves, A and B the constant of α_1 increases as d_1/d_c ratio is increased, and then reaches a maximum when d_1/d_c is approximately equal to one (1). As d_1/d_c is increased further, α_1 decreases.

Experimental results show the effect of the base configuration on the frequency-force relation. It not only effects α_1 , but also the linearity of the frequency-force relation, especially when the external force is very small. The external force will shift the effective stress nodal line of the flexure motion of the crystal in the lateral direction, especially when the base to the tines configuration is not symmetrical. The value of α_1 tends to decrease and is less than the calculated value based on Equation 6 if d_1 or d_2 is not near or equal to d_c .

Dynamic Range of Force Sensor

The maximum frequency change under a force T_{max} of a double-ended tuning fork quartz crystal is $T_{max} \cdot \alpha_1$, where T_{max} is the maximum axial force applied to the crystal. It can be shown that $T_{max} \cdot \alpha_1$ is proportional to the square of thickness-to-width ratio.

Figure 11 illustrates an example of the crystal frequency as a function of the axial force applied to a crystal. In this example the maximum frequency shift is approximately 4.4 kHz when the force is 760 grams.

Dynamic range will be reduced if there is a coupling between the main flexure mode to the other modes⁷. This coupling can decrease the Q of the crystal and causes the frequency instability, and sometimes crystal ceases to operate.

Table 2 lists examples of three different crystal designs and their dynamic ranges. The dynamic range can easily be increased if the thickness-to-width ratio is increased.

Selection of Flexure Modes

In the previous section the double-ended tuning fork quartz crystal operating at the fundamental and overtone flexure modes was described. For the force-sensor application, if sensitivity is the most important factor, the fundamental flexure mode is certainly a choice. However, in certain cases, if higher Q and higher frequency are required, then the overtone flexure modes provide an alternative choice.

The double-ended tuning fork quartz crystal may provide an alternative design in frequency control and filter applications in frequencies below the MHz range. In certain applications, if higher C_1 is required, the double-ended tuning fork will be better than the single-ended tuning fork⁸.

Summary

In this paper some design parameters and experimental results of a double-ended tuning fork quartz crystal have been shown. With photolithographic, chemical etching manufacturing techniques, one can produce a subminiature double-ended tuning fork quartz crystal. A frequency-force constant greater than 500 ppm/gram can be easily produced by a miniature double-ended tuning fork quartz crystal. With this high sensitivity, this crystal may be very useful in sensor applications.

It has also been shown that, with the combination of the fundamental and overtone flexure mode of a double-ended tuning fork quartz crystal, this crystal design can be useful in the frequency control and filter applications.

References

1. Erdley, H. F., U.S. Patent #3,238,789, March 8, 1966.
2. Serra, Norman R., AGARD Conference Proceedings #43, Inertial Navigation System and Components, Oxford, England, 1967, and Brunswick, Germany, 1968.
3. EerNisse, Errol P., U.S. Patent #4,215,570, August 5, 1980.
4. Staudte, J. H., Proceedings of the 27th Symposium on Frequency Control, 1973, pp50-54.
5. Rayleigh, J. W., The Theory of Sound, Vol.1, Dover Publications, Second Edition, 1945, pp255-305.
6. Albert, W., Proceedings of International Instrumentation Symposium, Las Vegas, Nevada, 1982, pp33-44.
7. EerNisse, Errol P., and Paros, J. M., U.S. Patent #4,372,173, February 8, 1983.
8. Chuang, S., Proceedings of the 35th Symposium on Frequency Control, 1980, pp130-143.

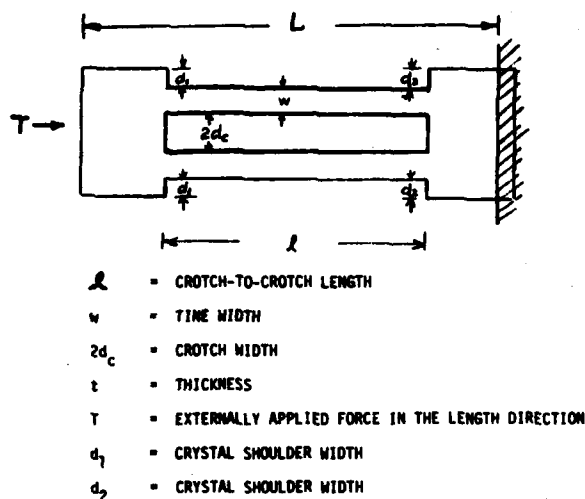


FIGURE 1:
AN EXTERNAL FORCE APPLIED TO A
DOUBLE-ENDED TUNING FORK CRYSTAL

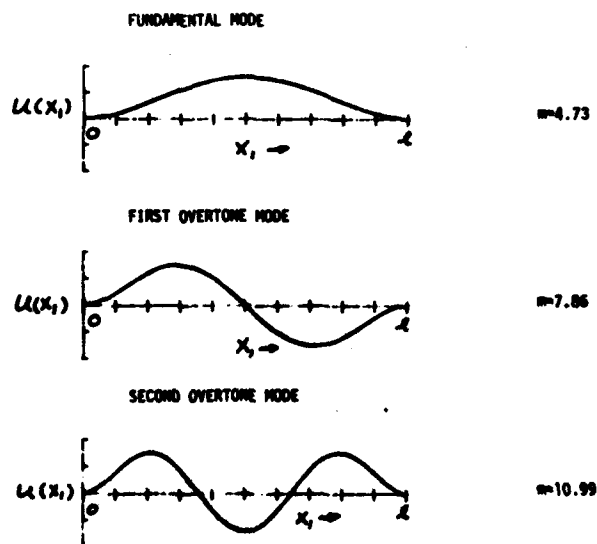


FIGURE 2:
RELATIVE LATERAL DISPLACEMENT OF THE CRYSTAL
TINE AS FUNCTION OF THE POSITION OF THE TINE
FOR DIFFERENT FLEXURE MODES.

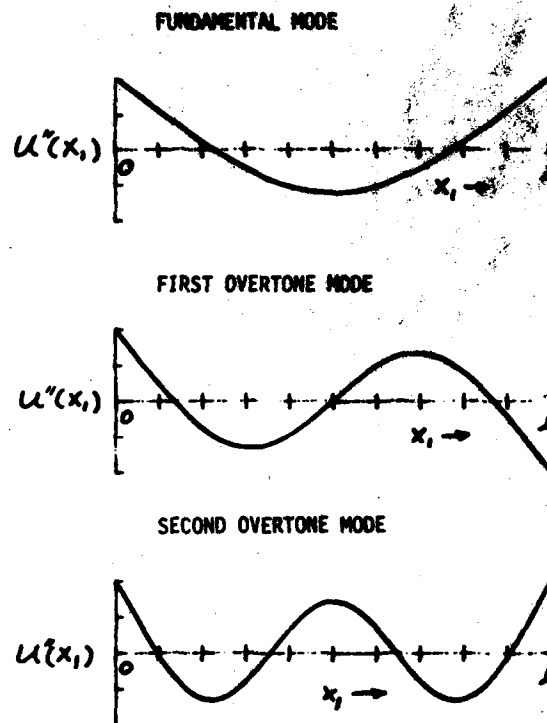


FIGURE 3:
RELATIVE STRESS OF TINE AS FUNCTION OF THE
POSITION OF TINE FOR FUNDAMENTAL AND OVER-
TONE FLEXURE MODES. QUARTZ CRYSTAL ORIE-
NTATION IS (ZYW)θ; ELECTRICAL FIELD IS IN
X DIRECTION.

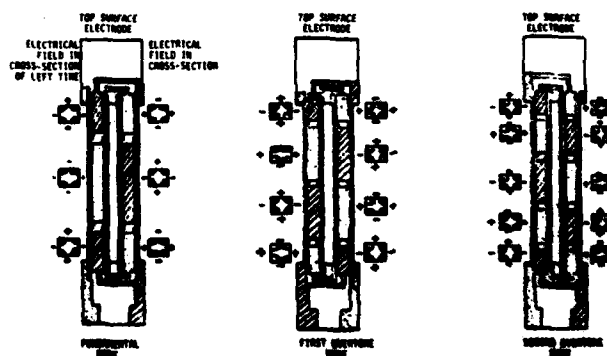


FIGURE 4:
ELECTRODE DESIGN A OF FUNDAMENTAL AND OVERTONE
FLEXURE MODES OF A DOUBLE-ENDED TUNING FORK CRYSTAL.

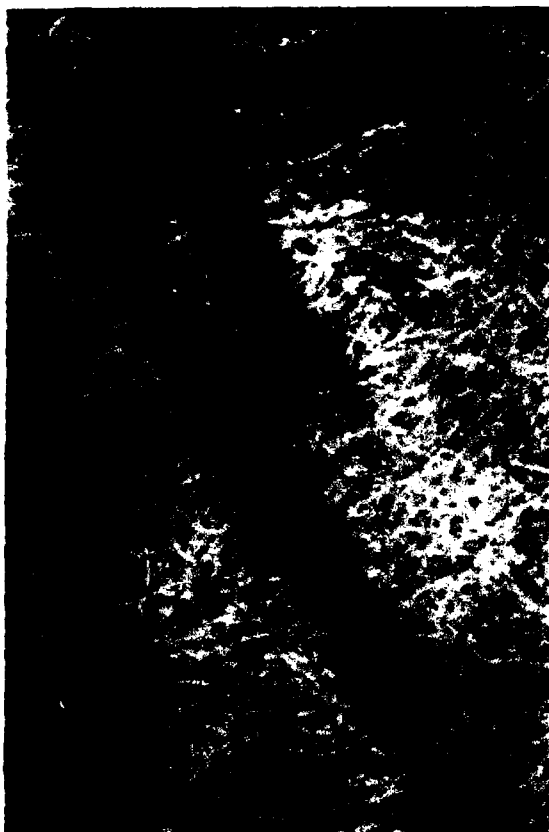


FIGURE 5.

AN EXAMPLE OF THE DOUBLE-ENDED TUNING FORK QUARTZ CRYSTAL MANUFACTURED BY USING PHOTO-LITHOGRAPHIC, CHEMICAL ETCHING TECHNIQUES

TABLE 1

	FUNDAMENTAL MODE	FIRST OVERTONE MODE	SECOND OVERTONE MODE
FREQUENCY (kHz)	41	109.9	211
Q	50×10^3	75×10^3	100×10^3
C_1 (pF)	7	7	7
R_1 (ohm)	11	2.8	1
$C_{a/1}$	289	280	280
α_1 PPM/GRAM	-135	-58	-39
α_2 10^{-7} /GRAM	5.1	-4	1
α_3 10^{-9} /GRAM	1.1	1.2	-0.3
TUNING POINT TEMP.	42°C	52°C	63°C

COMPARISON OF THE CRYSTAL PARAMETERS OF FUNDAMENTALS, FIRST AND SECOND OVERTONE FLEXURE MODES, OF A DOUBLE-ENDED TUNING FORK QUARTZ CRYSTAL.

CRYSTAL DESIGN

TINE WIDTH 10 mil
TINE LENGTH 227 mil
THICKNESS 5 mil
QUARTZ CRYSTAL ORIENTATION $(27^\circ)2^\circ$

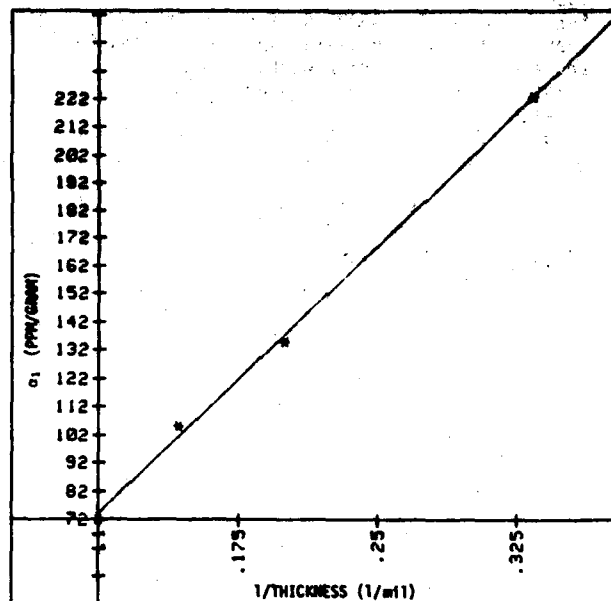


FIGURE 6.

FREQUENCY-FORCE CONSTANT, α_1 , AS FUNCTION OF CRYSTAL THICKNESS, t

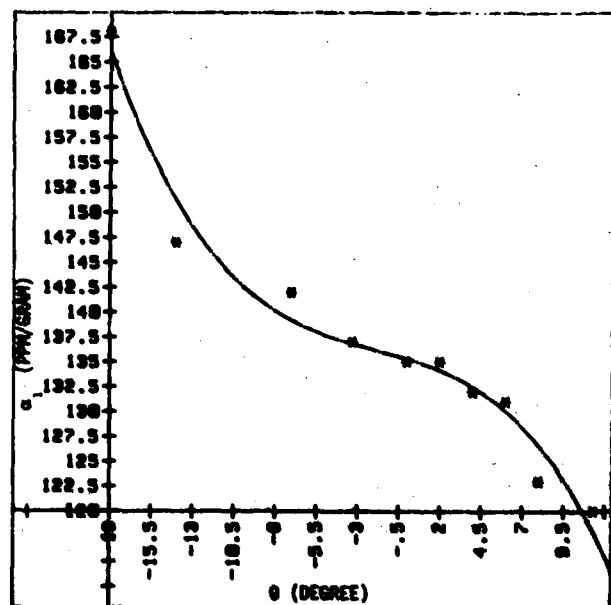


FIGURE 7.

THE FIRST ORDER FORCE FREQUENCY CONSTANT α_1 , AS A FUNCTION OF CRYSTAL ORIENTATION ANGLE θ

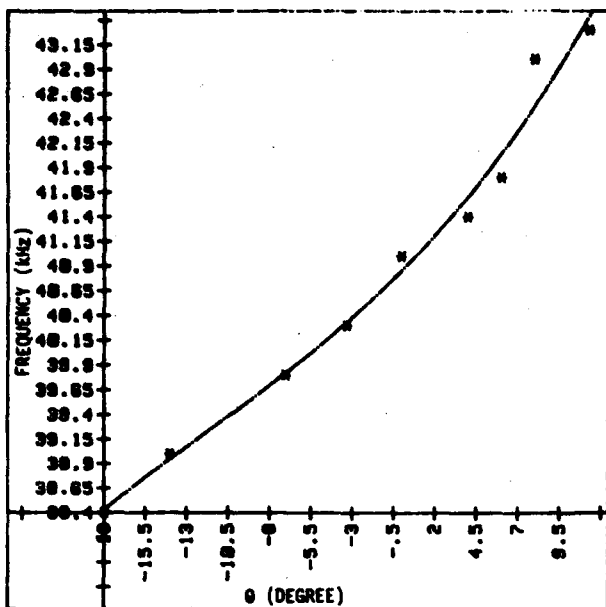


FIGURE 8.
FREQUENCY OF A DOUBLE-ENDED TUNING FORK
QUARTZ CRYSTAL AS A FUNCTION OF THE
CRYSTAL ORIENTATION ANGLE θ

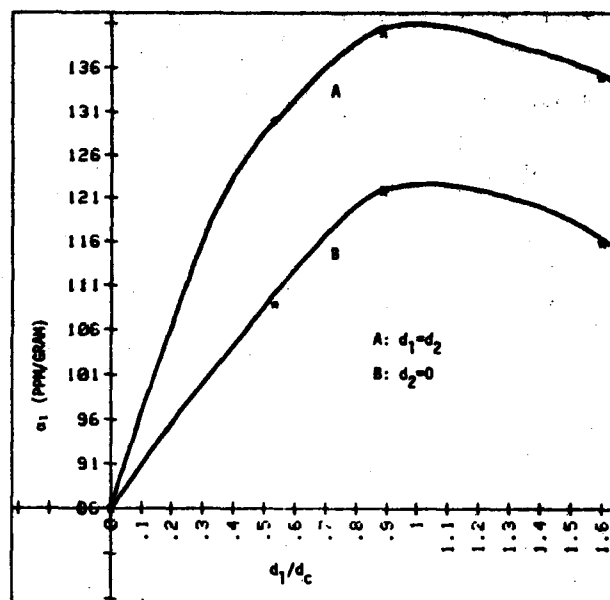


FIGURE 10.
FREQUENCY-FORCE CONSTANT, α_1 , AS FUNCTION OF
THE d_1/d_c ; $2d_c$ = CROTCH WIDTH; d_1, d_2 = SHOULDER
WIDTH

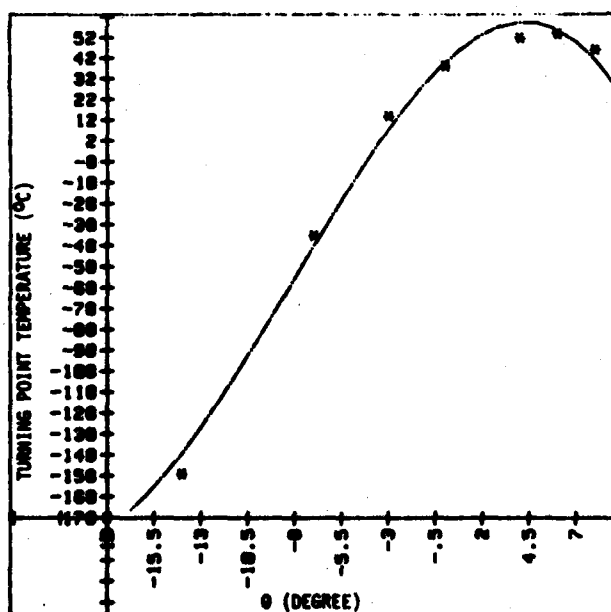


FIGURE 9.
TURNING POINT TEMPERATURE OF A DOUBLE-ENDED
TUNING FORK QUARTZ CRYSTAL AS A FUNCTION OF
THE CRYSTAL ORIENTATION ANGLE θ

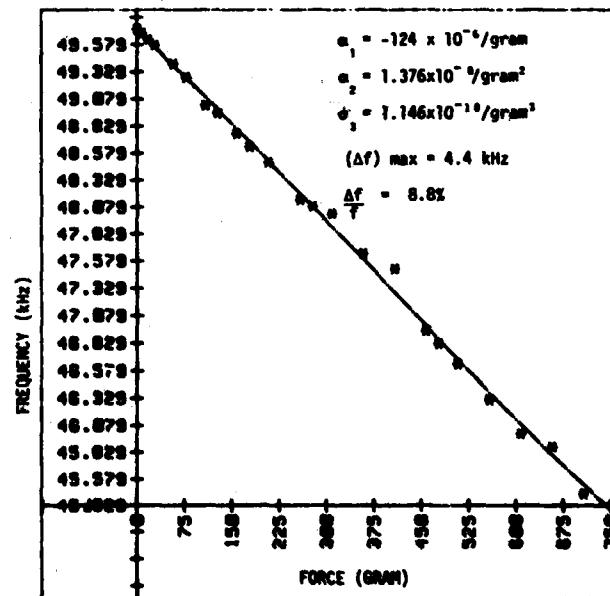


FIGURE 11.
FREQUENCY AS A FUNCTION OF THE AXIAL FORCE
APPLIED TO A DOUBLE-ENDED TUNING FORK
QUARTZ CRYSTAL

$w = 8.8 \text{ mil}$
 $t = 7 \text{ mil}$
 $\lambda = 187 \text{ mil}$

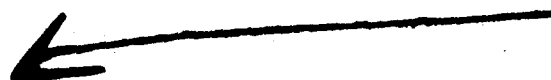
TABLE 2

MAXIMUM FORCE T_{max}

Maximum Force: $T_{max} = K_c \cdot \frac{t^3 w}{L^3}$

$(\frac{\Delta f}{f})_{max} = T_{max} \alpha_1 = K(\frac{t}{w})^2 \cdot (\frac{t}{L})^2$

CRYSTAL DIMENSION	T_{max} (GRAM)	α_1 PPM/GRAM	$(\frac{\Delta f}{f})_{MAX}$	$(\frac{\Delta f}{f})_{MAX}$ (kHz)
w = 10 mil t = 5 mil s = 227 mil L = 370 mil	250	135	2.9%	1.2
w = 8.8 mil t = 7 mil s = 187 mil L = 274 mil	800	124	8.8%	4.4
w = 5.6 mil t = 5 mil s = 187 mil L = 274 mil	170	500	6.4%	2.3



PRACTICAL CONSIDERATIONS FOR MINIATURE QUARTZ RESONATOR FORCE TRANSDUCERS

Errol P. Eernisse
Quartex, Inc.
1020 Atherton Dr., Bldg. C
Salt Lake City, UT 84107

Jerome M. Paros
Paroscientific, Inc.
4500 148th Ave. N.E.
Redmond Wa 98052

SUMMARY

A new device called the double-ended tuning fork has been examined for practical use as a force transducer. A number of problems unique to the use of the device are treated. Design rules, important considerations for applications, and an example of typical performance are presented.

INTRODUCTION

A recently introduced force transducer¹ has been studied in detail for practical applications. The device is a double-ended tuning fork wherein a longitudinal force causes frequency shifts through effects identical to the tuning of the frequency of a vibrating string. Frequency changes can be monitored and then reduced with simple equations, thus making the device an inherently digital, high resolution, low power consumption force transducer. Optimum design involves frequency shifts as large as 10% of the zero-force frequency. It has been found that such large frequency shifts cause some unique problems.

Several spurious modes exist that can result in activity dips if the desired frequency crosses one of the spurious mode frequencies as force varies. The spurious modes have been modeled approximately and confirmed by finite element calculations. Experiments have been performed to locate the modes and support the theories. The spurious modes have force-sensitive frequencies, which complicates the design of an activity-dip-free device.²

Unequal loading of the tines causes the natural frequencies of the tines to separate, which causes loss of Q and eventually bistable operation of the device in an oscillator circuit. In many cases oscillation ceases. Practical methods are shown to avoid the problem.³

The temperature behavior of the devices is modified from that of watch tuning forks by the thermal expansion of the mechanical structures used to apply forces. Some design considerations are included.

Finally, a practical application is given where large forces are measured with a compound lever system with resolution of better than 1/400,000 of full scale.⁴

MODE ANALYSIS

The device is shown in Fig. 1. It can be made by photolithographic techniques much in the same way a watch crystal is fabricated.^{1,5} The ends are enlarged for two reasons: attachment stresses are reduced and

the length L is well-defined independent of the actual attachment boundaries.² Several modes exist near the frequency of the desired tuning fork mode. These are seen in Fig. 2 for a typical device. Mode 4 is the one actually used by virtue of the symmetry of the electrode pattern. The frequencies are dependent upon the force F aligned along the length L. The solution of the resonant frequency of a beam with longitudinal stress is a transcendental equation.⁶⁻⁹ A simplified polynomial fit of order 1 can be used for most practical design considerations.² If one ignores the slot for Mode 1:

$$f_1 = 1.03 (E/P)^{.5} t \left(1 + \frac{w^2 L^2 S}{t^2 m^2} \right) L^{-2} \quad (1)$$

$$f_2 = 1.03 (E/P)^{.5} t \left(1 + \frac{w^2 S}{t^2} \right) m^{-2} \quad (2)$$

$$f_3 = 1.03 (E/P)^{.5} w (1+S) m^{-2} \quad (3)$$

$$f_4 = 1.03 (E/P)^{.5} w (1+S) m^{-2} \quad (4)$$

where E is Young's Modulus, P is mass density, the dimensions L, m, t, and w are shown in Fig. 1, and S is the force sensitivity relation for Mode 4

$$S = \frac{0.148 F m^2}{E^2 t w w^2} \quad (5)$$

which is the one used for the force measurement.

Finite element analysis has been used to confirm the mode shapes and frequencies. The mode shapes shown in Fig. 2 are from the finite element analysis. The frequencies from the finite element analysis and Eqs. 1-5 are shown in Fig. 2 for F = 0 and the dimensions shown in Fig. 2. Note that the simple theory is adequate for design. There is a small error in the beam calculation of less than 6% for Mode 3 because the simple calculation ignores the contribution from the short distance between the slot ends and the enlarged attachment portions.

An experiment has been performed to verify Fig. 2. The details will be published later. Briefly, the experiment uses a non-contacting fiber optic displacement gauge called a Fotonic Sensor (sold by the MTI Company) and a dual channel Fast Fourier Transform spectrum analyzer whose output gives both rms magnitude and phase angle of

the measured displacement relative to the driving force signal. Devices of dimensions as given in Fig. 2 were attached to a simple lever arm system. The lever arm system was driven with forces from a PZT piezoelectric drive system excited with a white noise source. The white noise drove all modes of the crystal. The displacement gauge was moved along the front of the device by an xyz table to investigate Modes 1 and 2, and along the side of the device to investigate Modes 3 and 4. The lever arm system allowed a static force to be applied simultaneously with the white noise forces so that the change in resonant frequency with force could be monitored. Figure 3 shows the results for Modes 1, 3, and 4. Mode 2 could not be seen because the displacement gauge was wider than both tines and the signals from the out-of-phase motion cancelled. Figure 3 also shows a comparison of Eqs. 1-5 with the experimental data. The general agreement is excellent. The differences are attributed to w being larger than designed due to the faceting that occurs when the devices are etched from the wafer of quartz.

Note in Fig. 3 that the frequencies shift by 10% or more. The dimensions in Fig. 2 have been selected carefully to avoid any activity dips from Mode 4 crossing Modes 1 or 2 as the force increases. This may seem trivial at first glance, but Modes 1 and 2 and their overtones cause severe design constraints, particularly if t is much smaller than w , which was the case for the early designs and is the case for watch crystals formed by chemical etching. A design criteria has been established based on Eqs. 1-5 and additional equations for the overtones of Modes 1-2. It is desirable for f_4 to not equal any other mode frequency. This can be accomplished if the ratios L/m and t/w are chosen appropriately. The following inequalities must be satisfied for the entire range of $F(S)$ to be used.

$$f_4/f_1 = \frac{w L^2 (1 + S)}{t m^2 [1 + S (L w / m t)^2]} \neq 1 \quad (6)$$

$$f_4/f_2 = \frac{w (1 + S)}{t [1 + S (w / t)^2]} \neq 1 \quad (7)$$

Additional inequalities arise for the overtones. For instance, the numerator of Inequality 6 should be multiplied by 0.363 to avoid $f_4 =$ the next overtone of f_1 .

Additional inequalities arise also because it is undesirable to have $2 f_4 = f_1$, etc. The complete set of inequalities can be found elsewhere.²

The inequalities must be evaluated for the range of $F(S, \text{force sensitivity})$ to be used. This has been done for the design of the device of Fig. 2 where $S_{\text{max}} = 0.1$ and $S_{\text{min}} = -0.1$ (in other words, $\pm 10\%$ frequency shift) were chosen for 3.4×10^6 dyne/cm² longitudinal stress in each tine in both tension and compression. Figure 4 shows the regions of L/m and t/w allowed when the

complete set of inequalities is used. Note that the device of Fig. 2 with L/m of 1.18 and t/w of 0.8 is in the middle of one of the largest allowed regions of Fig. 4.

UNEQUAL LOADING

Another problem arose in using the devices for force measurements. Each tine acts like its own force sensor, so if the force is not distributed equally between the two tines, the individual tine frequencies diverge and the mode coupling that makes Mode 4 possible ceases. This is a severe effect when one designs the devices for large S . Figure 5 shows how one might intentionally load the tines unequally. This configuration was used to evaluate the importance of the effect. Figure 6 shows what happens when there is a 1.8 or 3.6 % difference in the loading on the two tines. Note in Fig. 6 that there is a degeneracy that occurs when the two tines are no longer mode-locked together. Also note that the resistance at resonance degrades at forces below those causing the degeneracy. Clearly, if this circumstance were allowed to occur in practice, the device would have linearity problems at lower forces and stability problems at larger forces.

Figure 5 shows a practical configuration where the forces on the tines are equal enough to maintain mode-locking for the tines so that Mode 4 can be used to measure force. This configuration has proven useful if the flexure is integrally machined with the lever arm to form an integral suspension system.

TEMPERATURE BEHAVIOR

The frequency of Mode 4 of the double-ended tuning fork behaves with temperature quite similarly to the single-ended tuning fork used in the watch industry if there is no force on the device. Thus, the usual crystallographic orientations can be used to obtain a parabolic f vs. T behavior with turnovers near room temperature. However, as one can imagine from Fig. 5, if the metal used for the lever arm and flexure has a different thermal expansion than quartz, there is a temperature-dependent force acting on the device. In the case of Fig. 5, the force depends upon the torsional spring constant of the flexure, the difference in thermal expansion, the length of the free quartz, and the lever arm length. In addition, any mechanical structure attached to the input side of Fig. 5 must be considered. Most structures are too complicated to represent here by one equation. Some useful relations can be found elsewhere.⁴ A useful guideline is that when a 1 ppm/C change in frequency due to a temperature-dependent force from the structure is superimposed upon the inherent parabolic behavior, the turnaround temperature shifts by 12 °C. The shift can be up or down depending upon the net effect of the entire structure used to apply the force.

APPLICATION

The device of Fig. 2 has been used for measuring force in a space program application. The reader is referred to reference 5 for details. The application involves measuring the force applied to a magnetic plate (rotor) by an electromagnet. The force is used to point an optical experiment. The force measurement from the double-ended tuning fork is used in a feedback loop to correct for hysteresis and thermal effects in the magnet permeability. The maximum force to be measured is 40 N. The device of Fig. 2 can handle 3 N with a 3X safety factor. A lever system was designed to reduce the force from the rotor to a size compatible with the double-ended tuning fork. Figure 7 shows the lever system. It is a compound lever system isolated from the environment by a pair of opposing bellows. The crystal can therefore be in a vacuum while the force is applied with a pushrod through one of the bellows, as seen in Fig. 7. A simple CMOS oscillator circuit drives the crystal in Mode 4. The output of the oscillator circuit provides a measure of the frequency of Mode 4 of the device. Force measurement is accomplished by measuring the period of the oscillator output and using a polynomial expansion of the transcendental equation relating force and frequency. The following relation is adequate for most applications:

$$F = A (1 - f/f_0) - B (1 - f/f_0)^2 \quad (8)$$

where A and B are constants and f_0 is the frequency measured with no force applied.

The performance of the double-ended tuning fork is impressive. Figure 8 shows the resolution obtained with the lever system of Fig. 7. Resolution is defined here as the standard deviation of 10 consecutive period measurements using a 1 second gate. Note that the resolution is better than 1/400,000 of the full scale force of 40 N. This type of resolution is typically obtained if the mechanical structure is properly designed and the force is applied equally enough to the tines to maintain a stable crystal-controlled oscillation. Any one familiar with force measurement technology should appreciate that the results in Fig. 8 are at least an order of magnitude better than most other force measurement techniques.

The linearity of the results is shown in Fig. 9 for bi-directional force measurements. It is shown as the residual vs. actual force applied after Eq. 8 is used to reduce frequency shift data to force. The small residual seen in Fig. 9 is primarily due to the difficulty in applying the forces perfectly opposite one another when going from tension to compression in the crystal. Once again, the results are impressive when compared to other force measurement techniques.

CONCLUSIONS

A number of technical considerations have been presented for the design and application of a new force transducer called the double-ended tuning fork. A number of stumbling blocks in the use of the device have been pointed out and ways around them have been presented. The device proves to be a low cost, high performance force sensor. It has been successfully used for a critical force measurement in a pointing system application for space experiments. The results prove the viability of the double-ended tuning fork as a practical device if proper device design and careful application of engineering principles are used.

ACKNOWLEDGEMENTS

The authors express their appreciation for the contributions of T. Montague, G. Banik, and L. Clayton.

REFERENCES

1. E. P. EerNisse, "Miniature Quartz Resonator Force Transducer," U. S. Patent 4,215,570.
2. E. P. EerNisse, "Resonator Force Transducer," U. S. Patent 4,372,173
3. J. M. Paros, "Mounting System For Applying Forces to Load-Sensitive Resonators," U. S. Patent 4,384,495.
4. J. M. Paros, "Isolating And Temperature Compensating System For Resonators," U. S. Patent Pending (Claims Allowed, 1983).
5. E. P. EerNisse and G. J. Banik, "Vibrating Quartz Force Sensor For Magnetic Suspension Systems," 29th International Instrumentation Symposium, May 1-6, 1983, Albuquerque, NM.
6. S. S. Chuang, "Force Sensor Using Double-Ended Tuning Fork Quartz Crystals," 37th Annual Frequency Control Symposium, Philadelphia, June 1-3, 1983.
7. S. Timoshenko and D. H. Young, Vibration Problems In Engineering, D. Van Nostrand (3rd Edition, 1955), pp. 374-375.
8. J. M. Paros, "Precision Digital Pressure Transducer," ISA Transactions, Vol. 12, No. 2, 1973, pp. 173-179.
9. N. Serra, "Technical Report on the Quartz Resonator Digital Accelerometer, AGARD Conference Proceedings #43, Inertial Navigation Systems And Components, Oxford, England, 1967.

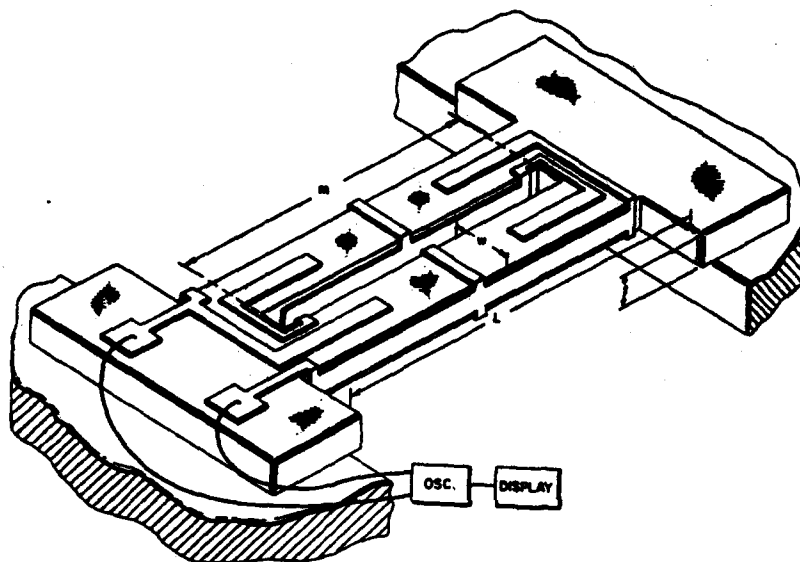


Figure 1. Double-Ended Tuning Fork configuration.

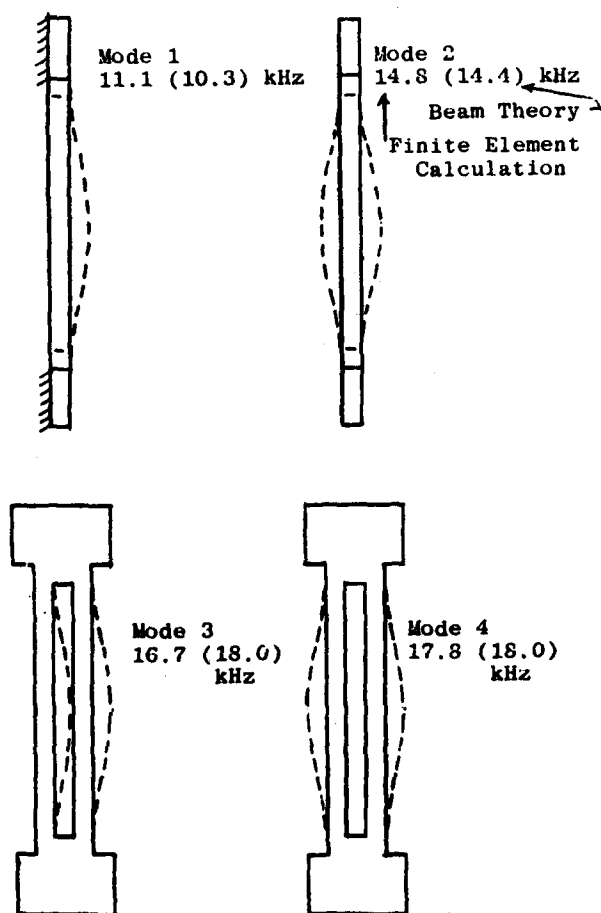


Figure 2. First four resonant modes of the Double-Ended Tuning Fork. $L=1.0$ cm, $m=0.85$ cm, $t=1.8 \times 10^{-2}$ cm, $w=2.3 \times 10^{-2}$ cm.

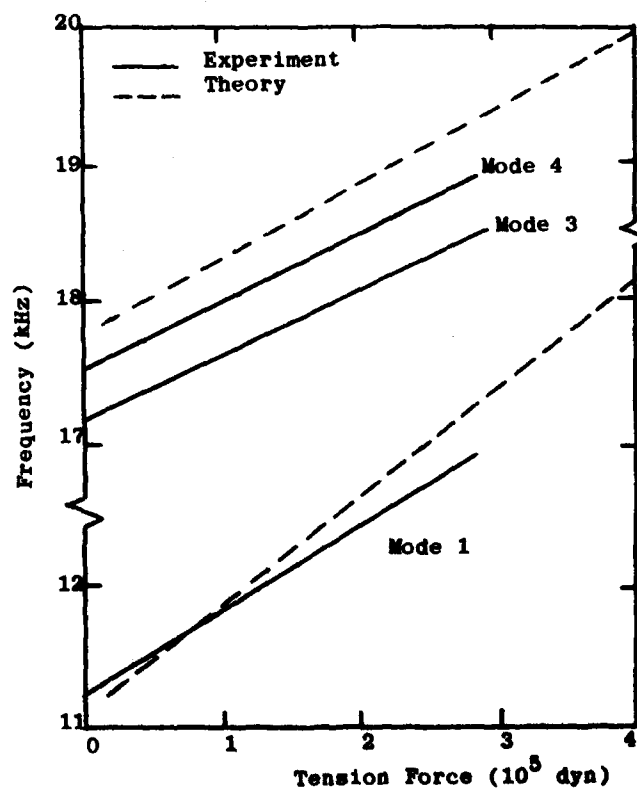


Figure 3. Frequency of Modes 1, 3, and 4 vs. force. Experiment and theory.

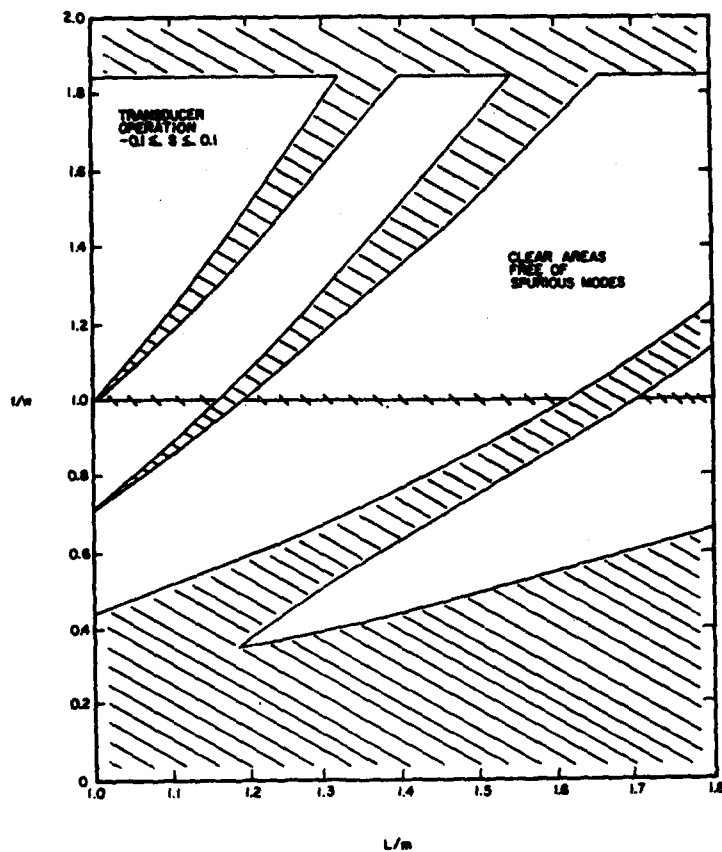


Figure 4. Plot of t/w vs. L/m showing regions where spurious modes exist for a specific choice of S_{\max} and S_{\min} .

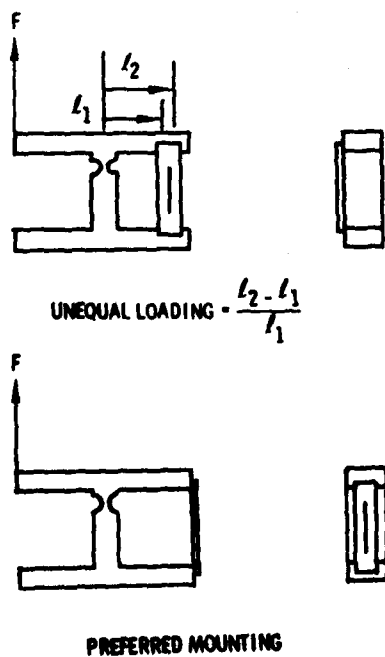


Figure 5. Method used for unequal loading of tines and preferred mounting where tines are equally loaded.

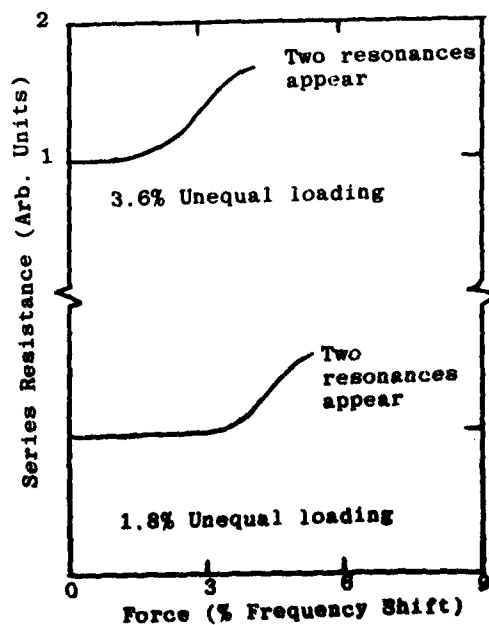


Figure 6. Effect of unequal loading of the tines.

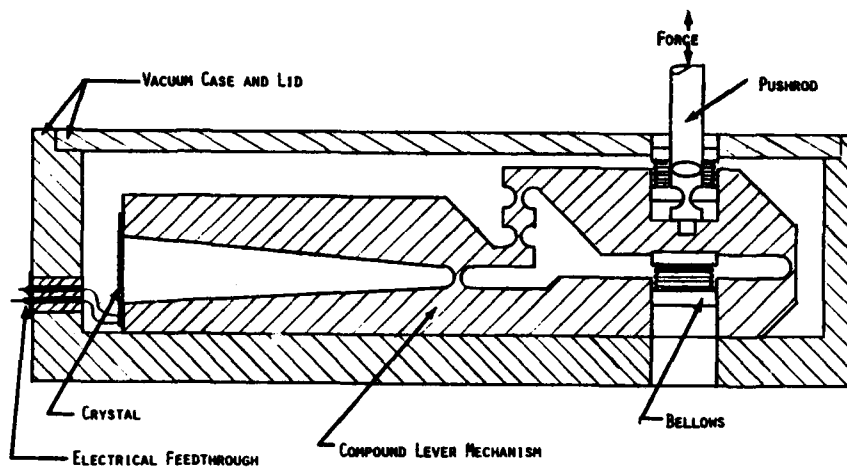


Figure 7. Compound Lever System for applying force to the Double-Ended Tuning Fork.

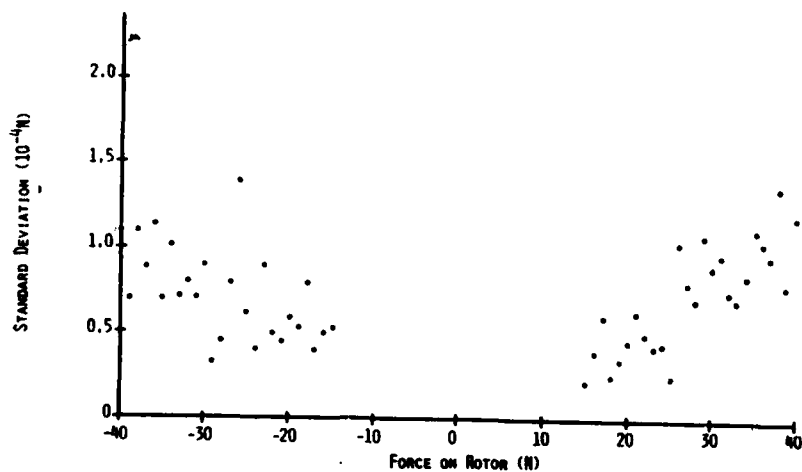


Figure 8. Measured Force Resolution of Crystal Assembly.

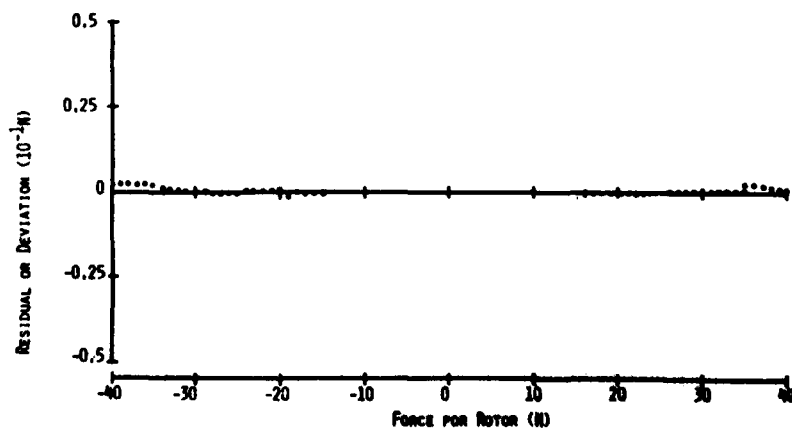


Figure 9. Measured Deviation from linear after Data Reduction with Equation 8.

AD P 002485

CHEMICALLY POLISHED HIGH FREQUENCY RESONATORS

by William P. Hanson
Piezo Crystal Company
Certeale, PA 17013

Introduction

The upper frequency limit of quartz resonators is determined by the physical properties of quartz and restrictions imposed by processing techniques. More often than not, the processing limitations are reached long before the physical limitations. Mechanical polishing has been used successfully for many years. The frequency limits determined by the mechanical polishing process have restricted the use of quartz resonators to frequencies less than 250 MHz. Chemical polishing is an etching process which removes quartz from a blank in a well defined manner. The upper frequency limit of chemically polished resonators is much higher than mechanically polished resonators. Chemically polished 'SC' resonators have surface topographies similar to highly polished resonators manufactured by conventional means. Currently, chemically polished 'SC' resonators can operate up to the seventh overtone, higher overtones may be achievable when a better quality quartz is used as a source material. Frequencies in the GHz range are thus manufacturable using chemical polishing.

Etching has been used in the manufacturing of resonators for many years. Rough cut blanks are regularly etched to check for twinning, +X and -X surfaces can be differentiated by deep etching in Ammonium Bifluoride. Partially because of these tests, the etching process is associated with the degradation of quartz surfaces. Quartz is an anisotropic solid; the physical properties vary with direction. Consequently, the etching rate varies with direction. There are many different types of resonators, 'DT', 'BT', 'AT', 'GT', 'FS', 'IT', 'SC'; to name a few. Each type is cut from a different crystalline orientation. The surface of a mechanically polished 'AT' blank turns frosty when etched in Ammonium Bifluoride. The quality of the etched surface is affected by many parameters: the chemicals used to make the solution, the temperature of the solution, the impurities in the solution, and the cut. Solutions developed by Dr. John Vig at ERADCOM etch 'SC' resonators with minimal degradation of the surface (See Figure 1). These are the solutions used in the manufacture of high frequency 'SC' resonators. A mechanically polished 'SC' resonator etched in a 5:1 solution of Ammonium Fluoride and HF at 75 degrees C will still have a highly polished surface after deep etching.

Benefits of Chemical Polishing

The following list describes some of the benefits chemical polishing has over mechanically polished resonators.

This is only a partial list assembled from one year of experience manufacturing high frequency 'SC' resonators.

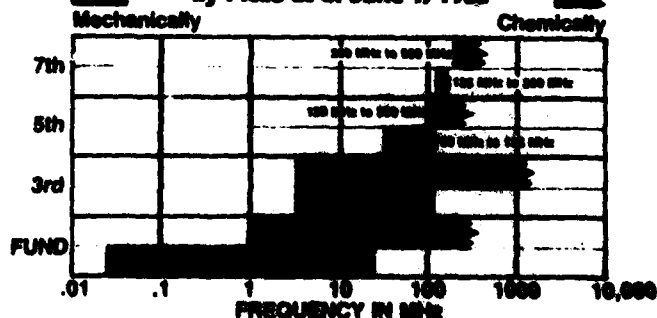
1. **HIGHER FREQUENCIES** can be manufactured using chemical polishing rather than mechanical polishing. Frequencies up to 400 MHz on the fundamental are being made. The response of a 371 MHz fundamental resonator is printed in Figure 7. The frequency limits of this new technology have not been reached.
2. Chemically polished resonators are **MECHANICALLY STRONGER** than mechanically polished resonators by several orders of magnitude (See Figure 1). Unplated resonators up to 400 MHz on the fundamental mode are handled with less than 8% breakage.
3. **COMPLEX GEOMETRIES** may be produced by selectively masking the quartz blank. Masses and rings are regularly etched out of the quartz. Quartz bridges and uncoupled multiple resonators have been made to check feasibility.
4. **HIGHER QUALITY** blanks are sorted out during the first etching. Blanks which have major lattice defects can be rejected. Blanks with etch pits, etch tunnels, and twinning are also rejected after the first etching.

Figure 1 is a graph of the frequencies manufactured by chemical and mechanical polishing. Mechanically polished fundamental resonators are available roughly from 3 KHz to 35 MHz. Conversely, chemically polished fundamental resonators are available from around 1 MHz to 400 MHz, extending the available frequency range for fundamentals ten fold. Third overtones mechanically polished

Figure 1

FREQUENCY SPECTRUM

for mechanically lapped and chemically
polished resonators manufactured
by Piezo as of June 1, 1983



ished resonators are available from 5 MHz to 135 MHz. Third overtone chemically polished resonators are available from 5 MHz to 1.1 GHz, an extension of eight fold. The extension of the frequency ranges has not been completed. Mechanically polished resonators have well defined manufacturing limitations. The manufacturing limitations of chemical polishing are just now being tested.

Chemical Solutions Used in the Manufacture of High Frequency Resonators

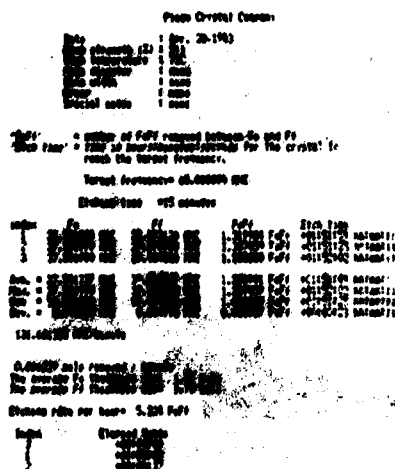
Many etching solutions have been tried for SC resonators. Dr. Vig has published several articles on solutions for chemical polishing (See Figure 2). The 5:1 solution is a mixture of 40% Ammonium Fluoride to 49% HF. Vig's 5:1 chemical solution polishes SC blanks quite well at 75 degrees C. Chemically polished SC resonators have been manufactured up to 400 MHz on the fundamental mode using the 5:1 solution. Vig's 10:1 solution chemically polishes slightly better than the 5:1 solution, but is much slower.

Etching Calculations

'F-squared', a frequency calculation, is a convenient method of calculating the etching rate since it is directly proportional to the thickness removed in the process and independent of the thickness (frequency). The etching rate is determined by etching several crystals for a pre-determined amount of time and using the change in blank frequency to calculate the etching rate. Figure 2 shows the printout from a computer program used to calculate the etching rate from a group of crystals. The calculated etching rate is an average etching rate for the group.

Figure 2

Computer Printout Determining Etching Rate



Some other interesting information is printed with the etching rate on the computer printout. The grouping information at the bottom of the figure is used to tighten the frequency spread of the blanks. The time each blank is to be placed in the etching solution is ordered from longest to shortest. Utilizing a digital watch to keep track of elapsed time and by placing the blanks in the solution at the elapsed time calculated, the group of crystals will theoretically reach the target frequency at the same time. Using this technique blanks can be grouped in frequency to within .01 "F-squared" easily; tighter grouping is possible with practice.

F-squared calculations in the program are based on the following three equations, where N is F-squared.

$$N = \frac{1000 \times (F_i - F_f)}{F_i \times f_i}$$

$$F_i = \frac{1000 \times F_f}{1000 - N \times F_i}$$

$$F_f = \frac{1000 \times F_i}{N \times F_i + 1000}$$

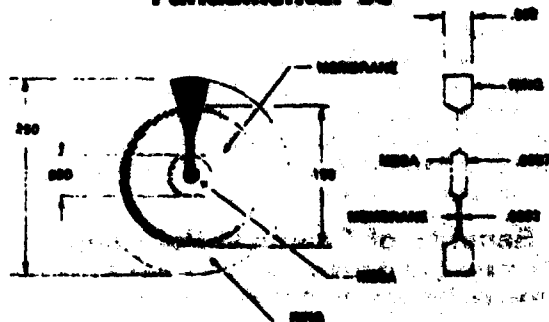
F_i: Initial Frequency in MHz
F_f: Final Frequency in MHz
N: "F-squared" in MHz

Resonator Design

Figure 3 is a drawing of a 100 MHz fundamental 'SC' resonator design. Several differences from conventional resonators are noticed immediately. A supporting ring of quartz is at the edge of the resonator. The ring thickness is determined by the thickness of the original mechanically lapped resonator. The central mass determines the frequency of the resonator, and the membrane serves as an electrical link to the ring. The supporting ring must be thick enough to use as a mounting structure and thin enough so that the etching process is optimized. If a large amount of

Figure 3

Chemically Polished 100MHz Fundamental 'SC'



quartz is etched away, the possibility of etch defects degrading the resonator is significantly increased. If a blank is too thin it will bend under the slightest stress. Mounting then becomes an impossible task.

Figures 4 through 10 are collections of data on the high frequency resonators manufactured by chemical polishing. The designs used are similar to the design for the 100 MHz fundamental 'SC' resonator described above. The data chosen is typical data, with the exception of the 371 MHz fundamental 'SC'. Taking into consideration that few have been manufactured, the typical values are not known. Some of the parameters of the higher frequency resonators are unknown and are marked UNKNOWN. This is due to the limits of available measuring equipment.

Difficulties with Chemical Polishing

The following list of difficulties with chemical polishing has been accumulated from a year of experiments. Some of the difficulties in chemical polishing are under control, but need constant attention.

1. Etched surfaces do not always keep their original SURFACE FINISH. 'AT' and 'BT' cut surface finishes degrade readily. 'SC' cut resonators keep their original surface finish with minimal degrading during etching. 'SC' resonators etched 30 F-squared still have a polished surface. In contrast, 'AT' resonators etched only 3 F-squared no longer have a polished surface; rather, their surface is more like a 1 micron finish.
2. Etching to a SPECIFIC FREQUENCY is a tedious process. The etching rate must be measured precisely. Each blank's frequency must be measured and the temperature of the etching solution must be maintained. The more a group of blanks is etched the greater the frequency spread becomes. If a tight grouping is necessary, then the crystals must be etched by different amounts.
3. Minimizing the concentration of ETCH DEFECTS is a prime concern. Only natural and swept premium Q cultured quartz are satisfactory for deep etching. High concentrations of etch tunnels and etch pits exclude unswept cultured quartz from being used for deep etching. Unswept natural quartz is minimally acceptable. Nearly 50% of the natural quartz is rejected for gross lattice defects, etch pits, or etch tunnels.
4. SURFACE IRREGULARITIES on the blank are propagated throughout the etching process. These are not etch defects, but rather scratches, chips, or nonparallelism. Nearly perfect blanks have to be used as a source material.
5. SAFETY is a prime concern when working with HF. Hydrofluoric acid has a stable vapor pressure and is considered a liquified gas under pressure. Protective clothing, fume hoods, and face masks are mandatory.

Figure 4

72.765 MHZ FUNDAMENTAL 'SC'

MAIN: 72.765 MHZ FUND SN 280



Sample #	: 280	Q (Roc)	: 58547
Series Resonance	: 72.765636 MHZ	Ratio (C0/C1)	: 948
Mode	: Fundamental	Plating	: A1, 2
Roc (Ohms)	: 44.9	Z-angle	: 34 degrees 10 min
C0 (pf)	: .93	Spur Free	: 1.23%
C1 (pf)	: .00004	Highest Spur	: 1.23%, -26db
L (mH)	: 4.884		

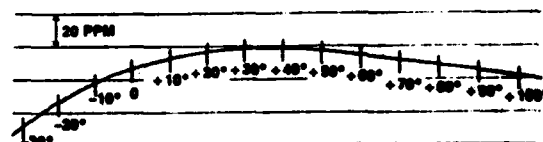


Figure 5

107.390 MHZ FUNDAMENTAL 'SC'

MAIN: 107.390 MHZ FUND SN 217



Sample #	: 217	Q (Roc)	: 32880
Series Resonance	: 107.390400 MHZ	Ratio (C0/C1)	: 1084
Mode	: Fundamental	Plating	: A1
Roc (Ohms)	: .98	Z-angle	: 34 degrees 12 min
C0 (pf)	: 1.00	Spur Free	: .17%
C1 (pf)	: .00117	Highest Spur	: .58%, -17db
L (mH)	: 3.521		



Figure 6

321.744 MHz 3RD OVERTONE 'SC'

MAIN: 321.744 MHz 3RD SC SN 217

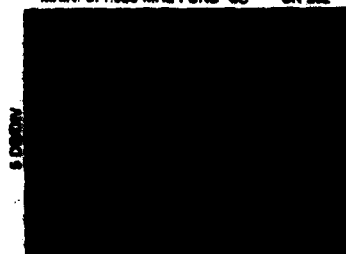


Sample #	: 217
Series Resonance	: 321.744 MHz
Mode	: 3 overtones
Res (Ohms)	: approx. 200
C0 (pf)	: 1.08
C1 (pf)	: unknown
L (mH)	: unknown
Q (3db)	: 11000
Ratio (C0/C1)	: unknown
Plating	: A1
Z-angle	: 34 degrees 12 minutes
Spur Free	: .03%
Highest Spur	: .34% out, -4.3db

Figure 7

371.588 MHz FUNDAMENTAL 'SC'

MAIN: 371.588 MHz FUND 'SC' SN 202

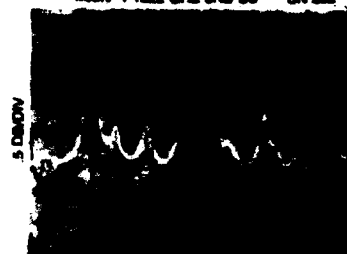


Sample #	: 202
Series Resonance	: 371.588 MHz
Mode	: Fundamental
Res (Ohms)	: approx. 150
C0 (pf)	: 1.20
C1 (pf)	: unknown
L (mH)	: unknown
Q (3db)	: 15000
Ratio (C0/C1)	: unknown
Plating	: A1
Z-angle	: 34 degrees 12 minutes
Spur Free	: .02%
Highest Spur	: .05%, -8db

Figure 8

1.1222 GHz 3RD OVERTONE 'SC'

MAIN: 1.1222 GHz 3RD SC SN 202

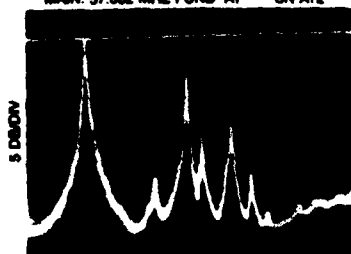


Sample #	: 202
Series Resonance	: 1.1222 GHz
Mode	: 3 overtones
Res (Ohms)	: unknown
C0 (pf)	: 1.20
C1 (pf)	: unknown
L (mH)	: unknown
Q (3db)	: 5000
Ratio (C0/C1)	: unknown
Plating	: A1
Z-angle	: 34 degrees 12 minutes
Spur Free	: .02%
Highest Spur	: .07%, -.7db

Figure 9

57.002 MHz FUNDAMENTAL 'AT'

MAIN: 57.002 MHz FUND 'AT' SN AT2

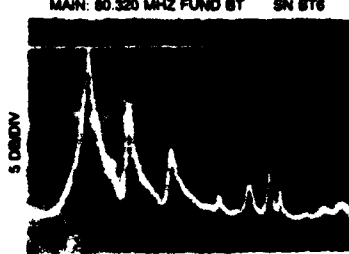


Sample #	: AT2
Series Resonance	: 57.002 MHz
Mode	: Fundamental
Res (Ohms)	: 360
C0 (pf)	: 1.18
C1 (pf)	: .00127
L (mH)	: 6.141
Q (3db)	: 6100
Ratio (C0/C1)	: 929
Plating	: Gold
Spur Free	: .74%
Highest Spur	: 1% out, -6.9db

Figure 10

80.320 MHz FUNDAMENTAL 'BT'

MAIN: 80.320 MHz FUND BT SN BT6



Sample #	: BT6
Series Resonance	: 80.320 MHz
Mode	: Fundamental
Res (Ohms)	: unknown
C0 (pf)	: .95
C1 (pf)	: unknown
L (mH)	: unknown
Q (3db)	: 3400
Ratio (C0/C1)	: unknown
Plating	: Gold
Spur Free	: .11%
Highest Spur	: .30%, -9.5db

Conclusions

1. Chemically polished blanks up to 400 MHz on the fundamental are manufacturable.
2. Etching to frequency is a controllable process.
3. Chemically polished resonators operate up to the 7th overtone.
4. Frequencies up to 1.2 GHz are manufacturable.

I wish to thank Wallace Samuelson for his generous guidance. His expertise in resonator technology, and his considerable patience has been a strong motivating force to me personally and professionally.

Bibliography

Vig, John R., 4th Quartz Crystal Conference Proceedings, pp47-84

Vig, John R., Proceedings of the 31st Annual Symposium on Frequency Control 1977, pp131-143

AD P 002486

FURTHER STUDIES ON THE ACCELERATION SENSITIVITY OF QUARTZ RESONATORS

Raymond L. Filler, John A. Kosinski and John R. Vig

US Army Electronics Technology & Devices Laboratory (ERADCON)
Fort Monmouth, New Jersey 07703

Abstract

Further studies on the acceleration sensitivity of quartz resonators have been performed. These studies have included conventional AT, strip AT, GT, IT, AK, and SC-cut resonators. No significant differences have been found in the best achievable acceleration sensitivities of the various cuts.

The dependence of the acceleration sensitivity of 4-point mounted SC-cut resonators on design parameters and operating conditions has been investigated. These parameters include blank geometry, angle of cut, mounting orientation, temperature, drive current, and overtone. No significant systematic variation of acceleration sensitivity has been found as a function of any of these parameters except overtone for the same resonator.

The advantages of 4-point mounting over 3-point mounting of SC-cut resonators have been demonstrated. The advantages have been shown to be 1) a lower total acceleration sensitivity for 14mm diameter blanks, and 2) more consistent manufacturing yields of low acceleration sensitivity resonators.

Key Words:

Quartz, quartz crystal, quartz resonator, quartz oscillator, SC-cut, AT-cut, AK-cut, IT-cut, GT-cut, acceleration, acceleration sensitivity, vibration, frequency control.

Introduction

The high degree of spectral purity required by some modern communication, navigation, and radar systems and the immunity to acceleration induced frequency offsets required by some precision timing applications require the development of resonator designs with lower acceleration sensitivities than are presently available. As part of a development program on high precision ceramic flatpack enclosed quartz crystal resonators, both SC-cut and AT-cut, the effects of blank geometry on various resonator performance parameters have been investigated. These effects, particularly those influencing AT-cut acceleration sensitivity, have been reported previously. In this paper, these investigations are expanded to include the effect of other design parameters on the acceleration sensitivity of SC-cuts. In addition, several other cuts of quartz and enclosure types have been studied.

Resonator Configurations

The majority of experiments were performed on SC-cut resonators enclosed in ceramic flatpacks. The details of the fabrication of such resonators have been reported previously²⁻⁹. The resonator blanks have a 14mm diameter and were bonded to four 1.5mm wide molybdenum ribbon "L" clips with silver filled polyimide adhesive. The clips were spaced 90° apart on the perimeter of the blank, two on the xx' axis and two on the zz' axis.

Other configurations tested included 3-point mounted AK-cut and SC-cut resonators enclosed in "C" holders, 4-point mounted IT-cut resonators in ceramic flatpacks, and miniaturized GT-cut resonators.

Acceleration Sensitivity Measurement System

Two methods of testing the acceleration sensitivity were employed. These were sinusoidal vibration on a shake table and the "2-6 tipover" test. The results of these two tests on any given resonator typically correlate to better than 1x10⁻¹⁰/g.

For the sinusoidal vibration tests, the resonators were connected to appropriate oscillator circuits and the acceleration sensitivity was determined by examining the response of the oscillators to the vibration. This technique minimized the temperature stability required and allowed a search for mechanical resonances in the resonator mounting structure. The oscillators were potted in beeswax to minimize the relative motions of oscillator components. The relative powers of the vibration induced frequency modulation sidebands were measured. The relative magnitude of the power in the first sideband to the power in the carrier is denoted by γ . γ is related to the acceleration sensitivity, $\frac{1}{f_0}$, of the resonator by

$$\gamma = \frac{2f_v}{af_0} \cdot 10^{\gamma/20} \quad (1)$$

where:

- f_v = vibration frequency in Hz
- f_0 = resonator operating frequency in Hz
- a = acceleration level in g's
- γ = sideband/carrier ratio in decibels

The value of γ was determined for several vibration frequencies in order to be certain that no mechanical resonances were present in the oscillator or resonator.

For the "2-G tipover" test, the resonators were set to turnover in appropriate precision ovenized oscillators. The acceleration sensitivities were determined by measuring the fractional frequency deviations corresponding to known changes in orientation with respect to the earth's gravitational field. The resonators were first rotated about an axis normal to the plane of the crystal plate to determine the radial component of the acceleration sensitivity. The resonators were then rotated 180° about the zz' axis to determine the normal component of acceleration sensitivity. As a self check, the response of the oscillator to a complete rotation about zz' was then predicted from the measured components, and verified by performing the rotation.

Results

For 4-point mounted SC-cut crystals in ceramic flatpack enclosures, the influence of several design parameters and operating conditions on acceleration sensitivity has been determined. The resonator designs tested were 5 MHz fundamental mode, 5 MHz 3rd overtone, and 10 MHz 3rd overtone SC-cuts. The design parameters examined were blank contour, blank thickness, angle of cut, and mounting orientation. The operating conditions examined were temperature, drive level, and overtone. The direction of maximum sensitivity to in-plane acceleration has been determined for units tested by the "2-G tipover" method.

Acceleration Sensitivity vs. Contour

It has previously been reported that the acceleration sensitivity of AT-cut quartz resonators depends strongly on blank contour¹. Similar measurements have now been made on SC-cut resonators. The results are shown in figures 1-3. Over the range of contours from 0.5 diopter to 5.0 diopter, no systematic variation of acceleration sensitivity has been observed. Although not noted in the figures, the data includes some blanks which have the positive-on-compression side contoured, some blanks which have the negative-on-compression side contoured, and some blanks which are biconvex. These different geometries do not appear to affect the acceleration sensitivity.

Acceleration Sensitivity vs. Thickness

Figure 4 presents the acceleration sensitivities of the SC-cut resonators tested as a function of blank thickness. There appears to be a slight trend toward lower achievable acceleration sensitivities by using thinner blanks. Recently however, tests have been made on 63.8 MHz and 100 MHz resonators of much smaller blank diameter. No improvements in the acceleration sensitivity were observed with these units. It is therefore possible that this potential influence on the acceleration sensitivity may be related to the diameter-to-thickness ratio.

Acceleration Sensitivity vs. Deviation From Reference Angle

SC-cut resonators for TCRO applications require the use of larger values of θ than oven controlled oscillator applications². Figures 5-7 show the acceleration sensitivity as a function of θ . Over the range of up to 60 minutes from the reference angle, the acceleration sensitivity of

these resonators does not appear to be affected by the value of θ .

Acceleration Sensitivity vs. Mounting Orientation

It has been reported that the magnitude of the acceleration sensitivity of 4-point mounted SC-cut crystals should be influenced by the location of the mounting points¹¹. Resonators have been fabricated with the mounting orientation, ψ , varying from 0° to 45°. No large variation in the best achievable acceleration sensitivity is observed with mounting orientation, as seen in figures 8-10. It appears that an unknown factor may be masking the predicted sensitivity to mounting orientation.

Acceleration Sensitivity vs. Overtone

The acceleration sensitivities of several resonators have been measured using overtones of the same resonator. The results are presented in figure 11. It is seen that for 9 of 10 resonators tested, the acceleration sensitivity was improved by using the higher overtones of the same crystal.

Acceleration Sensitivity vs. Drive Level

The testing of various overtones of the same crystal required operation of the crystals in oscillators of different frequencies. To eliminate the possibility that the observed changes with overtone were due to changes in the drive-level from oscillator to oscillator, a "2-G tipover" test was performed repeatedly using an ovenized oscillator with externally adjustable crystal drive current capability. For a 0.4 mA to 2.0 mA range of crystal drive current (which exceeds the range of oscillator drive current normally used), no variation in the acceleration sensitivity was observed.

Acceleration Sensitivity vs. Temperature

The oscillators used in vibration testing were not ovenized. As the vibration apparatus heated up during use, the temperatures of the test oscillators were measured as varying between 20° C and 60° C. No variation in acceleration sensitivity was observed as the temperature changed. Several of these crystals were later ovenized at temperatures between 75° C and 120° C for "2-G tipover" tests. As noted before, the acceleration sensitivities measured with the "2-G" tests matched those measured by sinusoidal vibration. The acceleration sensitivity therefore does not appear to be affected by operating over the range of 20° C to 120° C.

Direction Of Maximum Sensitivity To In-Plane Acceleration

The results of the "2-G" tests have been used to calculate the directions of the acceleration sensitivity vectors. Figure 12 shows the directions and relative magnitudes of the radial components of the vectors relative to the resonator flatpack structure. The magnitudes of the radial components shown vary from about 1×10^{-10} to about 11×10^{-10} /g. The zz' axis lies along the electrode tabs.

It is seen that the radial component of the acceleration sensitivity vector occurs only in certain sections of the blank. Due to the geometry

of the mounting structure and the ambiguities of mounting the blanks, it is possible to rotate the blank 180° about the zz' axis and xx' axis and end up with all the vector components lying in one quadrant and grouped close to the zz' axis. If we perform these rotations, we obtain the distribution shown in figure 13. This figure is an histogram of the number distribution of vectors lying off the zz' axis by each given angle. It is seen that the radial component of the acceleration sensitivity exhibits good repeatability and typically occurs 20° to 30° off the zz' axis for SC-cut resonators.

Four Point Mounting Structure

The effects of using a four-point mounting structure rather than a three-point structure have been investigated, and are shown in figure 14. The first two columns represent 14mm diameter blanks, half of which were mounted in 3-point mount "C"-type holders, and the other half in 4-point mount ceramic flatpacks. The set of blanks mounted at 3-points had acceleration sensitivities between $2.2 \times 10^{-9}/g$ and $3.4 \times 10^{-9}/g$, while the 4-point mounted resonators were between $0.4 \times 10^{-9}/g$ and $1.4 \times 10^{-9}/g$. The component of acceleration sensitivity in the plane of the blank was similar for both the 3-point and 4-point mounted resonators. The variation of the acceleration sensitivity occurred in the component normal to the blank.

It was necessary to make certain that the difference in acceleration sensitivity between these groups was due to the mounting structure and not the differing electrical characteristics of the metal "C"-type holder and ceramic flatpack enclosure. Therefore, a 3-point mounted resonator was tested with 1) the enclosure grounded, 2) the enclosure floating, 3) a load capacitor, and 4) no load capacitor. As shown in figure 15, no significant difference was observed in the acceleration sensitivity of the unit as it was tested under these conditions.

The last three columns of figure 14 all represent 10 MHz resonators. The column of 14mm blanks represents a group of ceramic flatpack enclosed crystals, while the two columns at 9.5mm diameter represent 3-point and 4-point mounted crystals in TO-8 type holders. Here it is seen that by using a smaller diameter blank the range of acceleration sensitivities achieved by using 3-point mounts was similar to that achieved by using 4-point mounts. However, it can be seen that the yield of low acceleration sensitivity resonators is both greater and more consistent for the 4-point mounted crystals.

Acceleration Sensitivity vs. Angles Of Cut

The variation of acceleration sensitivities among several angles of cut has been examined. Figure 16 is a summary of the data. The columns for AK and IT-cuts represent the testing of only a few developmental resonators, and the column for GT-cut resonators represents the testing of a group of 10 miniature GT-cut resonators. Over a wide range of angles of cut (including different modes of vibration) no significant variation in acceleration sensitivity can be observed. Nearly all of the resonators fall within a factor of four of $1 \times 10^{-9}/g$.

Conclusion

No single parameter has yet been identified which substantially lowers the acceleration sensitivity of SC-cut resonators.

Acknowledgements

The authors wish to thank R. Brandmayr, W. Washington, and R. Williams for their assistance in fabricating the resonators used, and D. Bowman and P. Thompson for fabricating the many test oscillators. We also wish to thank the General Electric Neutron Devices Department, Colorado Crystal Corporation, Frequency Electronics, Inc., Piezo Crystal Company, Motorola Inc., SEIKO Corporation, and the Air Force Rome Air Development Center for supplying resonators used in several of our experiments.

References

1. R.L. Filler, J.A. Kosinski and J.R. Vig, "The Effect of Blank Geometry on the Acceleration Sensitivity of AT & SC-Cut Quartz Resonators," Proc. 36th ASFC*, pp 215-218 (1982)
2. R.L. Filler and J.R. Vig, "The Acceleration and Warm-up Characteristics of Four-Point-Mounted SC and AT-Cut Resonators," Proc. 35th ASFC*, pp 110-116 (1981)
3. J.R. Vig and E. Hafner, "Packaging Precision Quartz Crystal Resonators," Technical Report ECOM-4143, US Army Electronics Command, Fort Monmouth, NJ, (July 1973), AD 763215
4. P.D. Wilcox, G.S. Snow, E. Hafner and J.R. Vig, "A New Ceramic Flatpack for Quartz Resonators," Proc. 29th ASFC*, pp 202-219 (1975), AD A017466
5. R.D. Peters, "Ceramic Flatpack Enclosures for Precision Quartz Crystal Units," Proc. 30th ASFC*, pp 224-231 (1976), AD A046089
6. J.R. Vig, J.W. LeBus and R.L. Filler, "Chemically Polished Quartz," Proc. 31st ASFC*, pp 131-143 (1977) AD A088221
7. J.R. Vig, R.J. Brandmayr and R.L. Filler, "Etching Studies on Singly and Doubly Rotated Quartz Plates," Proc. 33rd ASFC*, pp 351-358 (1979)
8. J.R. Vig, "UV/Ozone Cleaning of Surfaces: A Review," In Surface Contamination: Genesis, Detection and Control, K.L. Mittal, ed., Vol 1, pp 235-254, Plenum Press, New York, 1979
9. R.L. Filler, J.M. Frank, R.D. Peters and J.R. Vig, "Polyimide Bonded Resonators," Proc. 32nd ASFC*, pp 290-298 (1978)
10. J.R. Vig, R.L. Filler and J.A. Kosinski, "SC-Cut Resonators for Temperature Compensated Oscillators," Proc. 36th ASFC*, pp 181-185 (1982)
11. P.C.Y. Lee and K.M. Wu "The Influence of Support Configuration on the Acceleration Sensitivity of Quartz Resonator Plates," Proc. 36th ASFC*, pp 29-34 (1977), AD A046221

*ASFC = Annual Symposium on Frequency Control

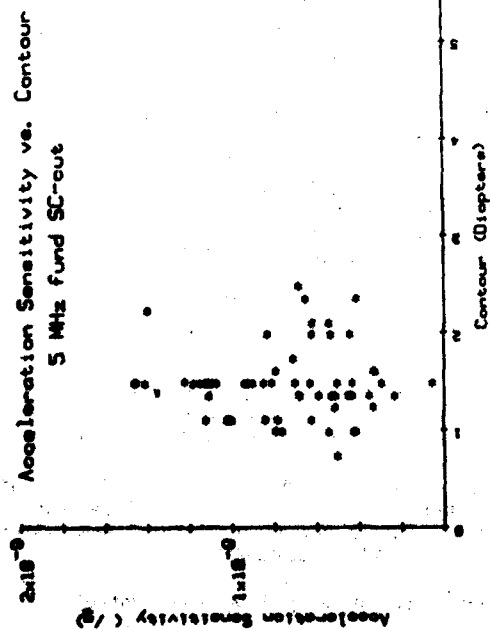


Figure 1 - Acceleration Sensitivity vs. Contour for 5 MHz Fundamental Mode 4-point Mounted SC-Cut Resonators.

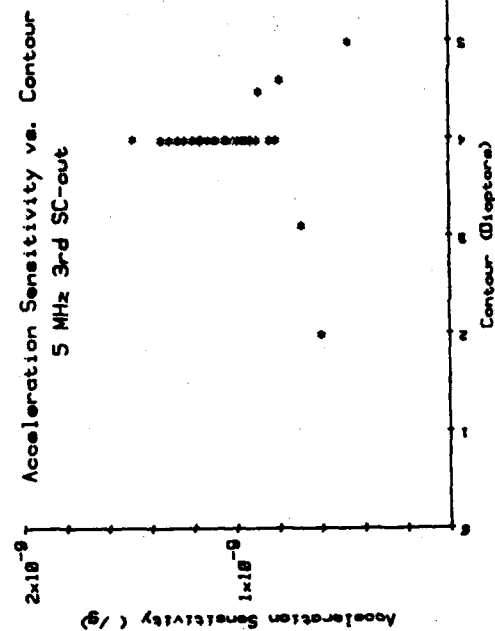


Figure 2 - Acceleration Sensitivity vs. Contour for 5 MHz 3rd Overtone 4-point Mounted SC-Cut Resonators.

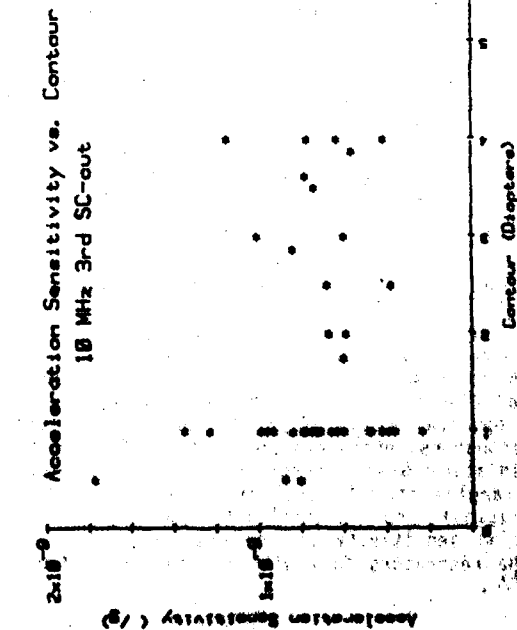


Figure 3 - Acceleration Sensitivity vs. Contour for 10 MHz 3rd Overtone 4-point Mounted SC-Cut Resonators.

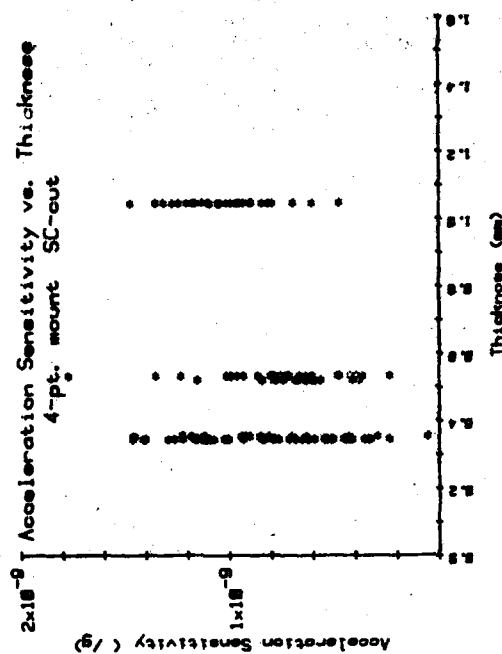


Figure 4 - Acceleration Sensitivity vs. Thickness for 4-point Mounted 14 mm Diameter SC-Cut Resonators.

SC-CUT RESONATORS
Accelerometer Sensitivity vs. Delta
5 MHz Fund SC-cut

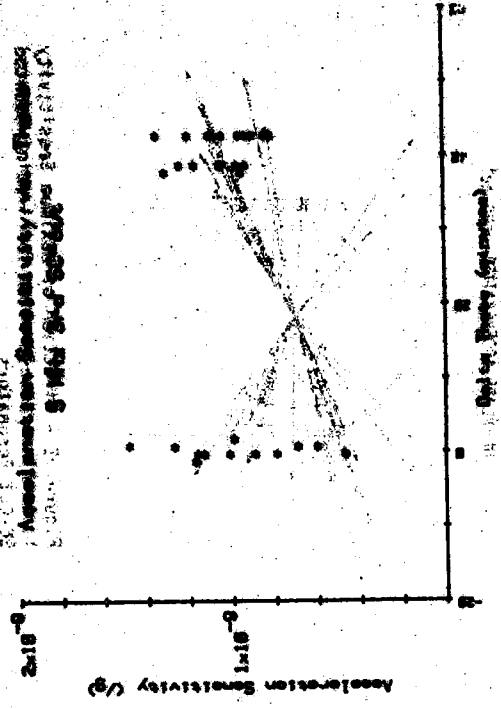


Figure 6 - Acceleration Sensitivity vs. δ for 5 MHz 3rd Overtone 4-point Mounted SC-Cut Resonators.

SC-CUT RESONATORS
Accelerometer Sensitivity vs. Delta
5 MHz Fund SC-cut

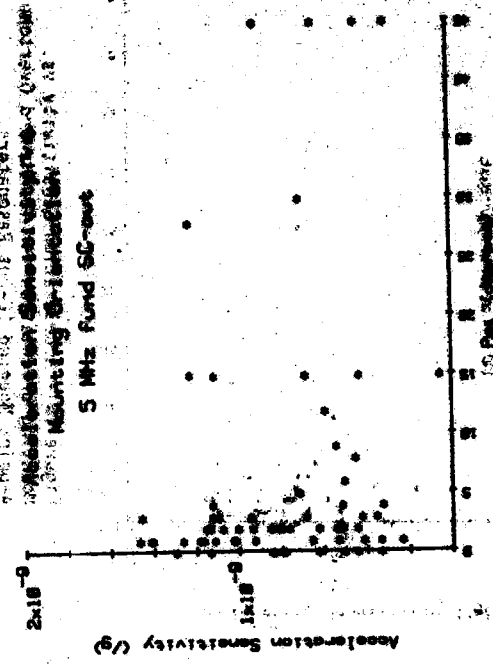


Figure 8 - Acceleration Sensitivity vs. δ for 5 MHz 3rd Overtone 4-point Mounted SC-Cut Resonators.

SC-CUT RESONATORS
Accelerometer Sensitivity vs. Delta
5 MHz Fund SC-cut



Figure 7 - Acceleration Sensitivity vs. δ for 5 MHz 3rd Overtone 4-point Mounted SC-Cut Resonators.

SC-CUT RESONATORS
Accelerometer Sensitivity vs. Delta
5 MHz Fund SC-cut

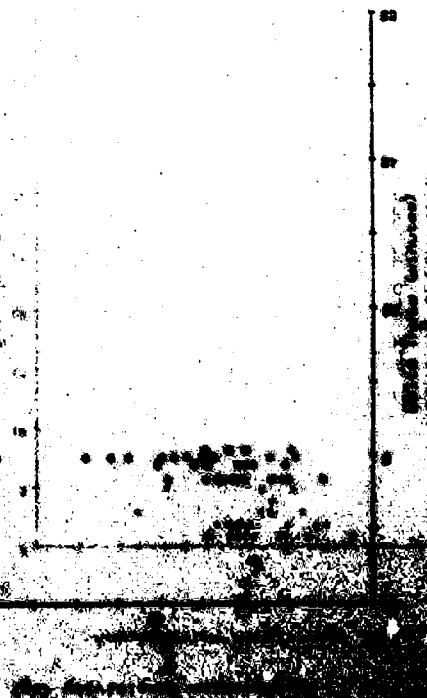


Figure 9 - Acceleration Sensitivity vs. δ for 5 MHz 3rd Overtone 4-point Mounted SC-Cut Resonators.

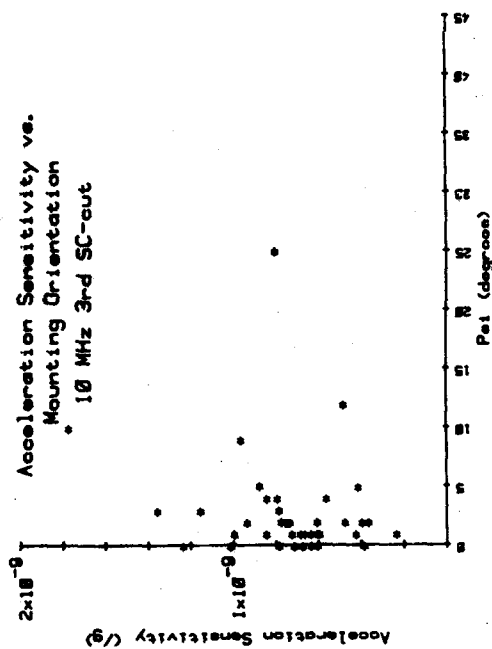


Figure 10 - Acceleration Sensitivity vs. Mounting Orientation for 10 MHz 3rd Overtone 4-point Mounted SC-Cut Resonators.

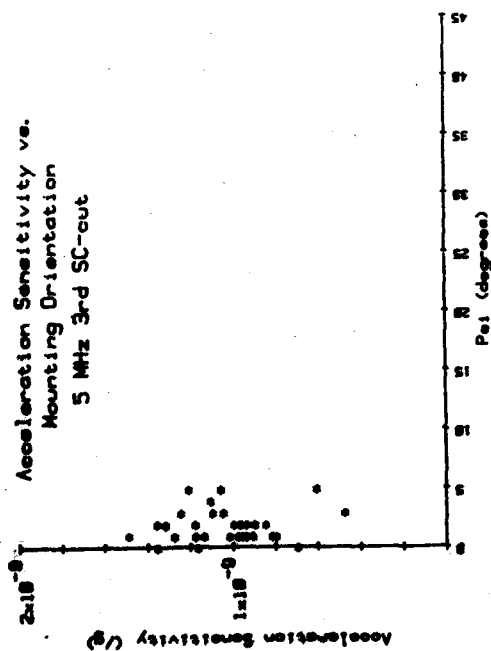


Figure 9 - Acceleration Sensitivity vs. Mounting Orientation for 5 MHz 3rd Overtone 4-point Mounted SC-Cut Resonators.

ACCELERATION SENSITIVITY^a vs. OVERTONE

TYPE	S/N	Fund	3rd	5th
5MHz Fund	CR-726	8.4	7.3	-
	CR-728	12.1	7.8	-
	CR-734	9.0	8.3	-
10MHz 3rd	109425	13.1	8.3	-
	109427	12.1	10.4	-
	109428	15.9	10.1	-
	109429	9.6	9.6	-
63.6 MHz 5th	CC-5	10.3	8.8	3.2
	CC-7	8.8	8.8	3.6
	CC-8	5.4	7.3	-

^a x 10⁻¹⁰ / g

Figure 11 - Acceleration Sensitivities of Various Overtones of the Same Resonator.

DIRECTION OF MAXIMUM SENSITIVITY TO IN-PLANE ACCELERATION—SC-cut

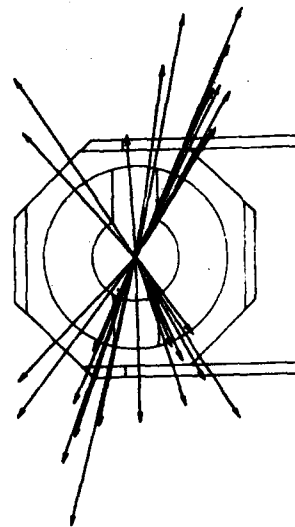


Figure 12 - Direction of Maximum Sensitivity to In-Plane Acceleration for 4-point Mounted SC-Cut Resonators.

IN-PLANE ACCELERATION - SC-out

Angle off Z' (degrees)	Number of Hits
0	2
5	2
10	1
15	1
20	1
25	1
30	8
35	3
40	2
45	2

Effects of Changes in Circuit Parameters on SC-out Acceleration Sensitivity

CIRCUIT MODIFICATION	ACCELERATION SENSITIVITY*
GROUNDING ENCLOSURE	$29.0 \times 10^{-10}/g$
FLOATING ENCLOSURE	$29.4 \times 10^{-10}/g$
LOAD CAPACITOR	$29.4 \times 10^{-10}/g$
NO LOAD CAPACITOR	$30.0 \times 10^{-10}/g$

Figure 15 - Effects of Changes in Circuit Parameters on SC-Cut Acceleration Sensitivity

[illegible]

ACCELERATION SENSITIVITY vs.
ANGLE OF CUT

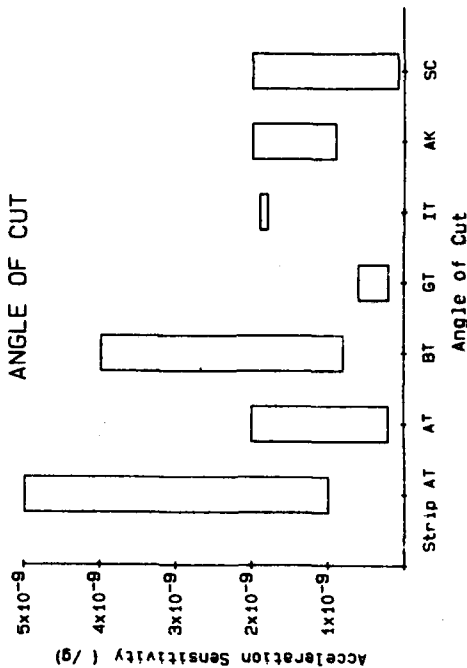


Figure 16 - Acceleration Sensitivity vs. Angles of Cut.

DETERMINATION OF ACCELERATION SENSITIVITY OF BULK MODE RESONATOR PLATES*

H. Rosman, J.T. Haynes
Systems Development Division
WESTINGHOUSE DEFENSE AND ELECTRONICS CENTER
Baltimore, Maryland 21203

SUMMARY

An approximate theory will be discussed for the determination of acceleration sensitivity of a high overtone yttrium aluminum garnet (YAG) bulk mode resonator plate operating directly at microwave frequencies. The high overtone bulk acoustic resonator facilitates implementation of low phase noise multiple frequency microwave sources with much less hardware than has been required by other means. Details of the oscillator design and tests are discussed in a companion paper, see reference¹. The vibration sensitivity is treated both analytically and experimentally in this paper.

Electrical measurements were performed on a YAG bulk mode resonator device in both quiescent and vibrating states. The computed value for the acceleration sensitivity (K) of approximately $4 \times 10^{-11}/G$ compared favorably with the tested value of $1.28 \times 10^{-11}/G$. The measured value for the bulk mode plate is about two orders of magnitude less sensitive to motion than a typical 3 point AT cut quartz crystal.

INTRODUCTION

In radar applications of oscillator devices, vibration will cause frequency modulation (FM) of the resonator. The FM, if excessive, will produce noise sidebands of sufficient magnitude to mask real targets during radar operation. The degree of sensitivity of the oscillator to vibration is primarily dependent upon the acceleration sensitivity (K) of the resonator.

The resonator is driven by very thin, on the order of 4000Å, ZnO transducers. The problem of deformation (strain) is solved in terms of its relationship to the frequency equation in the transposed direction. Several simple formulas are obtained relating elastic deformation to the acceleration sensitivity. Determination of the strain was achieved by first calculating the plate stresses. A brief description of the method for determining the plate stresses will also be included in the discussion which follows.

ANALYTICAL APPROACH

For this analysis the plate was assumed to be isotropic. It was further assumed that the plate was fixed at its interface with the mounting preform, on three sides, as depicted in figures 1 and 2. The electrically active area is about 8×10^{-6} in.² The problem of deformation caused by vibration of the resonator plate is solved in terms of its relationship to the frequency equation in the transposed (z) direction. The resonator frequency is

$$f = \frac{V_z}{2t} \quad (1)$$

where V_z is the velocity in the z direction and t is the initial plate thickness. The acceleration sensitivity is strain dependent and is caused primarily by the stresses that are developed at the plate surfaces, i.e., at the transducer mounting surfaces, during vibration. In terms of elastic constants, the generalized Hooke's law for strain in the transposed direction may be written as

$$\epsilon_z = \frac{\Delta t}{t} = \frac{1}{E} [S_z - \mu(S_x + S_y)] \quad (2)$$

where Δt is the change in thickness caused by the normal stresses S_x , S_y and S_z , E is the modulus of elasticity and μ is Poisson's ratio. Once the plate is stressed equation 1 becomes

$$f_s = \frac{V_z}{2(t \pm \Delta t)} \quad (3)$$

The + or - notations indicates an increase or a decrease in thickness respectively. The change in frequency (Δf) can be expressed as

$$\Delta f = f - f_s = \frac{V_z}{2} \left[\frac{1}{t} - \frac{1}{t \pm \Delta t} \right] \quad (4)$$

The acceleration sensitivity constant is governed by the known expression

$$K = \frac{\Delta f}{f} \left(\frac{1}{G} \right) \quad (5)$$

This equation represents the fractional frequency change per gravitational unit and by substituting equations 1 and 4 into equation 5 we now obtain the acceleration sensitivity constant in terms of strain where

$$K = \frac{\pm \Delta t}{t \pm \Delta t} \left(\frac{1}{G} \right) \quad (6a)$$

for $\Delta t \ll t$

$$K \approx \frac{\pm \Delta t}{t} \left(\frac{1}{G} \right) \quad (6)$$

Finally by substituting equation 2 into 6 an approximate expression relating the acceleration sensitivity

*This work was supported in part by contract F19628-82-C-0066, Rome Air Development Center, Electromagnetic Sciences Division, Hanscom AFB, MA.

to the generalized Hooke's law for strain is obtained where

$$K \approx \frac{\epsilon_z}{G} = \frac{1}{E} \left[S_z - \mu (S_x + S_y) \right] \left(\frac{1}{G} \right) \quad (7)$$

In order to obtain ϵ_z/G the normal elemental plate stresses at the location of interest (the electrically active area of $8 \times 10^{-6} \text{ in}^2$), were calculated. The technique used to determine the stresses was finite element analysis (FEA). In brief, FEA is a method of representing the plate structure by a network of finite structural elements. Figure 2 is a representation of the bulk mode resonator plate and its boundary support structure, depicted in figure 1. The elements are interconnected at nodal points each of which is identified by a number. For this analysis a 1G gravitational field was applied to the configuration. The FEA yielded values for S_x , S_y and S_z . These results were then inserted into equation 7 in order to obtain value for K. For the resonator plate described in this paper a value of $4 \times 10^{-11}/G$ was obtained.

EXPERIMENTAL SET-UP

Phase perturbations caused by sinusoidal vibration inputs, see figure 4, were measured with the test set-up shown in figure 3. This arrangement depicts a standard phase bridge implementation which is frequently used to determine the additional phase noise contributed by crystal oscillator devices during vibration. The system is calibrated by injecting a phase sideband of known magnitude and then observing the output of the low noise video amplifier with the spectrum analyzer. This yields a calibration point at which phase noise can be referenced.

The HP 3585-A can display phase noise as $V/\sqrt{\text{Hz}}$ where V is volts. Using the calibration, the $V/\sqrt{\text{Hz}}$ output is then converted to a fractional frequency sensitivity per G of acceleration.

In order to minimize any frequency modulation which could be caused by excessive motion of the RF cable during vibration appropriate mechanical clamping in conjunction with a low frequency cable loop was provided as indicated in figure 3. In order to establish the level of noise introduced by the cable/connector system a substrate structure (similar to the bulk mode resonator's structure) employing a transmission line was positioned where the resonator would normally be located. This system was measured in both quiescent and vibratory states. The resulting noise level was negligible. Once this was accomplished the actual device was vibrated. From the data observed the acceleration sensitivity constant was determined.

EXPERIMENTAL DATA

Setting up -40dB_c sidebands on RF signals and observing the discriminator output thru the IMA (on the analyzer) a calibration output at various vibration frequencies was obtained. Data taken at several frequencies of interest are cited below.

Vibration Frequency (HZ)	Calibration Output (CAL) $V/\sqrt{\text{Hz}}$
1000	0.147
2000	0.144
3000	0.147

The calibration sideband was then removed and the measurement $V/\sqrt{\text{Hz}}$ taken at the appropriate vibration frequency.

Vibration Frequency (HZ)	$V/\sqrt{\text{Hz}}$	Vibration Response at Resonator Substrate (G's)
1000	0.221	3.0
2000	0.278	3.8
3000	0.510	4.4

Note: 3G's rather than 1 G was used for the input from the vibration machine in order to obtain an observable response out of the phase noise test set-up. See figure 3 for the response data.

$$\text{Vibration Sideband} = SB_V = \text{CAL} + 20 \text{ LOG } \frac{V/\sqrt{\text{Hz}}}{\text{CAL}} = 20 \text{ LOG } \frac{f}{2f_V}$$

where f_V is the frequency of vibration, f is the resonator frequency ($1.65 \times 10^9 \text{ Hz}$) and

$$\Delta f = 2 f_V (10)^{-SB_V/20}$$

Sample calculation of K from the test data at $f_V = 2000 \text{ Hz}$:

$$SB_V = -40 + 20 \text{ LOG } \left(\frac{278 \times 10^{-6}}{144 \times 10^{-6}} \right) = -94.2 \text{ dB}$$

$$\Delta f = 2 \times 2000 \times 10^{-94.2/20} = 0.08$$

$$K = \frac{\Delta f}{f} \left(\frac{1}{G} \right) = \frac{0.08}{1.65 \times 10^9} \left(\frac{1}{3.8 G} \right) = 1.28 \times 10^{-11}/G$$

The experimental value for the acceleration sensitivity constant did appear to vary slightly with frequency of vibration, however, this was attributed primarily to errors associated with data observation and small variations in resonator temperature.

AD-A136 673

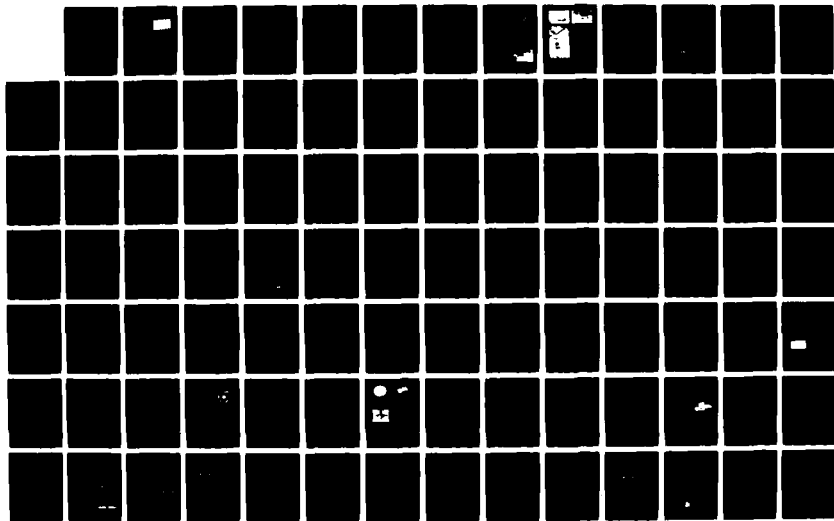
PROCEEDINGS OF THE ANNUAL SYMPOSIUM ON FREQUENCY
CONTROL (37TH) 1-3 JUNE 1..(U) ARMY ELECTRONICS
RESEARCH AND DEVELOPMENT COMMAND FORT MONMOUTH, 1983

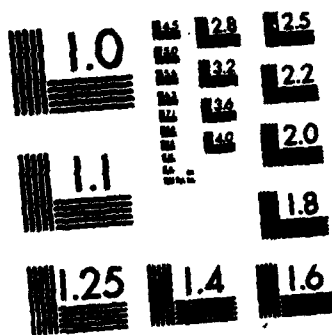
4/6

UNCLASSIFIED

F/G 9/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

CONCLUSION

While the calculated approach is an approximate solution, the resulting value for K equal to $4 \times 10^{-11}/G$ does compare very favorably with the experimental value of $1.28 \times 10^{-11}/G$. The foregoing method of calculating the acceleration sensitivity was applied to a 5 MHz convex-plane, AT cut 3 point mounted quartz resonator device. The calculation yielded a value of about $.5 \times 10^{-9}/G$ in the direction normal to the plane of the crystal, compared to generally published value of approximately $1 \times 10^{-9}/G$. The calculation procedure developed as part of this task appears to yield reasonable results for quartz resonator applications as well.

ACKNOWLEDGMENT

The author wishes to thank Brian T. Drude of Westinghouse Electric Corp. for the development of the finite element model of the resonator plate which was used to determine the stresses.

REFERENCES

1. J.T. Haynes, H. Salvo, R.A. Moore and B.R. McAvoy, "Low Phase Noise Multiple Frequency Microwave Source," Proc. 37th Ann. Freq. Control Symposium, 1983.
2. B.R. McAvoy and J.de Klerk, "High Q Microwave Bulk Mode Resonators," 1976 Ultrasonics Symposium Proc. IEEE.
3. Raymond L. Filler, "The Effect of Vibration on Frequency Standards and Clocks," Proc. 35th Ann. Freq. Control Symposium, 1981.

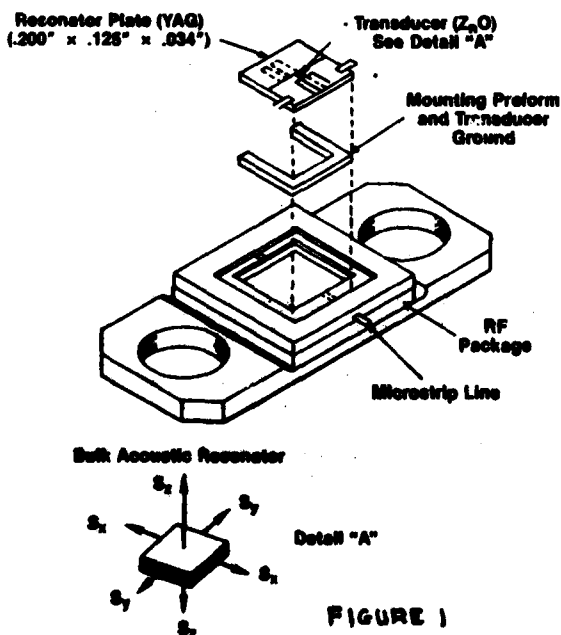


FIGURE 1

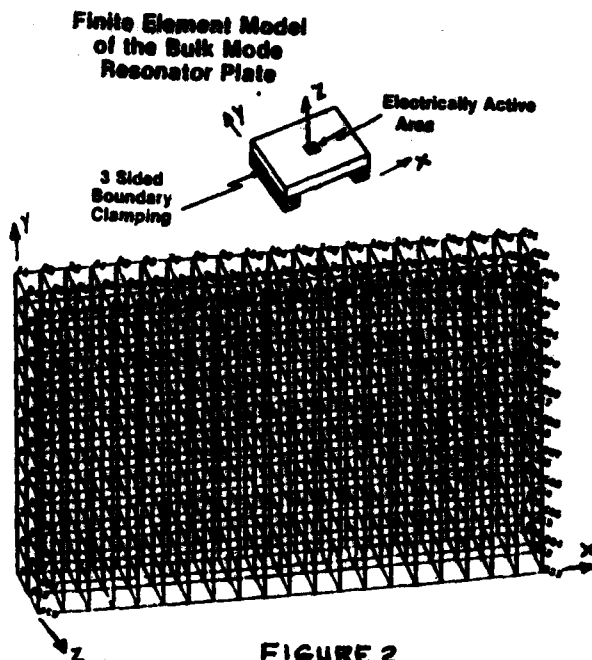


FIGURE 2

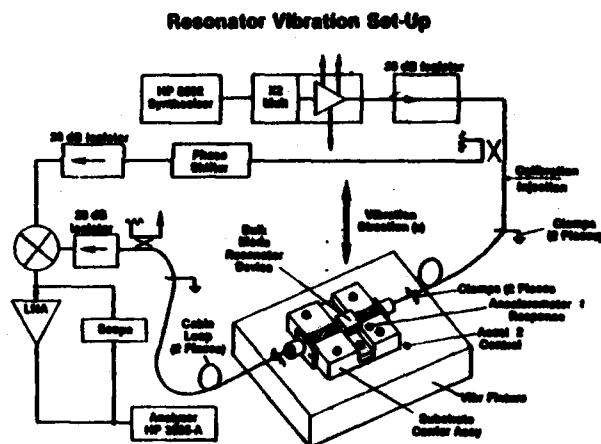


FIGURE 3

Sinusoidal Vibration Input and Response Data

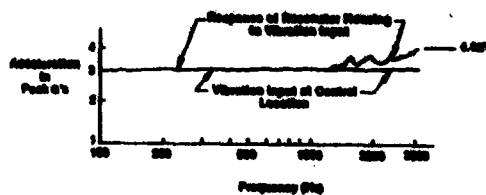


FIGURE 4

AD P 002488

AN INSTRUMENT FOR AUTOMATED MEASUREMENT OF THE ANGLES OF CUT OF DOUBLY ROTATED QUARTZ CRYSTALS

John L. Chambers

Advanced Research and Applications Corporation
1223 East Arques Avenue
Sunnyvale, CA 94086

Summary

An automated x-ray orientation system for accurate measurement of the angles of cut of both singly and doubly rotated quartz crystal plates is described. The instrument is capable of determining the conventional phi and theta angles-of-cut to a one-standard-deviation precision of 3 - 4 seconds of arc and is capable of measuring the general orientation of any cut on the bulk-wave zero-temperature-coefficient locus. Blanks are assumed to be pre-cut to within 15 minutes of the desired angles-of-cut.

The instrument, a laser-assisted Laue diffractometer, has been in operation for approximately one year. Thousands of measurements have been made on both singly and doubly rotated cuts of quartz, as well as other crystalline materials, with excellent results. Automatic handling equipment is currently being added, which will permit the instrument to run unattended for up to 16 hours.

Introduction

In this paper we describe a laser-assisted Laue diffractometer for accurate and automated measurement of the angles of cut of quartz crystals, including doubly rotated cuts -- in particular the SC cut. Earlier work on this instrument has been described in the Proceedings from last year's Symposium¹ and elsewhere.^{2,3,4,5} This instrument, designated the EXAQ-1200, utilizes a continuous x-ray spectrum rather than a monochromatic spectrum as in conventional diffraction instruments, permitting x-ray reflections to be easily located with use of a sample mount having only one degree of freedom in motion of the blank.

The primary function of the instrument is, given a blank cut to within about 15 arc-minutes of the correct nominal angles-of-cut (the conventional phi, theta, and psi), to measure these angles to a few arc-seconds accuracy in order to determine any required angular corrections. These measurements are made with respect to a reference standard crystal for each type of cut to be measured. The design goals we wished to meet included:

- 1) Types of Cuts: Blanks of any general orientation, in particular those on the bulk-wave zero-temperature-coefficient locus. Circular blanks as well as a variety of blank sizes and thicknesses must be accommodated.
- 2) Reproducibility and Accuracy: Reproducibility and accuracy of plus-or-minus 4 arc-seconds and 10 arc-seconds, respectively in the conventional phi/theta angles of cut for SC-cut plates, where the values refer to one standard deviation from the mean.

- 3) Rate of Measurement: 50 plates per hour.

Thus, the design goals specified an instrument with greater flexibility and accuracy than commonly found in conventional x-ray orientation equipment. We have achieved the first two goals set forth above. The time per measurement required to obtain this high precision is currently 2.5 minutes, although the instrument can measure at higher speeds if the requirement on precision is relaxed. To date, accuracy rather than speed has been the primary consideration, and significant improvements in the measurement rate are possible, keeping in mind that measurement of a general three-dimensional orientation is inherently a longer process than determination of a single angle-of-cut.

The Laue Method

The differences between the Laue method and conventional diffraction techniques as they apply to this system have been described earlier.^{1 - 5} An excellent general text on the theory and application of the Laue method is available.⁶ The Laue method has long been used for orientation of crystals to about 0.25 degrees, usually employing visual analysis of a diffraction pattern on film or on a fluorescent screen. Higher-precision Laue work (35-arc seconds in favorable cases) has been previously performed where the desired crystal orientation places a crystallographic symmetry axis exactly along the incident x-ray beam, by matching the intensities of symmetry-equivalent reflections⁷, but that method is less useful for an instrument which must measure a variety of general orientations.

Because the Laue method uses a continuous incident x-ray spectrum, small changes in crystal orientation do not in general move the crystal in and out of the diffracting condition. This is in contrast to the case of conventional Bragg diffraction, where a small change in the orientation of a diffracting crystal in general causes the diffraction maximum to disappear entirely. The search for maxima can thus require three degrees of freedom in motion of the sample and considerable amount of time. In the Laue case, since the peak is visible over a wide range of orientations, only a single degree of freedom in sample motion, a rotation about the psi axis, is required when the orientation in phi and theta is known to better than the separation of the prominent peaks in the Laue diffraction pattern. If it were not necessary to measure circular blanks; i.e., if psi were also known to 0.25 degrees at the outset, no motion of the blank would be required, considerably reducing measurement times. Determination of the locations of two diffraction peaks is sufficient

to completely specify the three-dimensional orientation of a crystal.

The major advantage of the Laue method is thus the fact that peaks are very readily located, leading to a mount with only one rotational degree of freedom, and therefore to a design which provides a great deal of accessibility to the sample. The major disadvantage lies in the lower diffracted intensities. Careful choice of instrument geometry and x-ray reflections has been made to maximize intensities.

Laue Instrument Concept

Given a Laue diffraction peak, as the orientation of the crystal is changed by a small amount the position of the peak changes just as if the reflection were occurring from a mirror parallel to the diffracting lattice planes, slightly changing the peak energy of the diffraction. At a 400 mm crystal-to-detector distance, as in the system to be described, the x-ray spot moves about 4 microns for a one arc-second change in crystal orientation in the plane of reflection. Figure 1 shows the concept employed in the Laue diffractometer for completely determining the crystal lattice orientation. A minimum of two x-ray detectors is required, although a larger number can be used (three are available on our system). The blank is mounted on a rotating stage with the axis of rotation approximately parallel to the psi-axis of the blank. The x-ray detectors, with apertures large enough (plus or minus 1 degree in this case) to accommodate any mounting errors as well as errors in the angles of cut, are positioned so that only at some unique stage rotation angle will x-ray reflections enter all counters simultaneously. The exact position of each reflection within its counter aperture is then determined. These measurements permit the normals to two or more sets of diffracting planes and, thus, the crystal orientation, to be computed.

Assuming that the blank has been mounted so that the normal to its face is exactly parallel to the rotator axis, this information is sufficient to determine the angles of cut. However, mounting errors are often in the one-minute range and a correction is therefore required. The method used, shown in Figure 2, employs a He/Ne laser and linear position-sensitive detector to measure the precession of a laser beam reflected from the face of the blank as it rotates, to yield both the height and orientation of the face. Determination of the severity of errors arising from sample mounting and the method of laser-assisted diffractometry have been described previously by Vig⁶.

Description of the Laue Diffractometer

The instrument has been described earlier in detail⁵. A brief summary will be presented here. An overall view of the prototype Laue diffractometer is shown in Figure 3. Although the He/Ne laser itself is not visible from this angle, the other major features of the instrument can be seen. A steel table serves as the base of the instrument. Mounted in the center of this table is a high-precision rotating stage which, like the other moving parts of the instrument, is under computer control. Just to the left of the rotary stage is the x-ray source collimator, attached to the x-ray source itself. The diffracted x-rays are detected by three x-ray detectors mounted on the three arch supports which can be seen at the right of the rotary stage. Each of the arch supports permits its detector to be varied in azimuth and inclination over a wide

range while maintaining a fixed sample-to-detector distance of 400 mm. These movements allow the instrument to be readily adapted to a wide variety of cuts, and a repositioning of the detectors is the only change in the instrument required to change from one type of cut to another. This method of measurement and reconfiguration makes no distinction between singly and doubly rotated cuts; i.e., every blank is measured in terms of its general three-dimensional orientation. While speed and accuracy therefore do not improve for a singly rotated cut, both angles of cut are obtained, and the instrument has proven uniquely useful, for example, in measurement of phi in AT-cut crystals.

We are in the process of adding automatic blank handling equipment to the instrument, so that it can run unattended for up to 16 hours. The handling equipment consists of four elevators which will hold stacks of blanks to be measured, four receivers into which the measured blanks will be sorted and a robot arm which can move in radius, azimuth, and height, which will operate in a pick-and-place mode.

A 2mW He/Ne laser (not visible here) is mounted near the central arch support, and the reflected laser beam is detected by an optical system including a position-sensitive diode detector, shown at the far left. Both the laser source and the laser detector are inclined at 8 degrees from horizontal with respect to the sample. The laser detector consists of a lens system to magnify the motion of the reflected laser beam, a bellows to decrease the amount of ambient light entering the system, and an electronics package consisting of a linear position-sensitive photodiode mounted with the position-sensitive axis vertical and circuitry to convert the diode output into a voltage proportional to spot position. This system enables the position of the reflected laser beam to be monitored by the computer in real-time during the initial rotational search for x-ray reflections.

Figure 4 is a closer view of the rotary stage on which the sample is mounted. On top of the stage is a three-point vacuum chuck, on which the blank is mounted. The collar around the sample mount is threaded to permit adjustment of the sample height. The viewing angle in this figure is approximately the same as in Figure 3, with the x-ray source collimator at the left, and the incoming He/Ne laser beam at the right.

The x-ray source is a standard sealed diffraction tube, excited below the characteristic K-line potential in order to produce a relatively smooth "white-radiation" spectrum. The incident x-ray beam approaches the sample at a glancing angle of 20 degrees. At the exit port of the x-ray tube housing is the x-ray shutter, which is also under computer control and has both "open" and "closed" fail-safe sensors for operator safety.

In front of each x-ray detector is a translating stage, which scans a pair of slits over its detector, in order to measure the spatial x-ray intensity distribution within the aperture. A least-squares profile fitting technique is employed to determine the position of the x-ray spot within the detector. Figure 5 is a close-up view of one of the outside translating stages, taken from the far "right" of the instrument. The nature of the slits used to determine the x-ray position can be seen in this figure. The two slits are 90 degrees to one another and each is at 45 degrees with respect to the direction of stage travel. This arrangement allows the X-Y coordinates of the x-ray spot within the detector to be determined with a single translational motion. Scans are performed on

all three detectors simultaneously, first with the "upper" slit, then with the "lower." Scan times can be varied to match the speed and accuracy requirements, and have ranged from 15 seconds to 2 minutes per scan.

The control subsystem is shown in Figure 6. A dedicated microcomputer is connected to the instrument through a CAMAC interface, which is a highly flexible standardized digital interfacing system. The CAMAC crate accepts standardized plug-in modules which in this case include four stepping-motor controllers, a digital I/O port (for control of the x-ray shutter, etc.), an A/D converter for the laser detector, and a six-channel scaler which counts the pulses received by the x-ray detectors. Also present are a NIM bin containing the high-voltage power supply and pulse-height discrimination circuitry for the x-ray detectors, and a power supply for the x-ray shutter and stepping motors. A CRT terminal for operator input and a 30-cps keyboard/printer for a hard-copy record of measurement results are also present.

The primary function of the control subsystem software is to perform the automatic measurement sequence, first for a reference standard crystal, then for the blanks to be measured. The program outputs a hard-copy record of the measurement results for each batch of crystals. The operator is prompted when any intervention is required, such as mounting the next blank when operating without the automatic blank handling equipment. Since we are dealing with low-level signals compared to those commonly observed in conventional orientation equipment, and we require very high accuracy, the computer is essential in extracting as much information from the signal as possible. Least-squares curve fitting techniques are employed to extract this information, for example, in the determination of crystal orientation from the computed diffraction vectors. The set of vectors obtained from the measured blank is treated as a rigid body, which is rotated to achieve the best match (in a least-squares sense) with the set of vectors measured from the reference standard crystal. The three rotation angles determined from this procedure correspond to the changes in the angles of cut (ϕ , θ , and ψ) between the measured crystal and the reference standard.

Quartz Crystal Measurement Results

The EXAQ-1200 has made several thousand measurements of various types of cuts of quartz. Whereas the measurements described at last year's Symposium tested repeatability without remounting the blank between determinations⁵, the ones described here involved removing the blank from the chuck, rotating it through an arbitrary angle in ψ , and replacing it on the chuck between repeated measurements. While the tilt due to remounting was often one arc-minute or more, the actual measurement precision was considerably better (e.g., a few arc-seconds) indicating that the laser subsystem was performing well. The measurements utilized data from all three x-ray detectors, and the measurement rate selected resulted in measurement times of 2.6 minutes. Among the types of samples measured to date have been SC cuts with $\phi = 21.93$ degrees and $\theta = 34.11$ degrees, SC cuts near $\phi = 23.75$ degrees, SC cuts near $\theta = 33.5$ degrees, and AT cuts. One or more crystals of each type of cut were selected for use as reference standards.

The only requirement which must be fulfilled for a given type of cut to be measurable (in addition to

the limits on physical characteristics and dynamic range) is that three x-ray reflections of suitable intensity be accessible to the x-ray detectors simultaneously. This requirement has to date been very easy to meet. Several such sets of reflections have been located for each cut tried so far, and the best set chosen in each case.

As an example of the high degree of flexibility of the instrument, experiments have been conducted not only on different types of cuts of quartz, but also on an entirely different material -- cadmium telluride (CdTe). Unlike quartz, which has a trigonal crystal lattice, the CdTe lattice is face-centered cubic with the desired orientation normal to the 111 planes. The 111 normal is an axis of three-fold rotational crystallographic symmetry. This symmetry was used to make an absolute measurement of the CdTe angles-of-cut much in the same manner as an absolute measurement of ϕ can be made in the case of AT-cut quartz, as discussed in the section on AT-cut measurement results. Six CdTe samples were measured with a repeatability of 5 arc-seconds in the angles-of-cut, and an absolute accuracy of 17 arc-seconds (root-mean-square values over the six samples measured) without availability of a reference standard crystal. Thus, the instrument appears capable of measuring angles-of-cut not only of quartz crystal plates of any orientation, but also of a variety of other crystalline materials as well. The remainder of this section will describe the results obtained on approximately 150 well-studied SC- and AT-cut quartz plates.

SC-CUT RESULTS

SC-cut blanks of both known and roughly known angles-of-cut were measured. The first batch was a set of 11 SC-cut blanks, each cut at a different angle near 21.93 degrees ϕ and 34.11 degrees θ , very kindly provided by Frequency Electronics, Inc. (FEI). The blanks had been selected to cover the entire dynamic range of plus-or-minus 15 arc-minutes about the nominal SC-cut angle and had been measured to about 20-30 arc-seconds in θ and 1 - 2 minutes in ϕ . These blanks were 14 mm (0.550 in) in diameter, 1.27 mm (0.050 in) thick, parallel to better than 0.5 micron, and had a 3-micron surface finish. Performance on these blanks was excellent.

Several sets of measurements were performed. For each set, the instrument was calibrated relative to a different one of the blanks. The results relative to blank #25 as the reference crystal are shown in Figure 7, a scatter plot of the twenty measurements made on each blank. The open circles represent the nominal angles, and a plus sign is plotted at each of the measured angles. A one arc-minute scale is shown near the lower right. The tight clustering of the plus signs is representative of the instrument reproducibility. The differences between the nominal angles and those measured on the EXAQ-1200 are consistent with the estimated precision in the nominal angles, and it can be seen that the largest differences are in ϕ (the less critical angle). Figure 8 summarizes these results numerically. Over all 220 measurements, the average reproducibility in ϕ was 3.6 arc-seconds, and in θ was 4.0 arc-seconds. The average total spread (maximum minus minimum) was 13.7 seconds in ϕ and 15.1 seconds in θ . The maximum observed ϕ spread for 20 measurements was 18.7 seconds and maximum θ spread was 19.4 seconds. The values in Figure 8 corresponding to $\Delta\phi$ and $\Delta\theta$ represent the difference between the angle-of-cut as measured on the EXAQ-1200 and the nominal angle. Similar results were obtained relative to other blanks (#21 and #31)

taken as the reference crystal. In all, 330 measurements were made, giving the following overall statistics (all in arc-seconds):

	<u>Phi</u>	<u>Theta</u>
Average Reproducibility	3.6	3.6
Average Spread	10.9	11.2
Maximum Spread	18.7	19.4

The fact that the spreads were within plus-or-minus three standard deviations from the mean suggests that the errors in measurements made within the dynamic range of the instrument conform to a normal distribution. Psi was measured absolutely to plus-or-minus 0.25 degree and was 70.6 degrees in all cases, at that orientation where the x-ray scans were performed. Overall, the root-mean square difference between the angles measured for these blanks on the EXAQ-1200 and the nominal angles was 1'48" in phi and 24" in theta. These differences agreed well with the estimated errors in the nominal angles.

The other set of SC-cut blanks for which angles were known was generously provided by Hewlett-Packard (HP). The set consisted of six blanks, all cut at angles phi and theta of 21.8611 and 34.1281 degrees respectively, as measured by HP. The blank-to-blank variation (one standard deviation) in angles-of-cut as measured by the EXAQ-1200 was only 9.6 seconds in phi and 6.4 seconds in theta, over all six blanks. The consistency within this batch was thus excellent and the results were consistent with the two-sigma (i.e., two standard deviations) in combined theta and phi errors of 12 arc-seconds estimated by HP for this set of blanks. Each blank was measured 20 times, and the measurement repeatability was very good, averaging 3.5 seconds in phi and 4.0 seconds in theta.

To test the stability and repeatability of the instrument, many measurements were made on a set of over 100 different SC-cut blanks whose angles-of-cut were known only approximately. These blanks were circular plates supplied by USAERADCOM, Fort Monmouth, which had been purchased from Sawyer Crystal Systems and lapped to a uniform five-micron surface finish at FEI. Additional physical characteristics were:

Diameter:	13.97 mm (0.551 in)
Thickness:	1.02 mm (0.040 in)
Nominal Phi:	21.93 deg (21 56')
Nominal Theta:	34.017 deg (34 01')

The results for these blanks are summarized in Figures 9 - 12. Performance on these samples was excellent over the entire batch of 1040 measurements, with an average one-standard-deviation reproducibility of 4.6" in phi and 4.2" in theta, and an average spread of 14.8 seconds in phi and 17.2" in theta, based on ten measurements per blank (see Figures 9 and 10). The maximum observed spreads in phi and theta for a ten measurement series were 22" and 30", respectively. The overall average angles-of-cut for the batch were phi=22.3575 and theta=33.9770, relative to FEI SC reference crystal #25 (phi=21.9601, theta=34.0912, as described in the previous section). The blank-to-blank standard deviation was 0.1395 degrees (8'22") in phi and 0.0197 degrees (1'11") in theta, taken over all blanks in the batch (see Figures 11 and 12).

AT-Cut Results

The other type of cut measured was the AT-cut. This cut is singly rotated, and while theta can be measured by conventional instrumentation very accurately, phi is typically not measured or controlled. The AT-cut Laue pattern contains an axis of mirror symmetry, which can be used to make absolute measurements of phi with the EXAQ-1200, based on measurement of both sides of the reference crystal. In the experiment described below, thirty-one circular AT-cut plates provided by USERADCOM, Fort Monmouth, and lapped to a 5-micron surface finish at FEI were used to obtain a batch of 620 measurements, 10 measurements on each side of each blank. Physical characteristics were:

Diameter:	13.97 mm (0.55 in)
Thickness:	1.09 mm (0.043 in)
Nominal Phi:	0.0
Nominal Theta:	35.383 deg (35 23')

The instrument performance on these samples was excellent, with an average reproducibility of 4.6" in phi and 3.0" in theta, and an average spread of 15.3" in phi and 9.6" in theta, based on ten measurements per side. The average value of phi overall was very close to zero -- 0.0001 degrees, and the average theta was 35.3989 degrees. The root-mean-square value of phi was 0.0747 degrees (4'29"), indicative of the fact that phi was not controlled during processing. The maximum observed spreads in phi and theta for the whole batch were 24" and 18" respectively. The blank-to-blank standard deviation for the batch was 0.0979 degrees (5'52") in phi and 0.0073 (26") in theta. The mean difference in theta measured on the two different sides of a blank was 5.4", indicative of the average parallelism of the two sides for this batch.

Stability

To test the long-term stability of the instrument, approximately 400 measurements were made on an SC-cut plate over a three-day period. The standard deviation for the entire set was 3.7 arc-seconds in phi and 2.9 seconds in theta. Figures 13 - 14 show the progress of the measurements for phi and theta, respectively. The measurements conformed to normal distributions, indicating that the major source of error in the measurements was random. These figures also indicate that drift over this term was very small. A thermometer was used to monitor the temperature at the instrument, and no dependence of the results on temperature has been observed over the 5 degree C temperature variation measured. The stability of the instrument with respect to time and thermal effects is thus excellent, as expected based on the materials used in its construction.

Precision vs Rate of Measurement

The measurements we are now routinely making on the EXAQ-1200 require 2.6 minutes each to achieve precision near 4 seconds of arc. X-ray scan times account for only about 1 minute. As shown in Figure 15, which is a plot of performance vs total measurement time, the largest component of the remaining overhead is the initial search for x-ray reflections, arising from the assumption that psi is unknown at the outset. Nevertheless it is feasible for many types of plates to determine psi (for example, by optical means), and a rough knowledge of psi at the start could reduce determination times by as much as 1 minute. The

anticipated effect on instrument performance is shown in the dotted curve of Figure 15, along with the performance vs total determination time in the current configuration (solid line). The only change to the system required to achieve these rates is a modification to allow the laser subsystem to determine the orientation of the crystal face from a single, stationary measurement.

In situations where the highest precision is not required, reduction of x-ray scan times can provide an improvement in the throughput. The scan time per step used in the 2.6 minute determinations is 2 seconds. If a factor of 1.41 (root 2) degradation in repeatability is acceptable, scan time per step can be cut to 1 second, saving 30 seconds per determination. If a factor of 2 degradation is acceptable, times can be cut to 0.5 second/step, saving 45 seconds per determination, etc., as shown in Figure 15.

Conclusion

The EXAQ-1200 prototype has provided automated measurements of the full three-dimensional orientation required for doubly rotated quartz crystals to a precision and at a throughput that is unmatched by any other currently available system.

This study has demonstrated that it is possible to obtain very high accuracy in measurement of the orientation of crystals by the Laue method, and the method need not be restricted to low-precision work. The prototype instrument has made several thousand measurements of different types of quartz plates, including AT and SC cuts. Measurement precision of 4 arc-seconds in phi and theta for both AT and SC cuts was achieved. Absolute accuracy of the measurements appears to be limited by the accuracy to which the angles of the reference standards are known. The distribution of measurement errors conformed well to normal Gaussian statistics; thus, the precision of measurements can be further improved to almost any desired level by averaging multiple measurements. The instrument is very versatile, since it can be readily changed by repositioning the x-ray detectors to accommodate different types of cuts of quartz, or even other crystalline materials. Moreover, measurements are made automatically, enabling the instrument to be integrated with automatic handling equipment.

Although fully suitable for measurement of singly rotated cuts, the determination of the general three-dimensional orientation is inherently a more time-consuming process than measurement of a single angle-of-cut. This consideration, along with the lower x-ray intensities obtained with the Laue technique compared to monochromatic techniques, leads to longer determination times, although the result does provide more information (e.g., phi in the case of AT cuts). Hence, the throughput of measured crystals is lower than for conventional x-ray orientation instruments designed for singly rotated crystal measurements. Measurements at the 4 arc-second level of performance require 2.6 minutes. Methods for improving these times include use of a stronger x-ray source, more x-ray detectors, faster mechanical stages, and elimination of the initial search for reflections by marking psi on the blanks.

Acknowledgements

This work was supported by the US Army Electronics Research and Development Command, Fort Monmouth, NJ. P. R. Hoffman Company, Carlisle, PA, was the prime contractor for the development phase, and was responsible for the mechanical elements of the prototype

instrument. Advanced Research and Applications (ARACOR) was a subcontractor to P. R. Hoffman Company during this initial phase, and is currently a subcontractor to Frequency Electronics, Incorporated who is the prime contractor for the automation phase. The automatic blank handling equipment has been developed by Brooks Associates, Incorporated.

I wish to thank my colleagues at ARACOR, Tom Workman, Myron Pugh, Alan Young, Kien Chau, Les Palkuti, and James Stanley, who have contributed substantially to this work.

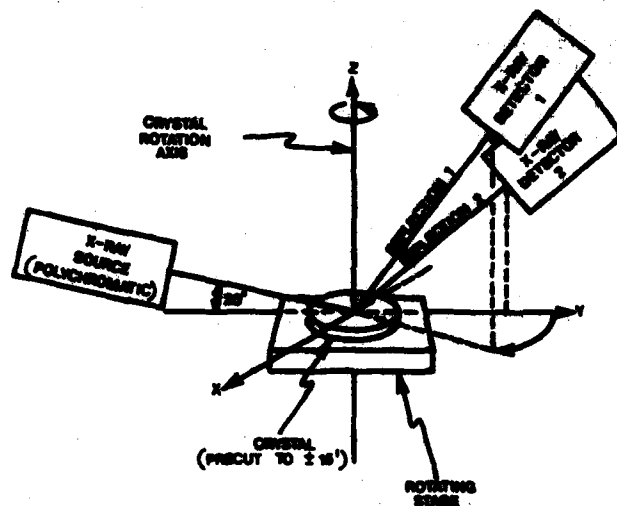
My thanks go also to Roger Ward for his advice and suggestions, including the idea that the Laue method could be used for this application; Jack Kusters and Charles Adams of Hewlett-Packard for their generous donations of reference crystals; and Martin Bloch, A. W. Warner, Jack Korman and Bruce Goldfrank of Frequency Electronics for their valuable comments, suggestions, and donations of reference and sample crystals. I am also very grateful to John Vig for his valuable advice and support throughout the course of this program.

References

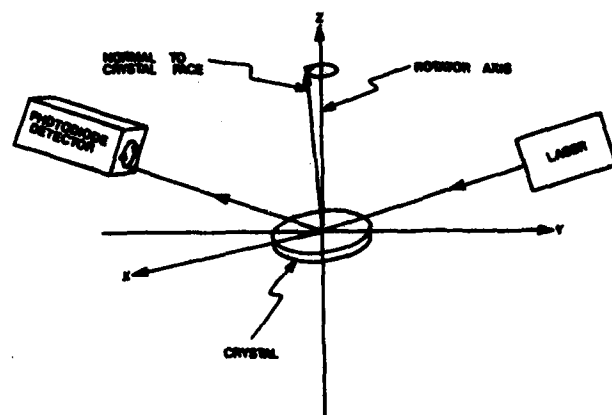
1. J. L. Chambers, M. A. Pugh, S. T. Workman, R. W. Birrell, and R. J. Valihura, "An Instrument for Automated Measurement of the Angles of Cut of Doubly Rotated Quartz Crystals," Proc. 35th Annual Symp. on Frequency Control, pp. 60 - 70, 1981. Electronics Industries Association, 2001 Eye St., NW, Washington, D.C. 20006
2. R. W. Birrell, R. J. Valihura, J. L. Chambers, M. A. Pugh, and S. T. Workman, Research and Development Technical Report DELET-TR-79-0254-1, USAERADCOM, Fort Monmouth, NJ, 1980.
3. R. W. Birrell, R. J. Valihura, J. L. Chambers, M. A. Pugh, and S. T. Workman Research and Development Technical Report DELET-TR-79-0254-2, USAERADCOM, Fort Monmouth, NJ, 1981.
4. R. W. Birrell, R. J. Valihura, J. L. Chambers, M. A. Pugh, and S. T. Workman, Research and Development Technical Report DELET-TR-79-0254-3, USAERADCOM, Fort Monmouth, NJ, 1981.
5. J. L. Chambers, "An Instrument for Automated Measurement of the Angles of Cut of Doubly Rotated Quartz Crystals," Proc. 36th Annual Symp. on Frequency Control, US Army Electronics Command, Fort Monmouth, NJ pp. 302 - 313, 1982. Electronics Industries Association, 2001 Eye St., NW, Washington, D.C. 20006.
6. J. L. Amoros, M. J. Buerger, and M. Canut de Amoros, "The Laue Method," Academic, New York, 1975, 375 pages.
7. G. Christiansen, L. Garverd, and I. Alstrup, Acta Cryst. A31, 142 - 145, 1975.
8. J. R. Vig, "A High Precision Laser Assisted X-Ray Goniometer for Circular Plates," Proc. 29th Annual Symp. on Frequency Control, pp. 240 - 247, 1975. National Technical Information Service Accession No. AD A017466.

Figure Captions

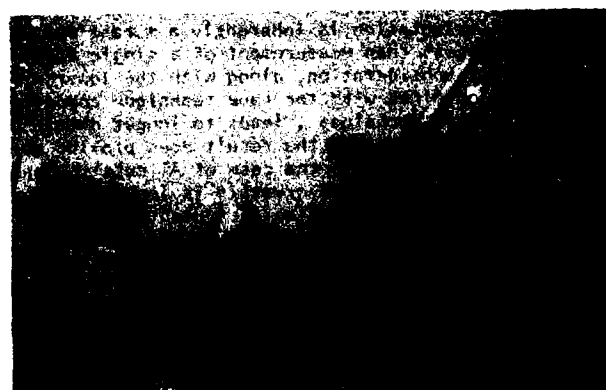
1. Concept for the x-ray portion of the measurement apparatus. A minimum of two detectors is required; three are available on the prototype instrument.
2. Concept for the laser portion of the measurement apparatus, which corrects for errors in mounting of the crystal to the rotating stage, as described in the text.
3. Full view of the prototype instrument.
4. Closeup of the rotating stage. Angle of view is about the same as in Figure 3.
5. Closeup of one of the outside translators. The pair of 1.5 mm slits used in the x-ray scans can be seen.
6. Control subsystem for the prototype instrument, with computer, computer peripherals, x-ray counting circuitry, and CAMAC digital control interface.
7. Scatter plot of the measurements on 11 SC-cut plates. The open circles represent the nominal angles. The clusters of +'s represent 20 measurements on each of the blanks. A 1 arc-minute scale is shown at the lower right. Agreement with the nominal angles was within their estimated error.
8. Comparison of nominal vs measured angles for 11 SC-cut plates.
9. Measured ϕ for 102 SC-cut plates, each point is the average of 10 measurements/plate.
10. Measured θ for 102 SC-cut plates, each point is the average of 10 measurements/plate.
11. Standard deviation in ϕ for 10 measurements each of 102 SC-cut plates.
12. Standard deviation in θ for 10 measurements each of 102 SC-cut plates.
13. Time dependence of ϕ for 392 measurements of an SC-cut plate over a three-day period.
14. Time dependence of θ for 392 measurements of an SC-cut plate over a three-day period.
15. Precision vs. total measurement time for SC-cut blanks. The solid line represents current performance; the dashed line represents the performance achievable if blanks were to be marked in psi prior to measurement.



X-RAY SUBSYSTEM CONCEPT

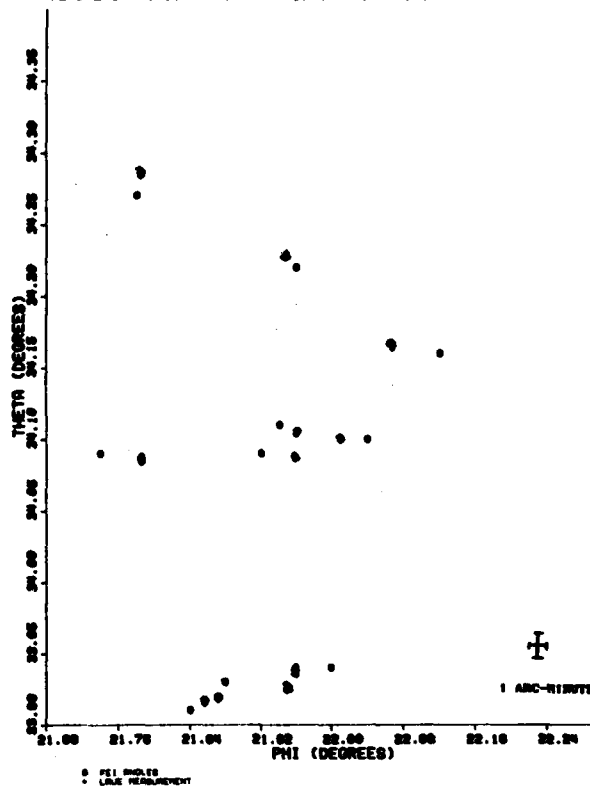


LASER SUBSYSTEM CONCEPT





REFERENCE 90-CUT CRYSTAL MEASUREMENTS 7/28/82

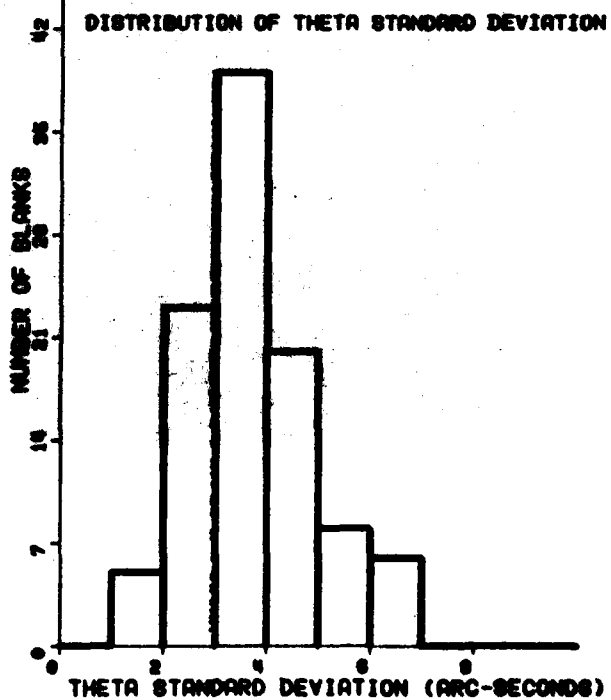
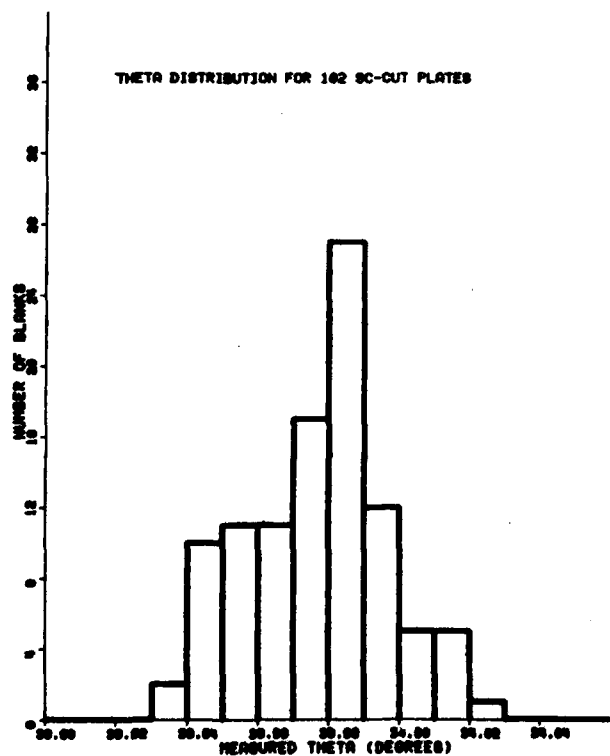
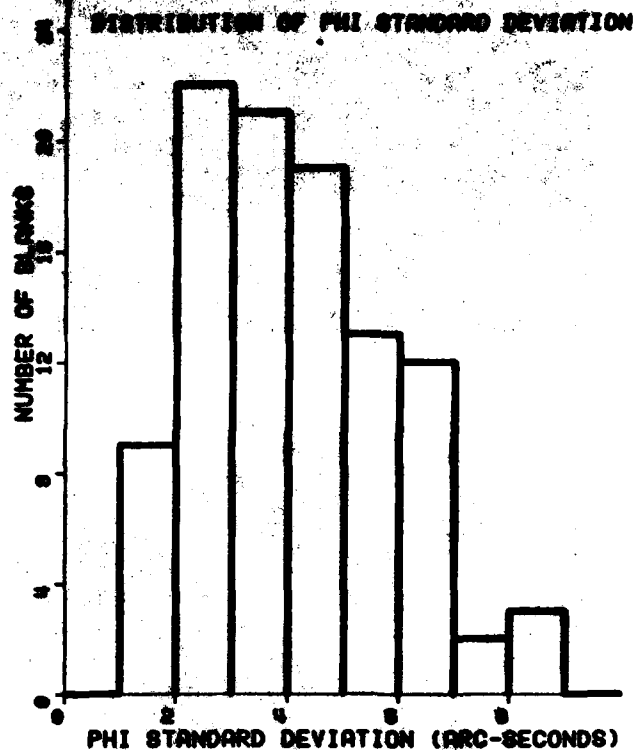
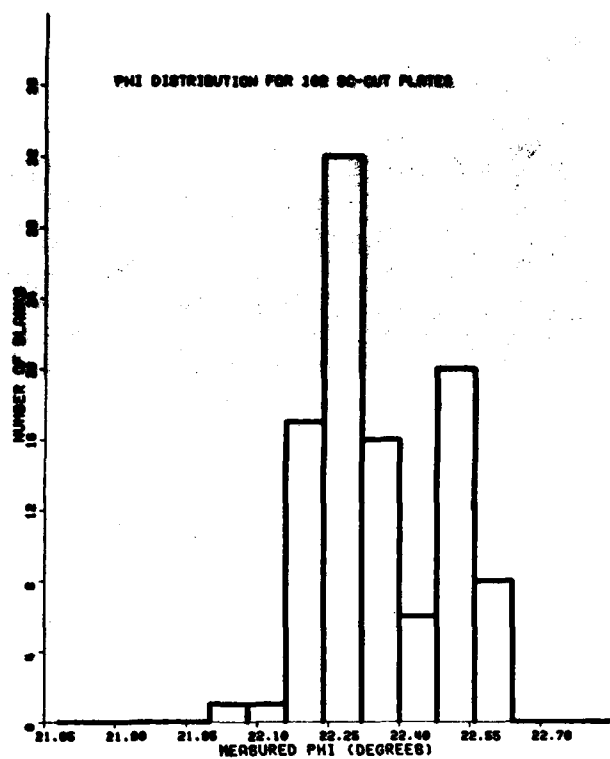


90-CUT REFERENCE CRYSTALS: COMPARISON OF REFERENCE ANGLES TO "TEST" DATA-MEASURED ANGLES

Slab #	Refined Phi (Degrees)	Refined Theta (Degrees)	TEST Phi (Degrees)	TEST Theta (Degrees)	Delta(Phi) (Seconds)	Delta(Theta) (Seconds)
3	21.704	20.574	21.7036	20.5807	16.6	26.6
4	21.948	20.941	21.9410	20.9388	-49.6	-28.6
16	21.800	20.914	21.8004	20.9087	77.6	27.6
17	21.741	20.880	21.7407	20.8760	100.0	-7.1
19	21.900	20.720	21.9000	20.7110	30.9	-5.0
21	22.004	20.901	21.9997	20.9000	-107.0	-7.1
26	22.120	20.107	22.0700	20.1097	-100.0	9.6
28	21.900	20.694	21.9001	20.6910	100.0	-10.0
29	22.004	20.100	21.9104	20.1000	-100.1	-10.0
31	21.948	20.580	21.9400	20.5804	-40.0	0.6
36	21.800	20.909	21.8000	20.9000	10.0	-20.0

Overall Root-Mean-Square 100.0 20.0

Notes:
1. "Delta" refers to TEST-REFERENCE-measured angle minus reference angle.



The graph shows the relationship between performance and total measurement time for two different methods. The y-axis represents performance in sec-seconds per unit, ranging from 2.00 to 20.00. The x-axis represents total measurement time in seconds, ranging from 0.00 to 300.00. A dashed line represents the 'CURRENT PERFORMANCE' and a solid line represents the 'PROPOSED PERFORMANCE'. Both curves show a rapid decrease in performance as measurement time increases, with the proposed method consistently showing higher performance (lower time per unit) than the current method.

Total Measurement Time (seconds)	Current Performance (sec-seconds per unit)	Proposed Performance (sec-seconds per unit)
0.00	20.00	20.00
20.00	10.00	10.00
40.00	5.00	5.00
60.00	3.33	3.33
80.00	2.50	2.50
100.00	2.00	2.00
120.00	1.67	1.67
140.00	1.43	1.43
160.00	1.25	1.25
180.00	1.11	1.11
200.00	1.00	1.00
220.00	0.91	0.91
240.00	0.83	0.83
260.00	0.77	0.77
280.00	0.71	0.71
300.00	0.67	0.67

[illegible]

100-443887-100

ALL INFORMATION CONTAINED HEREIN IS UNCLASSIFIED
DATE 08-14-2010 BY 60322 UCBAW/SJS

SECRET

A MEASUREMENT TECHNIQUE FOR DETERMINATION OF FREQUENCY vs ACCELERATION CHARACTERISTICS OF QUARTZ CRYSTAL UNITS

D.J. Healey, III
H. Hahn
S. Powell

Westinghouse Electric Corporation
Baltimore, Maryland 21203

Abstract

An apparatus that permits rapid measurement of crystal unit g-sensitivity under random vibration is described. It takes less than two minutes for a HP desk-top computer to set up a spectrum analyzer, to compute g-sensitivity vs. acceleration frequency, and also to record the results.

It is possible to obtain a noise floor of the instrumentation that allows measurements of g-sensitivity of VHF crystal units to less than $10^{-10}/g$ for accelerations of $10^{-4}g^2/Hz$ to $10^{-3}g^2/Hz$ over a frequency span of 500 to 15000 Hz.

The relationship between the measured voltages by the spectrum analyzer and the desired g-sensitivity will be discussed. Sample computer plots of g-sensitivity for AT, FC, IT and SC cut crystals indicating resonant frequencies associated with the mounting system of the crystal units will also be shown.

I. Introduction

The frequency vs. acceleration characteristic of quartz crystal units has been measured in the past either by static "2g-tipover" or centrifuge¹ tests, or by single-frequency dynamic^{2,3,4} tests applying sinusoidal acceleration forces to an entire oscillator circuit module. This paper describes a g-sensitivity measuring apparatus which tests a crystal unit by itself over a 15000-Hz band and also gives a simplified analysis of the measurement apparatus with a derivation of the lower limit of the measurable g-sensitivity.

The quantity of interest $\vec{\Gamma}$ or g-sensitivity of a crystal unit (crystal plate, support system, and package) is a vector. The measurement apparatus to be described can be used to obtain the g-sensitivity vs. acceleration frequency characteristic of a crystal unit alone in any one chosen axis when fitted with an appropriate crystal unit mounting fixture. For the application of the crystal units that required determination of their g-sensitivity, it was sufficient to measure the component of $\vec{\Gamma}$ normal to the crystal plate. A single setup is only practical for measurement of a specific component of $\vec{\Gamma}$ owing to the problem of attaching fixtures and maintaining adequately low mechanical crosstalk. Rapid measurement of all three components of $\vec{\Gamma}$ would only be practical if three separate measuring apparatus were provided.

In this paper, the symbol γ shall represent a g-sensitivity component in the direction of the applied acceleration (normal to plate) A in units of $g (= 9.8 \text{ m/sec}^2)$. Then,

$$\begin{aligned} \gamma &= \Delta f / (f_0 A) \\ &= \text{fractional frequency deviation per } g \\ &= \text{g-sensitivity} \end{aligned} \quad (1)$$

where Δf = frequency deviation from the series resonance frequency f_0 of the crystal unit under test. For the apparatus to be

described, the acceleration A is an RMS average of a gaussian random process in any one of 128 frequency bins. The bandwidth of each bin is either 60 Hz or 120 Hz. Δf and therefore γ are also functions of acceleration frequency (f_A). The apparatus makes indirect measurements of Δf and A needed to determine γ and then plots the γ vs. f_A (500 Hz to 15000 Hz) characteristic of a crystal unit in less than 2 minutes.

II. Measurement Analysis and Apparatus Setup

A measurement system block diagram and a more detailed diagram of the frequency discriminator block are shown in Figures 1 and 2.

II.1 Measurement Analysis

Eq. (1) is rewritten in Eq. (2) showing γ , Δf , and A as functions of acceleration frequency f_A .

$$\gamma(f_A) = \Delta f(f_A) / (f_0 A(f_A)) \quad (2)$$

The quantities measured directly are the RMS averages performed in the HP3582A spectrum analyzer of voltages V_A (channel A input) and V_B (channel B input), which are proportional to A and f, respectively. (See Figure 1.)

$$\begin{aligned} \text{Let } V_A &= K_A A \\ \text{and} \end{aligned} \quad (3)$$

$$V_B = K_B \Delta f.$$

Proportionality constants, K_A and K_B have to be known in order to evaluate Eq. (2) from the HP3582A spectrum analyzer data.

K_A (relating the channel A input voltage to the acceleration g level measured by a sub-miniature accelerometer (ENDEVCO) mounted very close to the crystal unit) is simply the accelerometer transducer constant (about $1.25 \times 10^{-12} \text{ coul/g}$) divided by the total capacitance seen by the accelerometer output charge. If the total capacitance is 500 pF, then

$$\begin{aligned} K_A &= 1.25 \times 10^{-12} (\text{coul/g}) / 500 \text{ pF} \\ &= 2.5 \text{ mV/g.} \end{aligned}$$

K_B (relating the channel B input voltage to f deviation of the crystal series resonance frequency) depends on the frequency discriminator constant and the LNA (Low Noise Amplifier) gain. In the frequency discriminator block of Figure 2, the input excitation frequency is constant at or very near the crystal series resonance frequency at rest. When subjected to vibration forces, the crystal series resonance frequency undergoes fluctuation. One is inclined to suspect that the actual mechanism^{5,6,7} of series resonance frequency shift results from changes in thickness, density, or even elastic constant experienced during plate deformation.

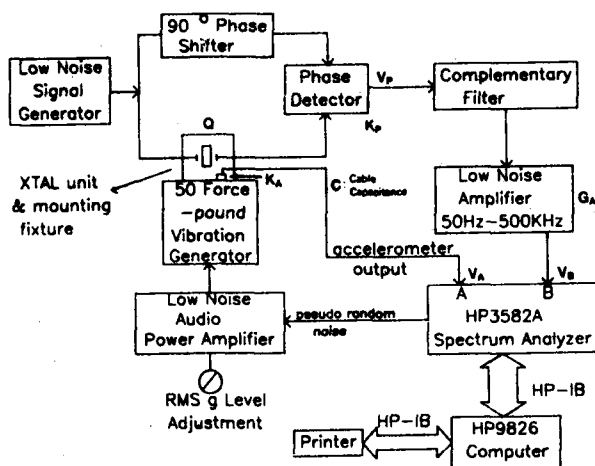


Figure 1. g-Sensitivity Measurement Apparatus

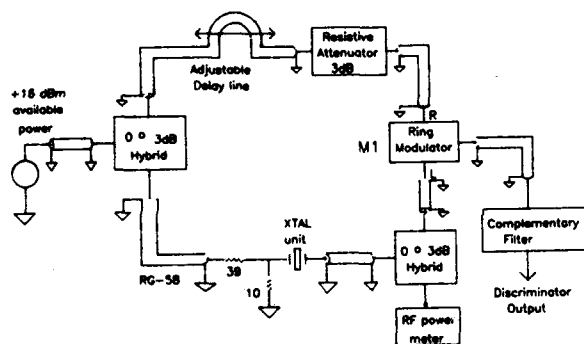


Figure 2. Frequency Discriminator

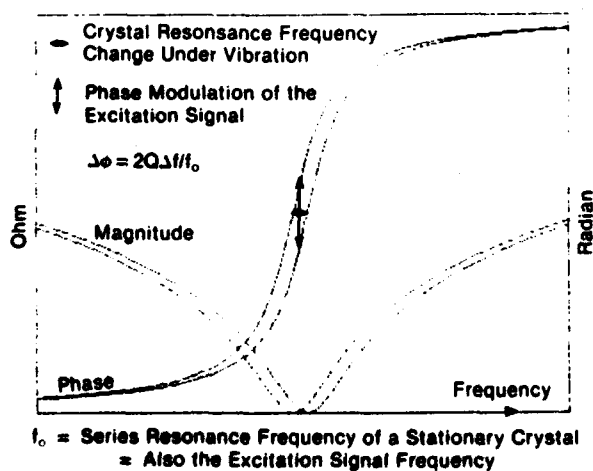
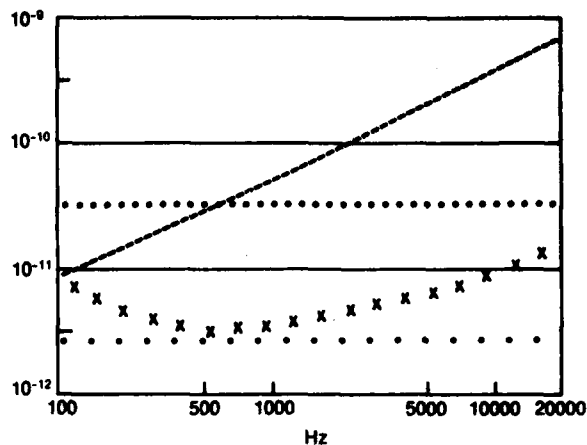


Figure 3. An impedance plot of the crystal unit used to depict phase modulation on the excitation signal



----- HP8662 Synthesizer
Phase Noise Limited

..... HP3582 Spectrum Analyzer
Input Noise Limited
When LNA Gain < 76 dB

x x x Low Noise Signal Generator
(Ref 8) Phase Noise Limited
When LNA Gain > 76 dB

• • LNA Input Noise Limited
When Gain > 76 dB and
If Signal Generator Is Ideal

Figure 4. Measurable g-sensitivity limited by various instrument noise levels ($Q = 40000$ and $A^2 = 5 \times 10^{-3} \text{ g}^2/\text{Hz}$ assumed in eqs. 1 and 2).

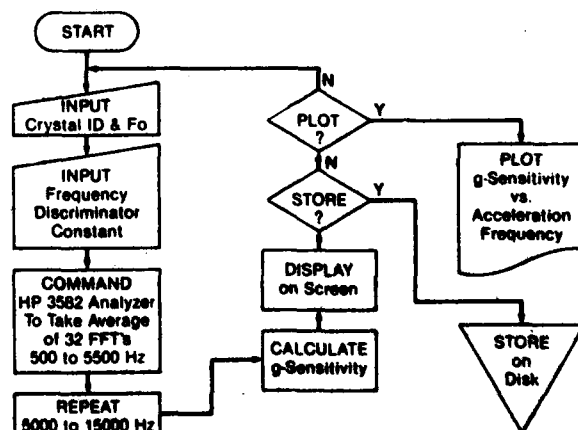


Figure 5. Measurement software flow chart - simplified.

From the preceding discussion, it can be seen that the net effect of vibration on the frequency discriminator is a phase modulation (see Figure 3) of the excitation signal according to $\Delta\phi = \Delta X/R$, where ΔX is the reactance change in the equivalent circuit of the crystal and R is the total equivalent series resistance. Near resonance, $\Delta\phi$ is related to Δf by (ignoring sign),

$$\Delta\phi = 2Q\Delta f/f_0 \quad (4)$$

For 100 MHz 5th overtone AT-cut crystals having Q_x of 100000, the passive Q in Eq. (4) is $Q = Q_x \cdot 40/(40 + 10 + 50) = 40000$, where 40 ohm is the crystal series resistance and the additional 60 ohm is an approximate resistance in series with the crystal unit as seen in Figure 2. Thus, for the case of 100 MHz 5th overtone AT-cut crystals, the phase shift given by Eq. (4) is

$$\Delta\phi = 80000 \Delta f/f_0 \text{ (radians)} \quad (5)$$

With signal levels at M1 ring modulator (phase detector) as shown in Figure 2, a phase detector sensitivity (K_p) of 0.1 volt/radian exists. The phase detector (M1) output voltage, V_p , is given by

$$\begin{aligned} V_p &= K_p \Delta\phi \\ &= 0.1 \Delta\phi \\ &= 8000 \Delta f/(100 \times 10^6) \\ &= 8 \text{ millivolts if } \Delta f = 100 \text{ Hz.} \end{aligned} \quad (6)$$

Finally,

$$\begin{aligned} V_B &= V_p G_{LNA} \\ &= (0.1 \times 80000 \times \Delta f/f_0) 550 \\ &= 0.044 \Delta f \\ &= K_B \Delta f \end{aligned} \quad (7)$$

and $K_B = 0.044$ volt/Hz is obtained. In Eq. (7), an LNA gain of 550 is used.

When the measuring apparatus is used, the phase detector output V_p is always measured for each crystal unit for Δf of ± 100 Hz in the signal generator frequency and K_B is found from Eq. (7). This is done to account for variation in crystal Q from unit to unit. The g -sensitivity in Eq. (2) can now be calculated from the measured RMS values of voltages V_A and V_B using the known proportionality constants K_A and K_B .

II.2 Limiting Factors of the Measurement Apparatus

In the present measurement system, the lower limit of measurable g -sensitivity is set by the spectrum analyzer (HP3582A) input noise and by the signal generator (HP8662A) phase noise. Figure 4 shows this limit together with a lower limit that can be achieved if a well designed oscillator is used as the signal source.

A simple Colpitts oscillator⁸ stabilized via phase shift property of the transconductance using an SC-cut crystal unit can be arranged for incremental tuning over a range of ± 15 ppm so that its frequency can be adjusted to resonance frequency of the crystal unit under test for g -sensitivity. Such an oscillator with a common base output buffer amplifier will provide noise levels 10 to 20 dB lower than the HP8662A synthesizer noise. As seen in Figure 4, it is possible to measure g -sensitivity less than 10^{-11} when testing 100 MHz crystal units using a low noise signal source.

After the phase detector, information of crystal resonance frequency fluctuation is in voltage fluctuation and therefore measurable g -sensitivity is limited by the spectrum analyzer input noise and by the LNA input noise. This g -sensitivity limit is given by Eq. (8).

$$\gamma = V_{NI}/(2Q K_p G_{LNA} A) \quad (8)$$

where Eq. (8) is derived from Eqs. (1), (4), (6) and (7). V_{NI} (equivalent noise input voltage) is the LNA input noise or the spectrum analyzer equivalent noise input referred to the phase detector output.

Before the phase detector, crystal resonance fluctuation information is in the signal phase fluctuation and therefore the signal generator phase noise is also a factor in determining the lower limit of measurable g -sensitivity. This limit is given by

$$\gamma(f_m) = 2f_m 10^{\mathcal{L}(f_m)/20}/(f_0 A) \quad (9)$$

where Eq. (9) is derived from Eq. (1) and the expression $\mathcal{L}(f_m) = 20 \times \log [\Delta f/(2f_m)]$ for the single sideband phase noise at offset frequency f_m . $\mathcal{L}(f)$ of the HP8662A synthesizer and of a VHF oscillator employing an SC-cut crystal unit were used to arrive at two of the limiting g -sensitivity curves shown in Figure 4. In practice, the actual HP8662A phase noise often exceeds the specified quoted resulting in improved g -sensitivity limits.

Bright lines exhibited by the HP8662A are at offset greater than 15 kHz so that measurement of g -sensitivity for acceleration frequencies of 500 Hz to 15000 Hz can be made without any degradation arising from generator bright lines.

Amplitude noise as well as phase noise can be important. The circuit symmetry of the balanced phase detector effectively suppresses the output fluctuation of the frequency discriminator arising from amplitude noise by typically 20 dB at 100 MHz. Signal generator additive noise occurs in output amplifiers and ECL dividers. Half of the noise power is AM and half is PM. Such additive noise levels are typically -150 dBc/Hz at 100 MHz, thus the AM noise is rarely a limitation.

Finally, it should be noted that the accelerometer itself has a non-ideal frequency response. At low frequency end, it acts as a high-pass filter determined by the total capacitance and resistance seen by the accelerometer output charge. At high frequency end, a useful frequency range of an accelerometer is limited by its mounted resonance (at 32 kHz in this example). The accelerometer frequency response is known, however, and therefore can be taken into account by a computer software. At 15000 Hz, for example, a correction factor of 1.4 is involved.

II.3 Apparatus Setup

The equipment used in the g -sensitivity measurement setup (Figure 1) consists of standard, commercially available units except an LNA (Low Noise Amplifier) and a complementary filter. The LNA has a nearly constant gain of 55 dB from 50 Hz to 500 kHz with an equivalent noise input of 1.5×10^{-9} volt/ $\sqrt{\text{Hz}}$. The complementary filter is a parallel network of a low-pass filter and a high-pass filter, of which the latter is terminated into a 50 ohm load so that the second harmonic (200 MHz) of the carrier does not overload the LNA input section. The line stretcher (delay line) provides the necessary 90 degree phase shift and also is used to compensate for phase detector (M1) offset.

The crystal unit together with its input and output network, an adjustable quarter-wave delay line, and the M1 ring modulator constitute a frequency discriminator whose function was described in the measurement analysis section (II.1).

Of the three costly items, the HP9826 computer can be replaced by a slower (and less convenient) HP85 and the HP8662A signal synthesizer by a simple crystal controlled Colpitts oscillator, as discussed earlier, in applications for which cost or low level g -sensitivity measurement is important.

III. Description of the Measurement Software

A BASIC program, of which a simplified flow chart is shown in Figure 5, written for the HP9826 scientific computer controls the measurement settings in the HP3582A spectrum analyzer and then reads the analyzer FFT data stored in memory. Then, the HP9826 computer calculates and plots the g-sensitivity vs. acceleration frequency characteristic of the crystal unit under test. The spectrum analyzer starts FFT analyses under computer command and accumulates the power spectra obtained in FFT repeated 32 times. The user is asked at the start of a measurement run to vary the excitation signal frequency by ± 100 Hz, enter the resulting frequency discriminator output voltage change, and correct a noise source to the vibration generator. The g-sensitivity measurement is otherwise automated taking a total of about 100 seconds per crystal unit.

The program also stores the g-sensitivity data on floppy disks for further analyses such as statistical treatment of mechanical peaks shown in Figure 11.

IV. Discussion of Results

(g-sensitivity vs. acceleration characteristics)

AT-cut crystal units with 3-point support - Figures 6 and 7 show typical computer plots of g-sensitivity for acceleration applied in the direction normal to the crystal plate of two sample units of 100 MHz 5th overtone AT-cut crystals using a 3-point support system.

Caution is required in determining that any of g-sensitivity peaks has not occurred as a result of transverse acceleration (associated with armature and fixture resonance of the vibration generator) experienced by the crystal but not sensed by the accelerometer. This is initially checked using three mutually perpendicular accelerometers mounted to a small cube of aluminum, one flat side of which is fastened to the vibration fixture with loctite. Crosstalk should be less than -15 dB at all frequencies when the vibration generator is driven at levels to be used in crystal unit testing. At most frequencies, crosstalk levels of -20 dB are observed using the VTS-100 vibration generator.

FC, IT, and SC - cut crystal units with 4-point support systems - A small quantity of FC, IT, and SC-cut crystal units have been tested for their g-sensitivity and sample g-sensitivity plots for each of three different types of crystals are shown in Figures 8, 9 and 10.

Firstly, it is to be observed that 4-point support systems in general excite low frequency (below 5000 Hz) resonances to a lesser extent than in the case of 3-point support systems of which examples are previously shown in Figures 6 and 7.

Secondly, as ϕ angle increases from AT-cut to SC-cut, g-sensitivity are seen to decrease as expected. Above 5000 Hz, this trend cannot be observed because of the present measurement system limitation, namely, the HP8662A Synthesizer phase noise.

It should be added that 3-point support systems are preferred for AT-cut crystal units to 4-point support systems because the latter have been found to cause an activity dip at an operating temperature range.

Statistics of mechanical resonance frequencies - The histogram in Figure 11 represents the number of resonance peaks vs. acceleration frequencies obtained for a sample of 450 crystal units.

While the g-sensitivity characteristic of a crystal unit can markedly be different from that of another unit (of similar electrical parameters), a general pattern of mechanical resonance locations emerges as shown in the histogram. The spread around a mechanical resonance frequency observed at three major peaks are caused by (1) manufacturing tolerances of crystal plate support systems and of their locations, by (2) variation in the unattached portion of packaged lead lengths between package and test fixture or by (3) effects on the crystal unit from transverse excitation owing to the vibration generator armature resonance.

Only the peaks occurring near 11500 Hz have been positively attributed to the crystal plate and the support system inside the package. A computer simulation shows that the resonances at 2500 Hz and at 7000 Hz involve the entire TO-5 package including the package leads transmitting energy to the crystal plate.

Narrow band measurements of mechanical resonance peaks show Q of 500 at 11500 Hz, but Q of only 50 at both 2500 Hz and 7000 Hz.

V. Conclusions

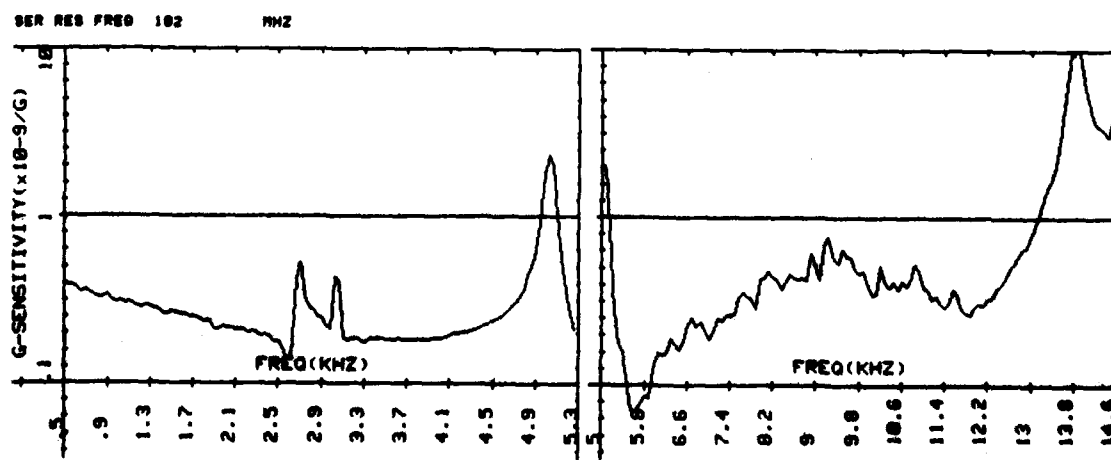
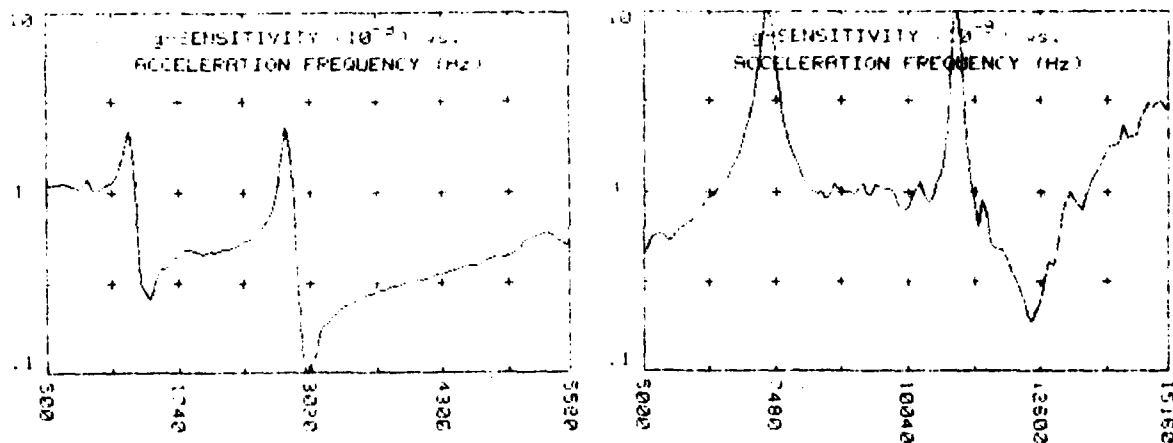
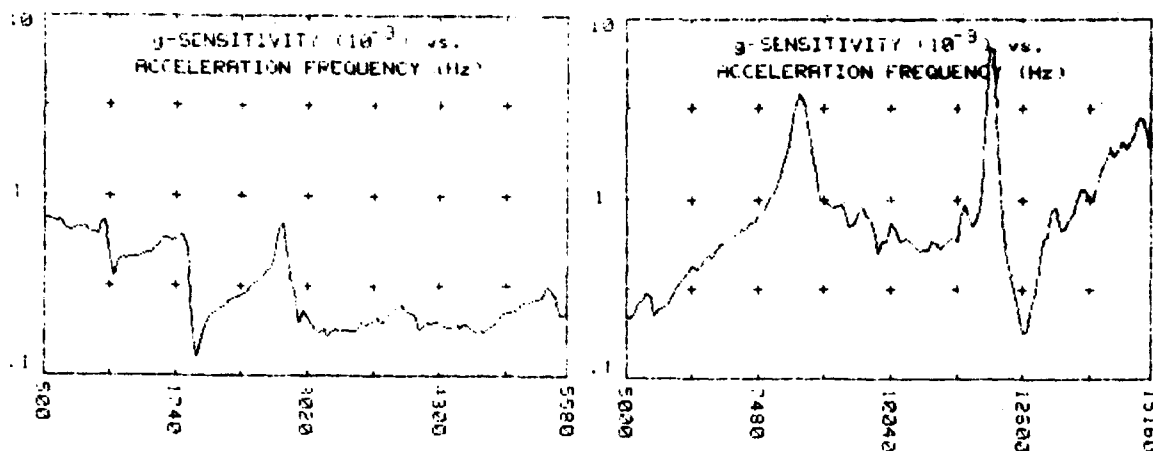
A computer-controlled g-sensitivity measuring apparatus has been assembled and used to screen 100 MHz AT-cut crystal units for any manufacturing defect. This apparatus obtains the g-sensitivity vs. acceleration characteristic of a crystal unit alone in less than 100 seconds in a given direction.

The basic equations governing computation of g-sensitivity from the measured data can be used to determine a new set of component requirements for the apparatus when measuring different crystal types to a specified g-sensitivity level. It has already been shown that measurement of g-sensitivity levels less than 10^{-11} is possible for 100 MHz quartz crystal units.

The described g-sensitivity measurement technique can be used to study mechanical resonance properties of crystal mounting and packaging schemes, simply screen defective crystal units, or investigate interfering modes of low frequency mechanical resonances of the crystal plate and supporting systems after a simple modification of the measurement software.

References

1. Warner, A.W. and Smith, W.L., "Highly Stable Crystal Oscillators for Missile Applications," Proceedings of the 13th Annual Symposium on Frequency Control, May 1959.
2. Filler, R.L., "The Effect of Vibration on Frequency Standards and Clocks," Proc. 35th ASFC, May 1981.
3. Warner, A.W., Goldfrank, B., Meirs, M. and Rosenfeld, M., "Low g Sensitivity Crystal Units and Their Testing," Proc. 33rd ASFC, May 1979.
4. Ward, R.W., "Design of High Performance SC Resonators," Proc. 35th ASFC, May 1981.
5. Bottom, V.E., *Introduction to Quartz Crystal Unit Design*, Van Nostrand Reinhold, 1982.
6. Lukaszek, T.J. and Ballato, A., "Resonators for Severe Environments," Proc. 33rd ASFC, May 1979.
7. Janiaud, D., Nissim, L. and Gagnepain, J.-J., "Analytical Calculation of Initial Stress Effects on Anisotropic Crystals: Application to Quartz Resonators," Proc. 35th ASFC, May 1981.
8. Healey III, D.J., and Kwan, S., "SC-Cut Quartz Crystal Units in Low-Noise Oscillator Application at VHF," Proc. 35th ASFC, May 1981.



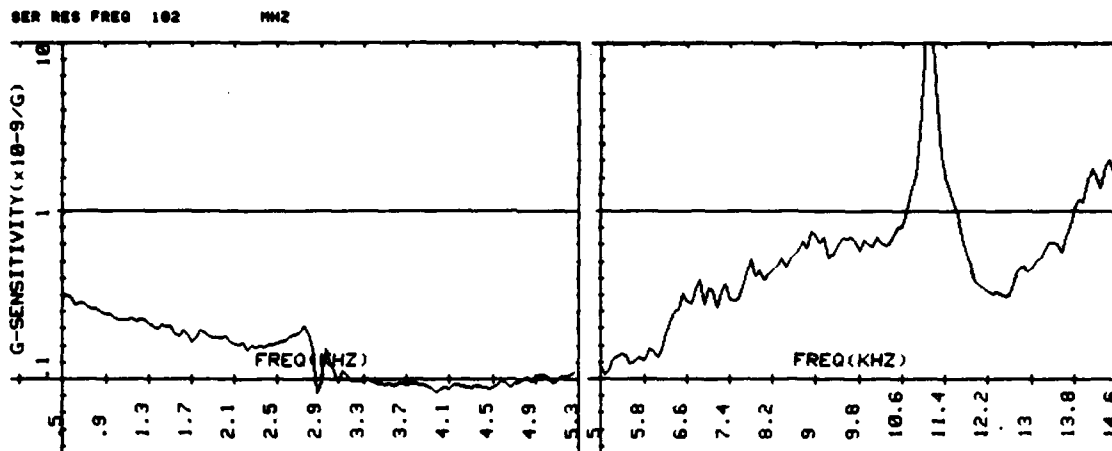


Figure 9. g-sensitivity vs. acceleration frequency 5th overtone 100 MHz IT-cut with 4-point support.

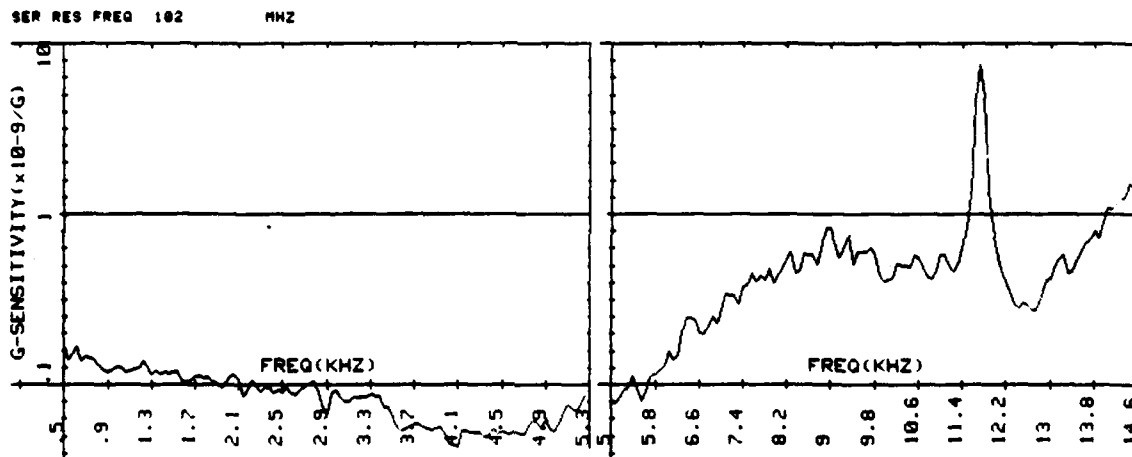


Figure 10. g-sensitivity vs. acceleration frequency 5th overtone 100 MHz SC-cut with 4-point support

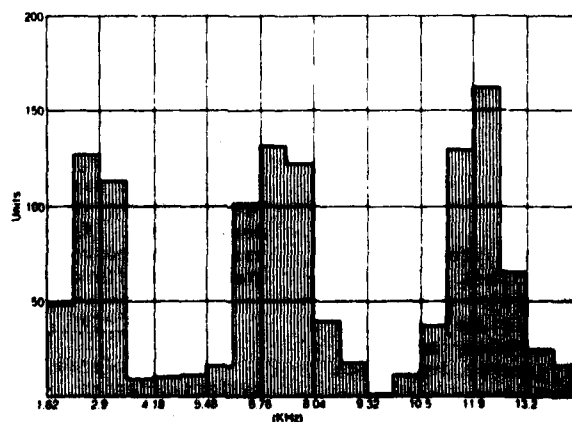


Figure 11. Number of mechanical resonance peak occurrences vs. acceleration frequency.

EVALUATION OF CRYSTAL MEASUREMENT SYSTEMS*

R.C. Smythe & W.H. Horton

PIEZO TECHNOLOGY, INC.
ORLANDO, FLORIDA 32804

SUMMARY

A number of commercially available instruments have been evaluated for use in precision quartz crystal resonator measurement. Since no single system is best for all the commonly required measurements, emphasis has been placed on basic performance factors so that users can judge the suitability of each system for a particular type of measurement. Comparative room temperature measurements of a range of crystal resonators, resistors, and capacitors are reported.

1. Introduction

This paper presents the preliminary results of a comparative evaluation of commercially available test instruments suitable for application to the measurement of precision quartz crystal resonators. More detailed results will be given later [1].

Four instruments were evaluated:

H-P 4191A RF Impedance Analyzer

H-P 4192A LF Impedance Analyzer

Polarad ZPV Vector Analyzer

Transat MCT Modular Crystal Test System

Of these, the Transat MCT is the only instrument specifically designed for crystal resonator measurement and the only instrument of the four which is capable of automatic resonator measurement without additional system components. Our evaluation of the MCT was limited to its use as a stand-alone instrument. The remaining instruments were incorporated into systems which included a separate controller and, for the H-P 4191A and Polarad ZPV, a separate frequency synthesizer. Software was provided by PTI and employed algorithms described earlier [2].

2. System Descriptions

2.1 H-P 4191A

The 4191A RF Impedance Analyzer is a self-contained, programmable instrument for the frequency range from 1 to 1000 MHz [3]. A built-in coaxial reflectometer (directional bridge) measures the reflection coefficient of the unknown impedance, which is converted into any of several immittance forms by an internal microprocessor.

*This work was supported by the U.S. Army Electronics Research & Development Command, Ft. Monmouth, N.J., Contract No. DAAK20-82-C-0390.

For precision crystal measurement the accuracy and resolution of the internal frequency synthesizer are inadequate, and an external synthesizer is required. Figure 1 shows the block diagram of the complete system, which includes a controller and peripherals.

The system has been described earlier [2]. For each resonator, C_0 is measured at frequencies above and below resonance. Resonance, f_r , or motional arm resonance, f_s , is found by a simple, two-stage search. The corresponding resistance, R_r or R_1 , is measured. Optionally, L_1 is measured by the $\pm f$ method. From the equivalent circuit parameters the load frequency and resistance, f_L and R_L , corresponding to an arbitrary positive or negative load capacitance may be calculated. In addition, f_L and R_L may be measured using the reactance offset method [2, 4]. Unwanted modes are measured in the same manner as the principal mode.

2.2 H-P 4192A

The 4192A Low-Frequency Impedance Analyzer is a self-contained, programmable instrument covering 5 Hz to 13 MHz. It is a three-terminal instrument; however, one of the floating terminals may be grounded if desired. The instrument measures the voltage across, and the current through, the unknown impedance, using a null technique [5], and presents the results in any of a number of immittance forms. A variety of fixtures are available.

The internal frequency reference can be locked to an external standard. Frequency resolution is 1 Hz at frequencies above 1 MHz, adequate for many purposes. An external synthesizer may be used for greater resolution.

Manual measurements of resonator parameters are possible with the 4192A, but more conveniently a separate controller is used. We have employed the same algorithms as for the 4191A.

In addition to impedance measurement, the 4192A can also function as a two-channel transmission test system measuring attenuation and phase shift. In this mode, it may be used with a separate pi-network, reflectometer, or other test network for resonator measurement. Measurements were made in this mode using a pi-network for the purpose of providing a comparison with other pi-network systems.

2.3 Polarad ZPV

The Polarad ZPV Vector Analyzer is a systems-oriented vector voltmeter. The internal microprocessor incorporates a number of features useful for impedance measurement. Using tuner ZPV-E3 the frequency range is 0.3 to 2000 MHz.

Figure 2 shows the block diagram of the complete system. Both PI-network and reflectometer test networks were used. The software is adapted from the H-P 4191A system. For the reflectometer measurements, it was necessary to add a calibration routine to the program.

2.4 Transat MCT

The Transat Modular Crystal Test System is a self-contained system for the automatic measurement of crystal resonators from 1 to 250 MHz using a separate pi-network. It may also be used with an external controller and also an external frequency source, but these modes of operation were not evaluated. Additionally, it has a phase-locked mode of operation.

The system comprises four modules plus the separate pi-network. They are 1) the voltage-controlled oscillator (VCO) module which provides outputs for the measurement channel and a reference channel; 2) the phase-locked receiver (PLR) module, a two-channel phase/amplitude detector; 3) the automatic lock module (ALM) which controls frequency sweep and phase lock functions; and 4) the automatic processor module (APM) which contains a pre-programmed microprocessor which controls the various test functions. Frequency is determined by a built-in counter with a resolution of .01 ppm per 1 second gate time. Provision is made for a TTL-level external 1 MHz frequency standard.

The full set of equivalent circuit parameters are computed from three admittance measurements (but additional measurements may have to be performed in obtaining these three.) The detailed measurement sequence is described in the operating manual [6]. The three measurements are: 1) a measurement of C_0 at af_1 where f_1 is a frequency near f_s and $a = 0.9$ for $f_1 > 32$ MHz. For $f_1 < 32$ MHz, a is an even integer such that $32 < af_1 < 64$ MHz; 2) a measurement at f_s ; 3) a measurement at a frequency f_1 near f_s such that the difference in vector voltmeter phase angle for the two measurements is sufficiently large. A spurious search mode is provided which measures frequency and resistance of unwanted modes.

3. Measurement Results

Comparative measurements were made of the equivalent circuit parameters of 13 crystal units at frequencies from 1.65 MHz to 173 MHz. For six of these units load frequency and resistance were measured. Unwanted mode measurements were made for some resonators. Measurements were also performed on a set of ten chip resistors and capacitors provided by E.T. & D. Lab, Ft. Monmouth.

In addition, numerous informal measurements were made with all four systems. The 4191A- and 4192A-based systems have been in daily use at PTI for over two years. As a result, our experience with these systems is much greater than with the ZPV and Transat systems. Moreover, our evaluation of the Transat system was performed using instruments provided at no cost by the manufacturer. During the nine-month contract period a series of modifications were introduced. Measurements made with early instruments are referred to in the tables as "Transat 1". Approximately two weeks before the end of the contract period an improved instrument was received. Measurements made with this system are referred to as

"Transat 2".

3.1 Chip Element Measurement

Typical comparative measurements of the chip resistors and capacitors are presented in figure 3 and Tables 1 through 8.

Figure 3 shows the result of amplitude linearity measurements on a 10 ohm chip resistor using the H-P 4192A, the ZPV with 12.5 ohm pi-network, and the Transat 2 with 12.5 ohm pi-network. The 4192 measurements were made at 11 MHz, the others at 30 MHz. These data represent averages of 7 to 9 measurements. From the figure it can be seen that the current range and linearity of the 4192 are clearly superior to the other two instruments. Also, the linearity and current range of the Transat 2 are very much better than those of the Transat 1. There are significant differences among the instruments in the standard deviation of the measurements and in the measured values of inductance for the chip resistor and resistance for the chip capacitor, and these will be reported later [1].

In Tables 1 through 8 a sample of the results of chip element measurements at 11 and 201 MHz are summarized. The unbracketed numbers are the mean of 7 to 9 measurements while those in parenthesis represent the standard deviations. All pi-network measurements used the same 12.5 ohm network.

Measurement of the series resistance of the capacitors presented some difficulty for the ZPV and the Transat 2. The measurements of chip resistors show, in general, good agreement among the several systems. The greatest differences are seen at high resistances. At 11 MHz the Transat 2 was unable to make usable measurements of a 10K ohm resistor, while the ZPV/12.5 ohm pi system gave a very low value. At 201 MHz both of these systems gave unreasonably low resistance values for the 10K ohm resistor [1].

3.2 Resonator Measurements

Comparative measurements were made on 13 AT-cut and SC-cut resonators ranging in frequency from 1.65 MHz to 173 MHz. Data for 5 of these are summarized in Tables 9-13, which, as usual, represent the mean of 7 to 9 measurements.

For each resonator, the equivalent circuit parameters were measured and the current at resonance measured or calculated. To minimize measurement differences attributable to current, all units were screened for linearity. For six units load frequency and resistance were measured with simulated load capacitance of 32 pF. Load measurements were not made with the Transat MCT since its firmware does not include a simulated load capacitance capability. (However, it does allow phase offset measurements.)

Examination of the tables shows good agreement for most measurements of equivalent circuit parameters. Small differences in resonance frequency value are attributable to temperature differences. VWF resonance frequencies obtained with the ZPV/PI appear to be in error by a few ppm.

The measurements reported as C_0 have different meaning for different measurement networks, giving rise to apparent measurement differences of the order of 0.5 pF. In our measurements the crystal can was

always floating, an ambiguous condition. Moreover, the three types of measurement network used differ as to terminal conditions. The reflection bridge makes a two-terminal grounded measurement; i.e., one measurement terminal is grounded. The pi-network makes (approximately [1]) a two-terminal floating measurement. Finally, the H-P 4192A makes a true three-terminal measurement.

Simulated load measurements presented greater difficulties, especially in the measurement of R_L . For example, the ZPV/Wiltron bridge measurements of R_L in every instance but one were essentially meaningless.

In addition, the Transat systems had difficulty with some very high Q resonators. For a 5 MHz 3rd overtone AT-cut resonator with a Q of 2.2×10^6 and resistance of 18 ohms, both the Transat 1 and 2 were unable to achieve phase lock, due to a too rapid frequency sweep, and measurements were not possible. (The manufacturer plans a modification to correct this problem.)

Figure 4 presents linearity measurements of a 10.2 MHz fundamental mode, AT-cut resonator. This resonator has increasing resistance at low drive. Measurements were made with the Transat 2, the ZPV with reflection bridge, and the H-P 4192A. No linearity measurements were made with the H-P 4191A since its test level is essentially fixed.

For several resonators unwanted mode measurements were made. Figure 5 shows the frequency spectrum of a 4.192 MHz fundamental mode, AT-cut crystal obtained using a 12.5/12.5 ohm pi-network in a transmission measurement system. The mode resistances shown were obtained with the 4191A system. Values obtained with the 4192A, the ZPV/bridge, and the ZPV/pi systems were within 10 ohms of these. The Transat 2 system was able to locate only the first of the two spurs and indicated a mode resistance approximately 30 ohms lower than that shown.

4. System Considerations

Our evaluation of the four systems included a number of factors, some of which are listed here.

1. "Turn-Key" status. The Transat MCT, as delivered, is capable of automatic resonator measurements. The H-P 4192A is capable of manual resonator measurements, but requires an external controller and software for automatic measurements. The H-P 4191A requires these and a programmable frequency source, while the ZPV requires all of the foregoing plus a measurement network.

2. Frequency range. The H-P 4192A is limited to 13 MHz maximum.

3. Drive level. Measurements over a range of drive level are possible with all instruments except the 4191A.

4. Remote measurement. Remote measurement is possible with all systems. The 4191A, 4192A, and ZPV/bridge systems permit coaxial line extension of the measurement port, so that the measurement network can remain with the instrument. Pi-network systems such as the Transat MCT require that the pi-network be located remotely. In addition the Transat system

requires phase equalization of the measurement and reference channels.

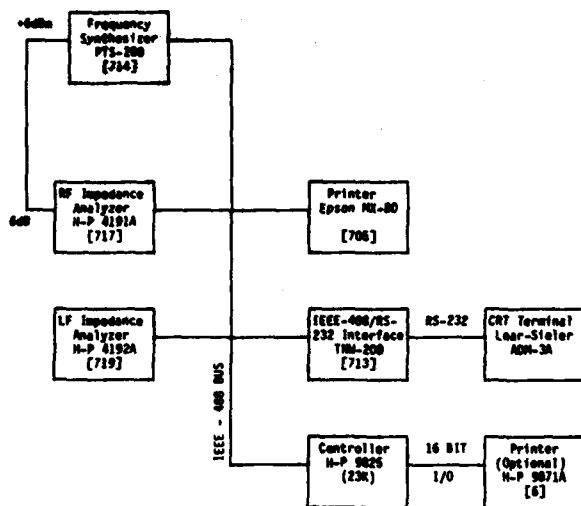
5. Frequency sweep/display. Only the ZPV (with suitable additional system components) is capable of rapid frequency sweep and display. The MCT, however, can rapidly search a given frequency band for a resonance and has a spurious search mode. Via a controller, the 4191A and 4192A systems can be given a display and plot capability, but the data rate of these instruments is inconveniently low for this purpose.

5. Conclusions

In manufacturing and characterizing crystal resonators a variety of measurements are required for different purposes. No single system can best perform all these measurements. This paper has attempted to set forth the salient features of four commercially available instruments as they apply to resonator measurement.

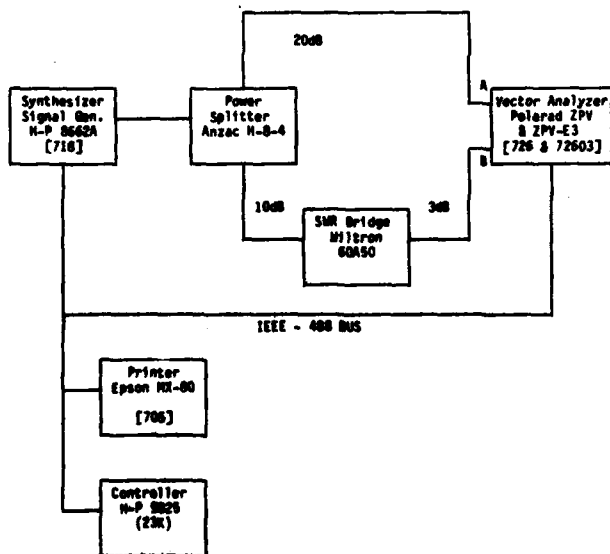
REFERENCES

1. R.C. Smythe & W.H. Horton, Final Report, Contract No. DAAK20-82-C-0390; in preparation.
2. R.C. Smythe, "An Automated Resonator Measurement System Using a Reflection Coefficient Bridge," Proc. 35th AFCS, pp. 280-285; 1981.
3. T. Ichino, H. Ohkawara, & N. Sugihara, "Vector Impedance Analysis to 1000 MHz," H-P Journal, vol. 31, no. 1, pp. 22-32; Jan. 1980.
4. W.H. Horton, et al, "Comparison of Methods for Measurement of Quartz Crystal Resonators with Load Capacitance," Proc. 35th AFCS, pp. 271-279; 1981.
5. Y. Narimatsu, K. Yagi, & T. Shimizu, "A Versatile Low-Frequency Impedance Analyzer with an Integral Tracking Gain-Phase Meter," H-P Journal, vol. 32, no. 2, pp. 22-28; Sept. 1981.



Notes: 1) Numbers in brackets are instrument addresses

FIG. 1 BLOCK DIAGRAM, RESONATOR MEASUREMENT SYSTEM USING H-P 4191A AND/OR H-P 4192A



Notes: 1) Numbers in brackets are instrument addresses

FIG. 2 BLOCK DIAGRAM RESONATOR MEASUREMENT SYSTEM USING POLARAD ZPV

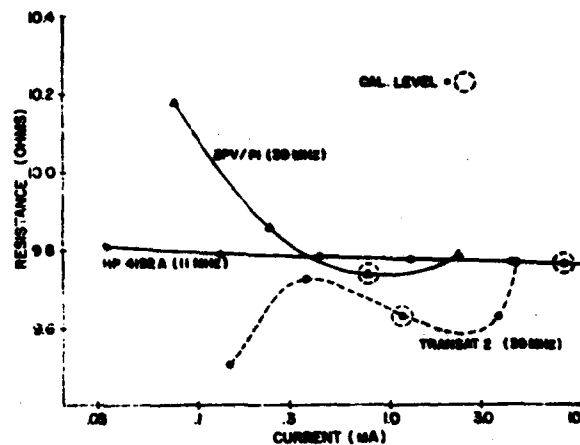


Figure 3 Resonance Resistance vs Current for a 50 Ohm Chip Resistor

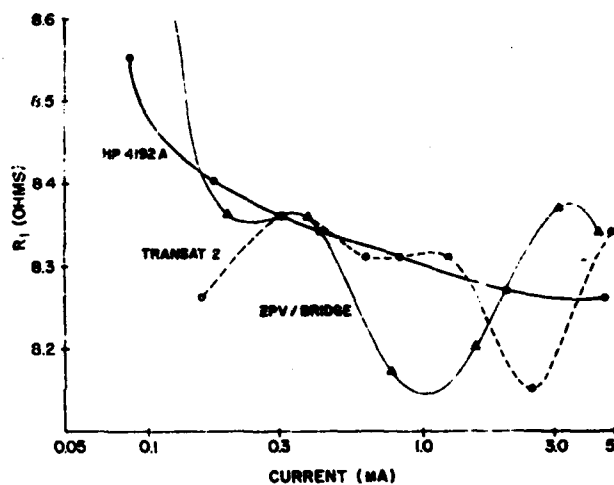


Figure 4 Resonance Resistance vs Current for a 30.2 MHz Fundamental Model AT-Cut Resonator

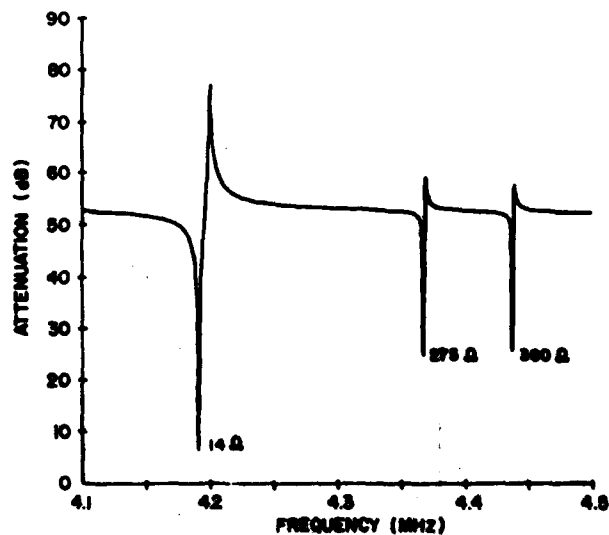


Figure 5 Standing Wave Spectrum of a 4.2 MHz Fundamental Mode AT-Cut Resonator. Indicated values of natural resistance measured with N-P 4192A.

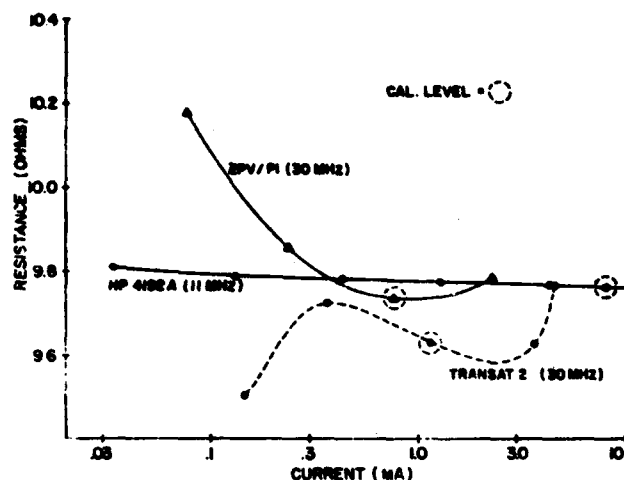


Table 1 Comparative Chip Capacitor Measurements,
5 pF Chip Capacitor
(Frequency: 11 MHz)

Measurement System	R_s (Ohms)	C_s (pF)
4191A	-3.7 (4.4)	4.88 (.006)
ZPV/Bridge	113 (11)	5.24 (.03)
ZPV/Pi Net.	367 37	5.25 (.08)
Transat 2/Pi Net.	-120.9 (.07)	4.79 (.00)
4192A	1.8 (.8)	5.02 (.001)
4192A/Pi Net.	-9.3 (16)	5.42 (.03)

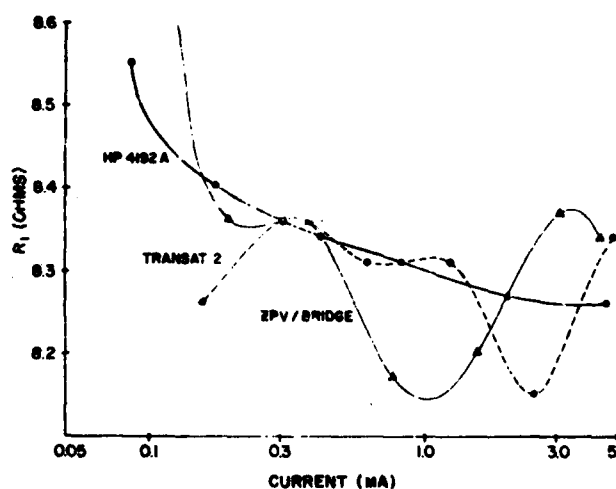


Table 2 Comparative Chip Capacitor Measurements,
5 pF Chip Capacitor
(Frequency: 201 MHz)

Measurement System	R_s (Ohms)	C_s (pF)
4191A	0.21 (.006)	5.142 (.0003)
ZPV/Bridge	-1.77 (.08)	5.245 (.004)
ZPV/Pi Net.	5.11 (.08)	5.343 (.005)
Transat 2/Pi Net.	5.30 (.74)	5.19 (.026)

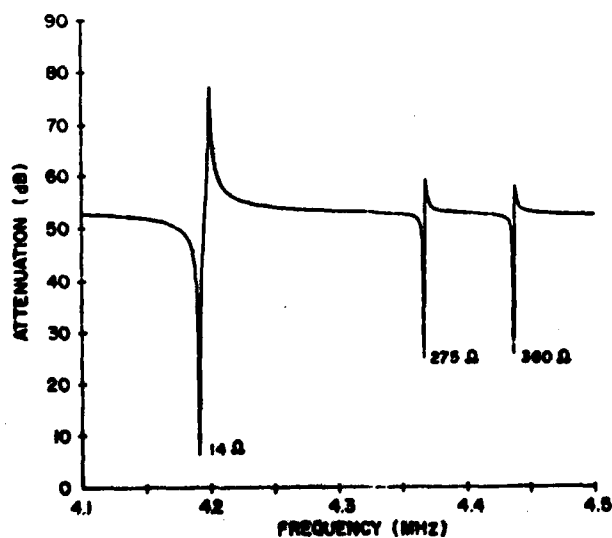


Table 3 Comparative Chip Capacitor Measurements,
39 pF Chip Capacitor
(Frequency: 11 MHz)

Measurement System	R_s (Ohms)	C_s (pF)
4191A	0.12 (.13)	38.87 (.02)
ZPV/Bridge	9.8 (.9)	39.77 (.09)
ZPV/Pi Net.	16.2 (.7)	39.06 (.09)
Transat 2/Pi Net.	-12.18 (.004)	38.51 (.01)
4192A	0.00 (.00)	38.40 (.00)
4192A/Pi Net.	0.9 (.35)	39.20 (.03)

Table 4 Comparative Chip Capacitor Measurements,
39 pF Chip Capacitor
(Frequency: 201 MHz)

Measurement System	R_s (Ohms)	C_s (pF)
4191A	0.07 (.003)	40.13 (.002)
ZPV/Bridge	-0.04 (.009)	43.70 (.04)
ZPV/Pi Net.	0.77 (.02)	39.65 (.05)
Transat 2/Pi Net.	0.90 (.28)	41.1 (1.2)

Table 5 Comparative Chip Resistor Measurements,
20 Ohm Chip Resistor
(Frequency: 11 MHz)

Measurement System	R_s (Ohms)	X_s (Ohms)
4191A	19.52 (.004)	0.07 (.00)
ZPV/Bridge	19.66 (.005)	0.20 (.00)
ZPV/Pi Net.	19.72 (.00)	0.21 (.04)
Transat 2/Pi Net.	19.65 (.00)	-0.04 (.00)
4192A	19.91 (.02)	0.07 (.02)
4192A/Pi Net.	19.67 (.007)	0.07 (.00)

Table 6 Comparative Chip Resistor Measurements,
20 Ohm Chip Resistor
(Frequency: 201 MHz)

Measurement System	R_s (Ohms)	X_s (Ohms)
4191A	19.66 (.003)	0.49 (.00)
ZPV/Bridge	19.57 (.01)	2.85 (.01)
ZPV/Pi Net.	19.54 (.02)	1.44 (.04)
Transat 2/Pi Net.	19.34 (.02)	1.49 (.04)

Table 7 Comparative Chip Resistor Measurements,
400 Ohm Chip Resistor
(Frequency: 11 MHz)

Measurement System	R_s (Ohms)	X_s (Ohms)
4191A	418.9 (.2)	-0.32 (.11)
ZPV/Bridge	416.2 (.3)	-4.6 (.2)
ZPV/Pi Net.	420.9 (1.0)	3.2 (.95)
Transat 2/Pi Net.	413.9 (.08)	-17.4 (.5)
4192A	425.5 (.0)	-2.50 (.00)
4192A/Pi Net.	420.9 (.5)	-6.3 (.6)

Table 8 Comparative Chip Resistor Measurements,
400 Ohm Chip Resistor
(Frequency: 201 MHz)

Measurement System	R_s (Ohms)	X_s (Ohms)
4191A	411.0 (.07)	-61.2 (.007)
ZPV/Bridge	406.4 (.7)	-64.1 (1.0)
ZPV/Pi Net.	399.0 (1.3)	-65.0 (.7)
Transat 2/Pi Net.	395.3 (.05)	-75.1 (.4)

Table 9 Comparative Measurements, Resonator 15
(SC-Cut, 3rd Fundamental Q=820K, N=840)

Meas. System	4191	4192	ZPV/ Bridge	ZPV/ 12.5 Ohm Pi	Transat 2/ 12.5 Ohm Pi
f_s (kHz)	4011.428	4011.420	4011.425	4011.429	4011.414 (.0004)
R_1 (Ohms)	9.99 (.00)	9.92 (.004)	9.97 (.011)	9.95 (.00)	10.03 (.011)
L_1 (nH.)	260.97 (.03)	270.36 (.30)	260.47 (.11)	260.90 (.25)	277.1 (7.5)
C_0 (pF)	4.13 (.000)	3.00 (.00)	4.47 (.15)	2.99 (.00)	3.04 (.000)
f_L (kHz)	4011.098	4011.098	4011.092	4011.708	-
$(f_L - f_s)$ (Hz)	260 (.4)	270 (0)	267 (1)	274 (.5)	-
R_L (Ohms)	12.2 (.4)	11.72 (.06)	1.2 (.4)	47.4 (4.4)	-
Current at f_s (mA.)	0.75	0.73	0.75	0.79	1.16

Table 10 Comparative Measurements, Resonator M-1
(SC-Cut, 3rd Overtone Q=750K, M=35)

Meas. System	4191	4192	4192/ 12.5 Ohm Pi	ZPV/ Bridge	ZPV/ 12.5 Ohm Pi	Transat 2/ 12.5 Ohm Pi
f_s (kHz)	9905.365	9905.347	9905.327	9905.360	9905.353	9905.239
R_1 (Ohms)	95.13 (.13)	95.73 (.06)	95.95 (.09)	94.72 (.01)	94.88 (.16)	94.7 (1.3)
L_1 (mH.)	1150.6 (.5)	1157.1 (.1)	1165.0 (7.6)	1131.8 (7.8)	1142.1 (8.4)	1166 (17)
C_0 (pF)	4.83 (.01)	4.33 (.00)	4.75 (.06)	5.02 (.03)	5.00 (.09)	4.23 (.002)
f_L (kHz)	9905.306	9905.377	9905.367	9905.390	9905.383	-
$(f_L - f_s)$ (Hz)	31.0 (.8)	30.7 (.5)	30.7 (.5)	31.0 (.4)	30.0 (.0)	-
R_L (Ohms)	130.5 (.8)	123.8 (1.0)	143.2 (1.2)	116.3 (.9)	130.2 (2.0)	-
Current at f_s (mA.)	.31	.30	.30	.31	.63	1.16

Table 13 Comparative Measurements, Resonator 10
(AT-Cut, 7th Overtone Q=55K, M=4)

Meas. System	4191	ZPV/ Bridge	ZPV/ 12.5 Ohm Pi	Transat 1/ 12.5 Ohm Pi
f_s (kHz)	173250.899 (.002)	173250.994 (.007)	173251.487 (.004)	173250.901 (.009)
R_1 (Ohms)	41.56 (.01)	41.25 (.02)	44.36 (.04)	41.58 (.03)
L_1 (mH.)	2.18 (.00)	2.17 (.005)	2.22 (.005)	2.13 (.00)
C_0 (pF)	5.44 (.00)	5.60 (.005)	5.46 (.008)	5.24 (.01)
Current at f_s (mA.)	0.49	0.49	0.39	1.39

Table 11 Comparative Measurements, Resonator 114
(AT-Cut, 3rd Overtone Q=71K, M=21)

Meas. System	4191	ZPV/ Bridge	ZPV/ 12.5 Ohm Pi	Transat 1/ 12.5 Ohm Pi
f_s (kHz)	83999.027 (.004)	83999.033 (.001)	83999.293 (.003)	83999.108 (.003)
R_1 (Ohms)	14.42 (.01)	14.49 (.01)	14.72 (.02)	14.30 (.04)
L_1 (mH.)	1.93 (.00)	1.95 (.005)	1.94 (.01)	1.908 (.005)
C_0 (pF)	6.30 (.005)	6.44 (.01)	6.71 (.01)	6.082 (.006)
Current at f_s (mA.)	0.70	0.70	1.68	2.28

Table 12 Comparative Measurements, Resonator A
(AT-Cut, Fundamental Q=11K, M=15)

Meas. System	4191	ZPV/ Bridge	ZPV/ 12.5 Ohm Pi	Transat 2/ 12.5 Ohm Pi
f_s (kHz)	151392.853 (.001)	151392.830 (.009)	151394.570 (.007)	151393.315 (.011)
R_1 (Ohms)	46.59 (.006)	46.58 (.023)	47.45 (.08)	45.56 (.02)
L_1 (mH.)	0.519 (.001)	0.519 (.001)	0.510 (.01)	0.515 (.0006)
C_0 (pF)	1.47 (.004)	1.50 (.007)	-	0.98 (.001)
Current at f_s (mA.)	0.45	0.45	0.46	0.56

A NEW FREQUENCY FOR PIEZOELECTRIC RESONATOR MEASUREMENT

M.H. Horton and R.C. Smythe

PIEZO TECHNOLOGY, INC.
ORLANDO, FLORIDA 32804

AD P 002491

When a quartz crystal resonator unit is used in a reflectance bridge to stabilize the frequency of an oscillator it is found that the frequency is insensitive to the length of the cable between the crystal unit and the bridge (1,2,3,4). Properties of the reflection coefficient of a crystal unit, one of which explains this insensitivity, and some related crystal unit measurements will be described.

When the crystal unit can be represented in the vicinity of an isolated mode of vibration by a linear electrical impedance Z_e having the usual simplified equivalent circuit of fig. 1, the corresponding

crystal unit reflection coefficient is $\Gamma_e = \frac{Z_e - R_0}{Z_e + R_0}$.

Just as the Z- or Y- plane immittance circle diagram of the crystal unit is useful in describing the relationships between the resonance frequency f_r , the motional arm resonance frequency f_s and other characteristic frequencies of Z_e , the corresponding plots in the Γ or Smith-chart plane are useful in analyzing the characteristics of Γ_e . Steps in making such a plot are shown in fig. 2, fig. 3 and fig. 4.

The characteristic frequencies f_s and f_r can be located on the Γ plane plot as in the Z- or Y- plane plot. In addition the condition of minimum $|\Gamma_e|$ identifies another frequency, f_g , which is not obvious on the Z- or Y- plane plots. Since the output of a reflectance bridge is proportional to Γ , the minimum value of $|\Gamma_e|$ corresponds to a minimum of transmission in the reflectance bridge. It can be shown this is also essentially the reference condition used in oscillator stabilization. It is therefore useful to examine properties of f_g .

If the crystal unit is connected to a reflectance bridge of impedance R_0 by a cable of the same impedance the effect on the Smith-chart is to rotate the crystal immittance characteristic about the chart origin by an angle of twice the angular length of the cable as shown in fig. 5. Since this rotation does not change the location of the minimum value of $|\Gamma_e|$ on the immittance circle, that characteristic of f_g is unchanged. The positions of f_r and f_s on the rotated circle have no correspondingly simple method of identification. Moreover, from fig. 4 it can be seen that, unlike f_r , f_g continues to exist when the crystal unit figure of merit M is less than 2. Other properties of f_g evident from graphical analysis are shown in fig. 6.

From other analysis (5) using the work of Hafner (6) it can readily be shown that $f_g < f_r$ if and only if

$$m + m^{-1} > M \quad (1)$$

or, equivalently,

$$R_1 > \frac{R_0}{1+m^2} \quad (2)$$

where m is defined in fig. 7.

In addition the expressions in fig. 7 for the fractional frequency difference between f_g and f_r can be obtained. The well known corresponding relations for f_r are also shown in fig. 7 for comparison. By graphical analysis or from equation (3) of fig. 7 it can be seen that f_g is independent of R_1 .

Because of insensitivity to cable length, use of f_g for remote measurement is of particular interest. With instrumentation capable of measuring or calculating Γ_e the procedure is straightforward. Typical results from measurement of an ovenized crystal unit using an HP 4191A Impedance Analyzer (7) are shown in table I. The crystal unit equivalent circuit element values are given in table II.

REFERENCES

- (1) F.L. Walls and S.R. Stein, "A frequency-lock system for improved quartz crystal oscillator performance," IEEE Trans, Instrum. Meas., vol. IM-27, no. 3, pp. 249-252, Sept. 1978.
- (2) S.R. Stein and J.P. Turneaure, "The development of the superconducting cavity stabilized oscillator," Proc. 27th Annual Symp. on Frequency Control, pp. 414-420, 1972.
- (3) S.R. Stein, "The superconducting-cavity stabilized oscillator and an experiment to detect time variation of the fundamental constants," Ph.D. dissertation, Stanford University, Stanford, California, 1974.
- (4) C.S. Stone and O.J. Baltzer, "A frequency domain reflectometer for quartz resonator investigations," Proc. 36th Annual Symp. on Frequency Control, pp. 321-326, 1982.
- (5) M.H. Horton and R.C. Smythe, "A new frequency for piezoelectric measurement," Proc. IEEE, vol. 71, no. 2, pp. 280-282, Feb. 1983.
- (6) E. Hafner, "The piezoelectric crystal-Definitions and methods of measurement," Proc. IEEE, vol. 57, no. 2, pp. 179-201, Feb. 1969.
- (7) R.C. Smythe, "An automated resonator measurement system using a reflection coefficient bridge," Proc. 35th Annual Symp. on Frequency Control, pp. 280-286, 1981.

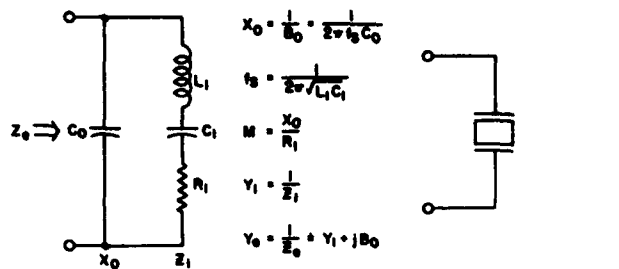


FIGURE 1 Crystal Unit Equivalent Circuit

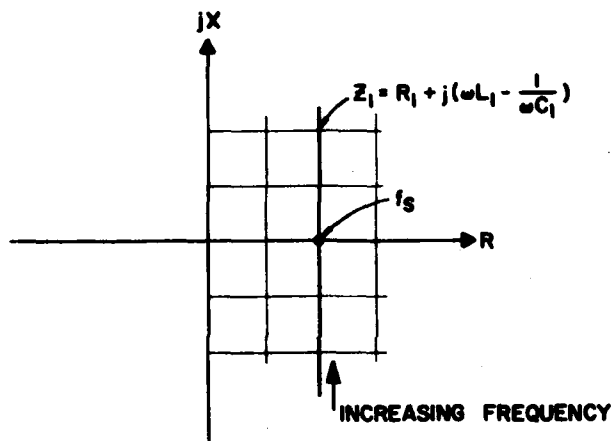


FIGURE 2 Z-Plane plot of Z_1

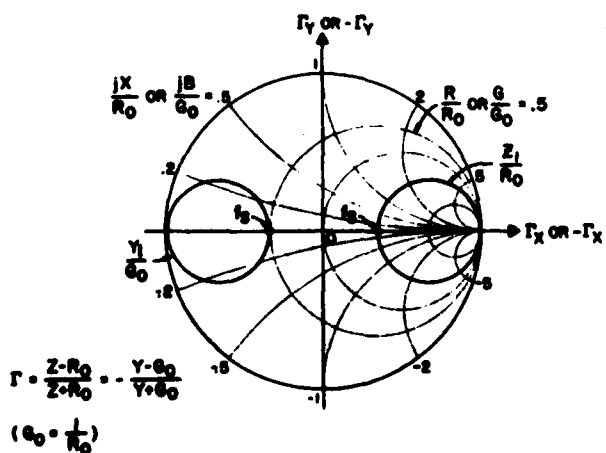


FIGURE 3 Γ -Plane Plot of Z_1/R_0 and Y_1/G_0

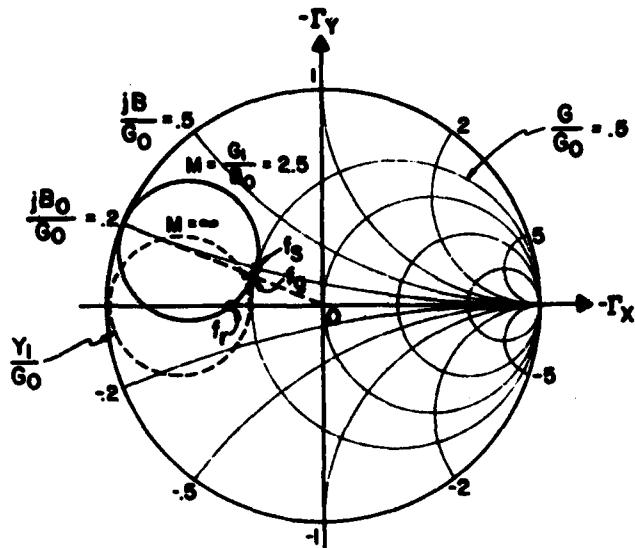


FIGURE 4 Γ -Plane Plot of $Y_0/G_0 = Y_1/G_0 + jB_0/G_0$

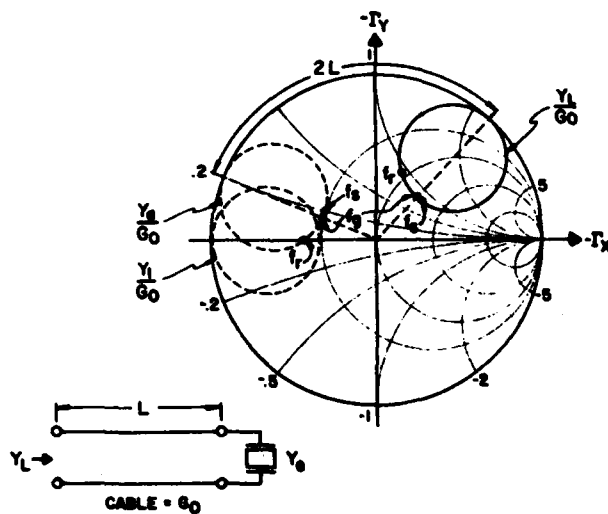


FIGURE 5 Γ -Plane Plot of Y_0/G_0 at end of line of characteristic admittance G_0 and length L

RESISTANCE RELATION

$$R_r > R_0$$

$$R_r = R_0$$

$$R_r < R_0$$

FREQUENCY RELATION

$$f_s < f_g < f_r$$

$$f_g = f_r$$

$$f_s < f_r < f_g$$

FIG. 6 RELATIONSHIP BETWEEN f_s , f_g AND f_r FOR DIFFERENT VALUES OF R_r WITH RESPECT TO R_0

	GENERAL EXPRESSION	INTERMEDIATE APPROXIMATION	APPROXIMATION FOR VERY LARGE m, M
δ_g	$\frac{C_1}{2C_0} \frac{1}{1+m^2} \quad (3)$		$\frac{C_1}{2C_0} \frac{1}{m^2}$
δ_r	$\frac{C_1}{2C_0} \frac{1}{2} [1 - (1 - 4/M^2)^{1/2}]$	$\frac{C_1}{2C_0} \frac{1}{M^2} (1 + \frac{1}{M^2})$	$\frac{C_1}{2C_0} \frac{1}{M^2}$

DEFINITIONS:

$$\delta_x = \frac{1}{2} \left(\frac{f_r}{f_s} - \frac{f_s}{f_x} \right)$$

$$m = X_0/R_0$$

$$M = X_0/R_1$$

$$\delta = \frac{f_x - f_s}{f_s}$$

FIG. 7 COMPARISON OF GENERAL AND APPROXIMATE
EXPRESSIONS FOR δ_g AND δ_r

TABLE I COMPARISON OF f_g MEASUREMENTS OF
CRYSTAL UNIT M-3 IN OVEN (71°C)

Cable	$\frac{\Delta f_g}{f_g} \times 10^8$	f_g^{**}	Physical Length (Ft)	Electrical Length (Deg)	Type	Nominal Impedance (Ohms)
A	*	.073 -29.1	1.5	9	RG-174	50
B	-1.9	.0601 -73.5	4	23.85	RG-55	53.5
C	-1.0	.0212 -156	12	64.05	RG-55	53.5
D	+1.1	.0137 -21	16	84.52	RG-55	53.5
C & D	+2.6	.0315 29.5				
E	+1.8	.0594 5.0	Approx 600	***	RG-58	53.5

* Reference $f_g = 9,905,116.72$ Hz. Cable A is outside connection to oven. All other cables are connected to A.

** w.r.t. 50 Ohms

*** Reel of Cable; Delays = 1.09 NS, Loss = 9.25 dB @ 9.905 MHz

TABLE II
CRYSTAL UNIT (M-3, 3rd OT - SC Cut) ELEMENT VALUES (25°C)

f_g (kHz)	R_1 (Ohms)	I(mA)	L_1 (mH)	C_1 (pF)	C_0 (pF)	r	Q(K)
9904.904	56.16	.057					
9904.906	55.80	2.90	1122.49	.23	4.37	19016	1262

AD P 002492

MEASURING METHOD OF EQUIVALENT SERIES CAPACITANCE AND NEGATIVE RESISTANCE OF QUARTZ CRYSTAL OSCILLATOR CIRCUITS

M. Toki, Y. Tsuzuki

Yokohama National University
Hodogaya-ku, Yokohama 240 Japan

and

T. Mitsuoka

Toyo Communication Equipment Co., Ltd.
Samukawacho, Koza-gun, Kanagawa Prefecture, Japan

Summary

A measuring method has been developed to obtain the oscillator circuit input impedance characteristics, so that the oscillator circuit operation performance can be estimated without connecting a crystal resonator.

The operation performance of the crystal oscillator can be estimated without connecting a crystal unit if we measure how the input impedance, which is usually represented by an equivalent capacitance and a negative resistance, behaves with the frequency and amplitude of the driving signal current of the circuit.

In this paper a measuring method of the equivalent capacitance and the negative resistance of the oscillator circuit has been considered. The equivalent capacitance and the negative resistance are measured by comparing them with a reference capacitance and a reference resistance, respectively.

A bridge circuit is adopted as a measuring circuit, and a measuring unit is developed so that the reference capacitance and the reference resistance of suitable values can be inserted. The measuring system is the same, except the measuring unit, as that of the measuring method of a crystal unit.

Measurements have been made on a Colpitts oscillator circuit. It has been shown from the experimental results that the input impedance characteristics can be measured by the newly developed measuring method and the oscillator circuit performance can be evaluated without connecting a crystal unit. It has also been shown that adjustment can be made on the equivalent capacitance of the oscillator circuit so that the frequency of oscillation for a given crystal resonator can be obtained.

Introduction

Crystal oscillators have been in wide use as important devices of electronic equipments, and extensive efforts have been made to improve performances of both the crystal unit and the electronic circuit of the oscillator. (1) It is, then, important to check a crystal oscillator circuit prior to putting it in use.

Crystal oscillators have usually been checked by observing how they operate upon connecting a certain crystal resonator of a given frequency. (1) The operation performance of the crystal oscillator, however, can be estimated without connecting a crystal unit, if the input impedance characteristics of the circuit facing the crystal unit are measured. In this paper, a measuring method of the input impedance of the crystal oscillator circuit is described.

The oscillation characteristics of an ordinary crystal oscillator such as a Colpitts oscillator as shown in Fig.1, can usually be represented by an equivalent circuit as shown in Fig.2. In Fig.2, the crystal resonator is represented by a set of equivalent series circuit parameters, an equivalent resistance R_e and an inductive reactance X_e at a given frequency. Facing the crystal unit, the input impedance of the oscillator circuit is represented by the series circuit composed of an equivalent capacitance C_L and a negative resistance $-R_L$.

Both the equivalent capacitance C_L and the negative resistance depend on the frequency and on the amplitude of the driving signal current. The operation performance of an oscillator circuit for a given set of crystal resonator parameters, then, can be estimated if we measure how the equivalent capacitance C_L and the negative resistance $-R_L$ behave with the frequency and amplitude of the driving signal current.

Measuring Method

Principle of measuring method

The principle of the measuring method of the equivalent capacitance C_L and the negative resistance $-R_L$ is as follows. When a certain resistance R_s is connected

cted in series with the oscillator circuit represented by CL and -RL and a RF signal is applied, the negative resistance -RL of the oscillator circuit can just be compensated by the resistance Rs on a certain signal current level of the circuit. Then, the total impedance of the series circuit becomes equal to the reactance of the equivalent capacitance CL. The value of CL is measured by comparing it with a reference capacitance Cs.

Measuring circuit configuration

A bridge circuit as shown in Fig.3 has been adopted as a measuring circuit. In Fig.3, the circuit circumscribed by the broken line is the oscillator circuit which is represented by CL and -RL. Rs is the reference resistance with which the negative resistance -RL is compensated. The series circuit of Rs, -RL and CL makes an element of the bridge circuit. Another element Cs is the reference capacitance with which the equivalent capacitance CL is compared. The two capacitances of the same value C are the remaining two bridge elements. Vs is the RF signal source voltage. V1 and V2 are the two output voltages of the bridge and are led to a voltage-phase meter.

Measuring unit

A measuring unit as shown in Fig.4 has been designed, so that the reference capacitance Cs and the reference resistance Rs of a suitable value can be used. In Fig.4, the portion of the unit indicated by 'A' is the measuring terminal through which a testing oscillator circuit is connected. 'B' and 'C' indicate the sockets for the reference capacitance Cs and the reference resistance Rs. 'D' indicates the input connector through which the RF signal is applied to the unit. 'E' and 'F' indicate the receptacles for the probes of the voltage-phase meter for measuring the bridge circuit output voltages V1 and V2, respectively.

Measuring procedure

In the measurement, the signal source voltage Vs is adjusted so that the phase difference of the two output voltages V1 and V2 becomes zero. The magnitude of the negative resistance -RL is, then, equal to the value of the reference resistance Rs. The value of the equivalent capacitance CL is calculated from the measured voltages V1 and V2 by the relation:

$$CL = (Cs + C)/(V1/V2) - C. \quad (1)$$

If V1 is equal to V2, the equivalent capacitance CL is equal to the reference capacitance Cs. Possible effects of capacitance differences of the bridge elements due to the stray capacities of the probe receptacles and the circuit wiring of an actual measuring unit can be eliminated, if we calibrate the CL measurement by means of connecting a standard capacitance in place of the series circuit of Rs and the testing oscillator circuit.

It is expected that the measuring method is also useful for adjustment of the equivalent capacitance CL, by means of an adjusting series capacitance, for instance.

Features of measuring bridge circuit

It can be seen that the measuring bridge circuit has features such as:

(1) Measuring errors due to the difference in the magnitude of V1 and V2 can be minimized, since the measurement can be made in a state that the magnitude of V1 is near that of V2.

(2) By the symmetrical configuration of the bridge circuit, effects of the stray capacitance and the residual inductance of the measuring unit can be eliminated if we make use of a measurement by means of comparing the equivalent capacitance CL and the negative resistance -RL with the reference capacitance Cs and the reference resistance Rs, respectively.

Measuring system

Fig.5 shows the measuring system using the above described measuring unit. The system is the same, except the measuring unit, as that of the measuring method of the crystal resonators by means of a measuring unit such as the pi-network or Fr-meter. (2),(3),(4) In the system, a YHP vector voltmeter is utilized as a voltage-phase meter. A variable attenuator is used in order to adjust the driving signal current level of the oscillator circuit to be measured.

Measurement of Oscillator Performance

Measurements have been made on the oscillator circuit of Fig.1. As the input impedance characteristics, the frequency characteristics of the equivalent capacitance CL have been measured by keeping the value of the negative resistance -RL on a certain constant value. The measured results are shown in Fig.6(a). The signal current I has also been calculated and its frequency characteristics are shown in Fig.6(b). The characteristics have been measured for several parameter values of the negative resistance -RL over a frequency range up to 40MHz. The range of the magnitude of the negative resistance -RL has been chosen so that it covers the equivalent resistance of ordinary crystal resonators. From these measured results, the following oscillator circuit performances can be seen.

(1) Suppose, for instance, a crystal resonator which has an equivalent resistance Re of 30 to 40ohms is used and a magnitude of -RL two or three times as large as the crystal resistance Re is required for a sure oscillation to take place. If this is the case, it can be seen from Fig.6(a) that the oscillator circuit of Fig.1 can be used in a frequency range from 6 to 22MHz. The frequency range is also seen more clearly on the frequency axis in Fig.6(b).

(2) It can be seen from Fig.6(a) that, over the above evaluated frequency range from 6 to 22MHz, the equivalent capacitance CL has a value between 30 and 60pF if the above described crystal resonator with the equivalent resistance Re of 30 to 40ohms is used.

(3) It can also be seen from Fig.6(a) that at a particular frequency near 17MHz, the value of the equivalent capacitance CL is nearly constant for a wide range of the magnitude of -RL. Consequently, as far as the oscillator circuit of Fig.1 is concerned, it is especially desirable to use it at a frequency near 17MHz.

Thus it has been shown that the oscillator circuit performances can be estimated from the measured results of the input impedance characteristics of the circuit.

Adjustment of Equivalent Capacitance

In order to show a usefulness of the measuring method, adjustment has also been made on the equivalent capacitance CL of the oscillator circuit of Fig.1. The frequency of the crystal resonator is usually made so as to have the nominal frequency at which the crystal resonator is antiresonant with the so-called load capacitance.(1) This frequency is called the load resonance frequency and gives the frequency of oscillation when the crystal resonator is connected with the oscillator circuit whose equivalent capacitance CL is equal to the load capacitance of the crystal resonator.(1)

By adjusting a variable capacitor connected in series with the circuit, the equivalent capacitance CL of the oscillator circuit has been adjusted to 30pF, which is one of typical values of the load capacitance for fundamental mode AT-cut crystal resonators. A couple of crystal resonators have been used for each testing frequencies (except for 30MHz). Frequency of oscillation F_{osc} has been measured upon connecting the test crystal resonator with the above adjusted oscillator circuit. The adjustment of the equivalent capacitance CL and the measurement of the frequency of oscillation F_{osc} have been repeated twice for each testing crystal resonator. The differences between the load resonance frequency FL and the frequency of oscillation F_{osc} are shown in Fig.7. As Fig.7 shows, the differences are less than 2ppm over the frequency range from 10 to 30MHz. The difference of 2ppm in frequency corresponds approximately to that of 0.3pF in the equivalent capacitance CL . Thus it has been shown that the measuring method is also useful for adjustment of the oscillator circuit equivalent capacitance.

Conclusion

It has been shown from the experimental results that the oscillator circuit input impedance characteristics can be measured by the newly developed measuring method and the crystal oscillator performances can be evaluated without connecting a crystal unit. It has also shown that the adjustment of the oscillator circuit equivalent capacitance can be made by the measuring method.

Thus it is concluded that the newly developed measuring method is useful for the development and adjustment of crystal oscillator circuits.

Acknowledgement

The authors wish to thank Mr. S. Okano and Mr. K. Hiram of Toyo Communication Equipment Co., Ltd. for their participation in the development of the measuring method.

References

- (1) Marvin E. Frerking: "Crystal oscillator design and temperature compensation", Van Nostrand, 1978.
- (2) IEC 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", 1973.
- (3) Issac Koga: "Measurement of a VHF crystal unit using a "Fr-meter" and an "Auxiliary reference resistor", 1980.
- (4) T. Adachi, M. Toki and Y. Tsuzuki: "Digital Measurement Method of VHF Quartz Crystal Resonators" (in Japanese), Trans. IECE, '82/8. Vol.J65-A, No.8, 1982.

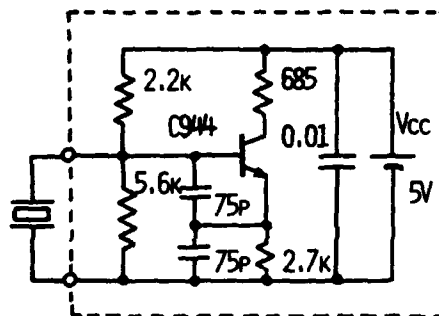


FIG.1. A COLPITTS CRYSTAL OSCILLATOR CIRCUIT.

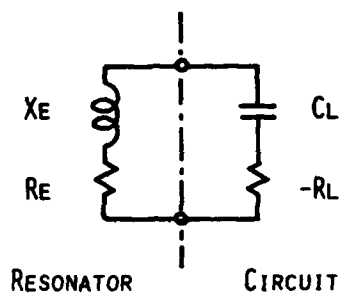


FIG.2. CRYSTAL OSCILLATOR EQUIVALENT CIRCUIT.

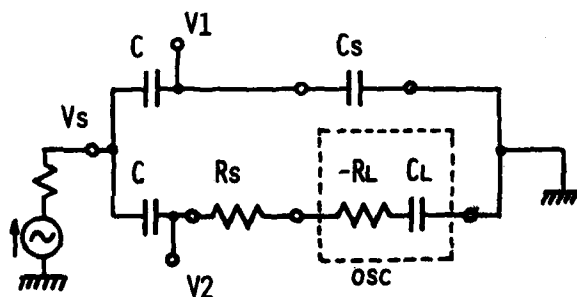
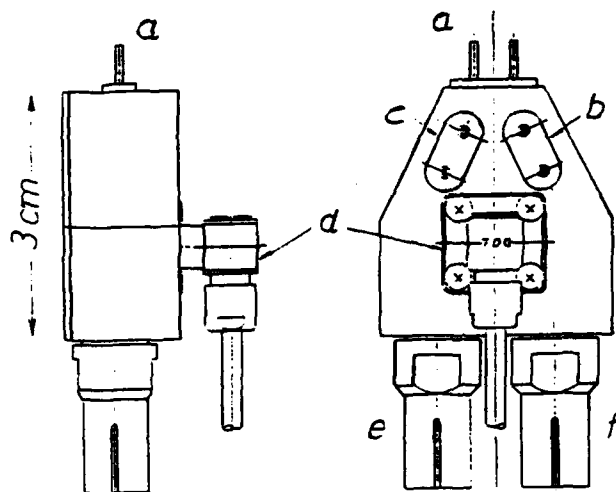


FIG.3. MEASURING BRIDGE CIRCUIT OF CRYSTAL OSCILLATOR.



a: MEASURING TERMINAL FOR OSC CIRCUIT
 b: SOCKET FOR REFERENCE CAPACITOR C_s
 c: SOCKET FOR REFERENCE RESISTOR R_s
 d: CONNECTOR FOR RF SIGNAL INPUT V_s
 e: PROBE RECEPTACLE FOR MEASURING V_1
 f: PROBE RECEPTACLE FOR MEASURING V_2

FIG.4. CONFIGURATION OF OSCILLATOR CIRCUIT MEASURING UNIT.

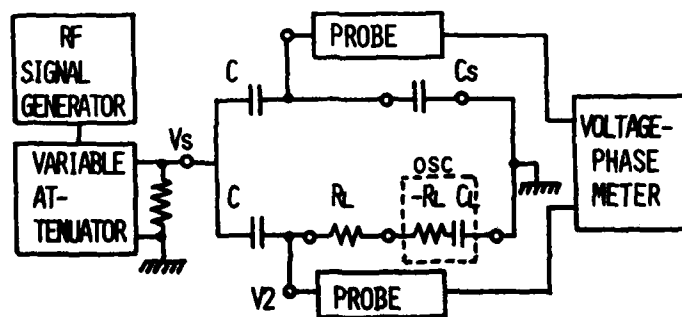


FIG.5. MEASURING SYSTEM OF CRYSTAL OSCILLATOR.

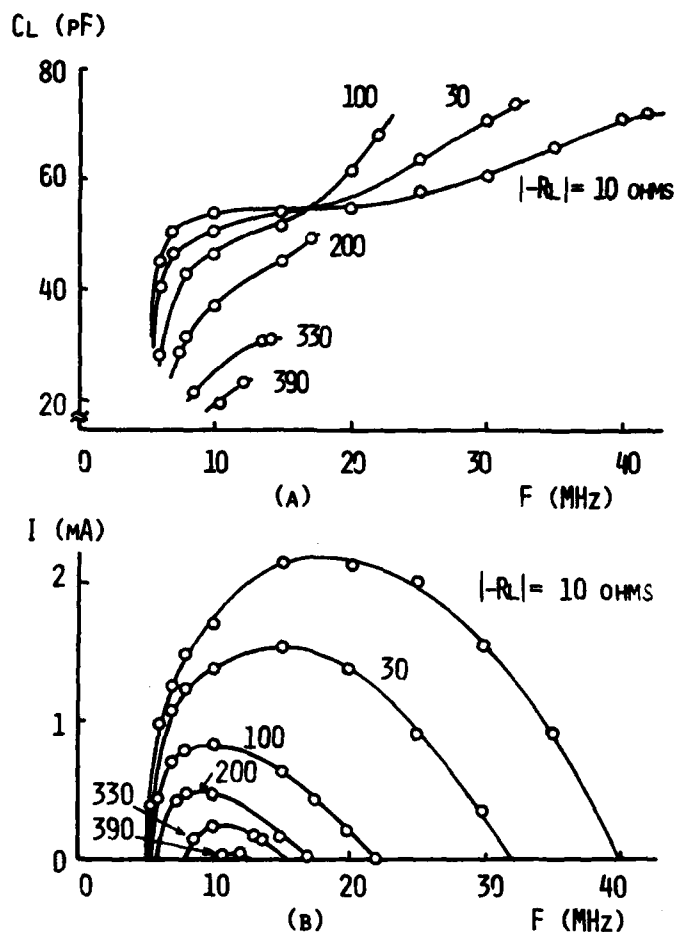


FIG. 6. FREQUENCY CHARACTERISTICS OF (A) THE MEASURED EQUIVALENT CAPACITANCE C_L AND (B) THE CALCULATED CURRENT LEVEL OF OSCILLATION I .

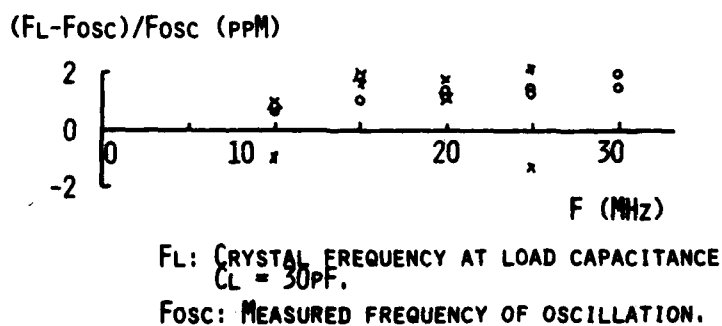


FIG. 7. COMPARISON OF FL WITH F_{osc} . (OO AND XX AT EACH FREQU. INDICATE THAT ADJUSTING OF C_L AND COMPARING HAVE BEEN REPEATED TWICE.)

AD P 002493

S.Y. PARAMETERS METHOD FOR ACCURATE MEASUREMENTS OF BULK WAVE CRYSTAL RESONATORS AT FREQUENCIES UP TO 2 GHz

J.P. AUBRY, E. GERARD and S. LECHOPIER

COMPAGNIE D'ELECTRONIQUE ET DE PIEZO ELECTRICITE
95100 ARGENTEUIL (FRANCE)

SUMMARY

Bulk wave quartz crystal resonators with very-high Q factors or of very high frequencies have received great attention in recent years and they do require new methods of characterisation.

A general method allowing accurate measurements of high Q or high frequency crystal resonators is described.

The basic idea is the use of the admittance matrix of the resonator.

For low frequency resonators (≤ 200 MHz), the transmission method through the classical IEC π network leads to the determination of Y_{12} (f) and gives accurate results of motional parameters and frequency. For VHF crystals (up to 2 GHz) the Y matrix is computed from the four S parameters.

The equipment used the basic steps of the method and the calibration and measurement procedures are described. Typical results are presented. Accuracy and reproducibility of measurements are discussed for both high Q and VHF crystals.

I - INTRODUCTION

Among the applications of modern piezo electricity⁽¹⁾, crystal resonators with very high Q factor or of (2-3) very high frequencies are more and more studied.

Without neglecting the difficulties to overcome when manufacturing such crystals, a major problem to solve is to perform reliable and accurate measurements of these crystals.

Numerous methods and equipments have been described⁽⁴⁻⁷⁾ to reach more or less accurate measurements of crystals. These methods can be active (Crystal Impedance meter) or passive (π or T network transmission). Impedance bridges have also been used and more recently reflexion bridges⁽⁸⁾ or S parameters network analyser⁽⁹⁾ have been described.

All of these methods have their own advantages but they also have their own limitations.

Classical methods of passive networks measurements, such as the IEC - π network⁽⁵⁾, are not suitable for the crystals we are dealing with. The limitations come from :

- frequency range mismatch between crystal and network
- bandwidth of high Q crystals
- crystals having no null-phase response
- operator dependence of measurement.

Other methods can require a large number of experimental measurements in a wide frequency range to allow precise fitting or accurate results⁽⁸⁾.

In other hand, industrial purpose requires that the operator contribution must be as reduce as possible, by using computer controlled equipment.

In order to overcome these limitations, CEPE has developed a fully automatic equipment. The basic idea is to use the admittance matrix of the crystal. Variations of the admittance matrix are recorded as a function of frequency, in a range less than ± 20 ppm around f_0 .

II - THE ADMITTANCE PARAMETERS

Around its resonant frequency, a crystal resonator can be described by the simplified equivalent circuit of figure 1a or 1b. In these figures :

f_0 is the resonant frequency of the serie arm
 R_1, L_1, C_1

C_0 is the static capacitance

C_2, C_3 are the holder capacitances.

Looking the crystal as a two-port network (fig. 1.b) leads to the definition of the admittance matrix

$$Y = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix}$$

From fig. 1.b one has : $Y_{12} = Y_{21} = -Y_Q$)

$$Y_{11} + Y_{12} = Y_2 \quad (1)$$

$$Y_{22} + Y_{21} = Y_3 \quad)$$

It can be easily shown that :

$$Y_Q = j C_0 \omega + \frac{1}{R_1 + j (L_1 \omega - \frac{1}{C_1 \omega})}$$

which transform in :

$$Y_Q = \frac{R_1}{R_1^2 + X_1^2} + j (C_0 \omega - \frac{X_1}{R_1^2 + X_1^2}) \quad (2)$$

in which $X_1 = L_1 \omega - \frac{1}{C_1 \omega}$

The values of Y_Q as a function of frequency can be reached from (1).

It is possible to use the full relation (2) to reach the equivalent circuit of the crystal. This is the well known determination of the circle of admittance. When calculating this circle, it appears that a good fit is obtained with a large number of experimental points (20), carefully distributed in a wide frequency range, from resonant to parallel frequencies. It appears also that stray capacitances must be taken into account (mainly in the test circuit). Additional difficulties arise from the large difference of measured impedances, giving a mismatch of adaptation of impedances in the covered frequency range.

An alternative solution in using the admittance parameters is to consider only the reciprocal of the real part of (2) :

$$\frac{1}{R_e(Y_Q)} = \frac{R_1^2 + X_1^2}{R_1} = G^{-1} \quad (G = \text{conductance})$$

Using the well known narrow band approximation :

$$X_1 = L_1 \omega - \frac{1}{C_1 \omega} = 2 L_1 (\omega - \omega_0) = 2 L_1 \Delta \omega$$

G^{-1} becomes :

$$\frac{1}{R_e(Y_Q)} = R_1 + \frac{(2 L_1 \Delta \omega)^2}{R_1} \quad (3)$$

This relation describes a parabolic curve as a function of frequency. Values of the equivalent circuit are then calculated from the coefficients of a parabolic curve, fitted by a least-square method from the experimental points.

Advantages of this method are as follows :

- small numbers of points are required ($N=7$)
- points can be taken in the close neighbourhood of f_0 .
- small variations of impedance of the crystal in the covered frequency range, allowing good impedance matching of test circuit
- measurement independent of C_0 .
- general method usable disregarding any zero phase response.

The general procedure of measurement can then be described as follows. After an appropriate calibration procedure, the crystal is inserted in a test circuit. A set of measurements is performed stepping in frequency around the assumed resonant frequency. Computer aid will allow the adjustment of the measurement conditions on the crystal under test (frequency range, location of G^{-1} parabola versus frequency ...)

After completion of this set of measurements, least square fitting is performed and interpreted. The minimum of the G^{-1} curve gives the value of R_1 , the frequency at this minimum is the series arm resonant frequency f_0 , and the second order term of G^{-1} gives L_1 (fig. 2).

Criteria for a "good" measure can be introduced in the computer program.

III - LOW FREQUENCY CRYSTALS : Y TRANSMISSION METHOD

For crystal resonators of frequency below 200 MHz, the transmission measurement through a π network gives enough information to reach the motional elements. This method can be successfully used with high Q crystals.

Description of the equipment

The equipment used for transmission measurement is described on figure 3. This system includes :

- an ADRET frequency synthesiser (0,01 Hz step) delivering a highly stable signal (cesium controlled)
- a RHODE et SCHWARZ vector voltmeter (ZPV) allowing complex ratios and voltage measurements
- a power splitter
- a IEC 444 π network

Both synthesiser and voltmeter are computer controlled through the IEEE 488 bus.

Calibration and measurement procedure

The first step to perform, assuming that apparatus have been carefully calibrated is to calibrate the full equipment. This is obtained by inserting in the network the IEC recommended shorting blank.

Voltage attenuation is measured at nominal frequency (phase and amplitude). These values will be taken as references in all further measurements performed at $f \leq f_0 \pm 1\%$. Voltmeter can store these references in its own memories in order to display only the induced attenuation. This calibration procedure is automatically performed and checked. The shorting blank being inserted, voltage ratio must be read as 1 : 0.

The measurement procedure starts by a rapid determination of f_0 . This is done by an iterative phases reading at two frequencies (f_1, f_2) close to the previously estimated f_0 (starting from f_N). Sign and magnitude of both phases give the new determination of f_0 . Three iterations generally lead to a correct estimation of f_0 .

The following step is to perform a run of N measurements at N frequencies centered around f_0 . At each frequency, phase and amplitude attenuations are measured, and the value of $G^{-1} = \frac{1}{R_e(Y_Q)}$ is calculated

through the classical voltage transfer relation of the π network :

$$\frac{1}{G} = R_{\pi} \times \frac{1 + \left(\frac{|A_c|}{|A_s|}\right)^2 - 2 \left(\frac{|A_c|}{|A_s|}\right) \cos(\theta_c - \theta_s)}{\left(\frac{|A_c|}{|A_s|}\right) \cos(\theta_c - \theta_s) - 1} \quad (4)$$

in which $A_s = \frac{V_B}{V_A} = |A_s| e^{j\theta_s}$ is the voltage

attenuation between input and output of the π network when the crystal is inserted.

If the calibration procedure on the short circuit has been performed first, one can take in (4)

$$|A_c| = 1 \text{ and } \theta_c = 0 ;$$

$A_c = |A_c| e^{j\theta_c}$ being the voltage transfer on the shorting blank.

The frequency range covered by the N steps of frequency is given mainly by the overtone order of the resonator (from $\pm 20 \cdot 10^{-6}$ for a fundamental mode to $\pm 1 \cdot 10^{-7}$ for a fifth or higher overtone).

Various requirements must be met to allow a correct interpretation of the measurements :

- minimum of $\frac{1}{R_e(Y_Q)}$ must be seen between 2nd and N-1st point.
- standard deviation of the parabolic fitting (given in n) must be lower than a given percentage of the computed value of R_1
- minimum of the G^{-1} curve must be positive.
- the coefficient of the 2nd order term must be positive
- ratio between the greater and the smaller $\frac{1}{R_e(Y_Q)}$ observed must be found between given limits.

If one of these conditions is not fulfilled, an other run of measurement is operated after a possible modification of center frequency or frequency range.

Accuracy and reproducibility

Accuracy of the motional parameters and frequency comes from :

- accuracy of the method which will be discussed
- accuracy of the test circuit (π network) introducing systematics errors.

The accuracy of the measurements is related to the G^{-1} fitting curve, which comes from the accuracy of each $\frac{1}{R_e(Y_Q)}$ experimental points.

This accuracy can be calculated from (4) by taking into account the specifications of the vector voltmeter. It can also be estimated by a run of 100 successive measurements of $\frac{1}{R_e(Y_Q)}$ at a given

frequency.

For a given calibration, this error is found about 1 %.

The effect of this possible error on each $\frac{1}{R_e(Y_Q)}$

points on the G^{-1} fitted curve is estimated by affecting 1 to N experimental points (randomly distributed in the N points) by 1 % error. Motional elements and frequency are recalculated at each time. Table I summarizes this computational analysis for various resonators

	Q (10^6)	$\Delta f/f$	$\Delta L/L$	$\Delta R/R$
BVA 5 MHz AT P/5	2.3	$\pm 1.35 \cdot 10^{-9}$	$\pm 0.8 \%$	$\pm 0.8 \%$
10 MHz SC P/5	1.3	$\pm 3.5 \cdot 10^{-9}$	$\pm 2.5 \%$	$\pm 0.7 \%$
10 MHz SC P/3	1.3	$\pm 3 \cdot 10^{-9}$	$\pm 2 \%$	$\pm 0.7 \%$
100 MHz AT P/5	0.1	$\pm 5 \cdot 10^{-8}$	$\pm 2 \%$	$\pm 0.7 \%$

Table I

Reproducibility of measurements is given by figure 4, in which a set of 50 successive measurement is shown. These measurements have been performed on 5 MHz AT P/5 CEPE BVA design.

Both accuracy and reproducibility are found in the range of some parts in 10^{-9} .

Typical results

Typical results of measurement are presented on figures 5 to 7.

The horizontal axis is in $\frac{\Delta F}{F}$ around a nominal frequency (or around a pre-estimated f_0). The vertical axis gives the $\frac{1}{R_e(Y_Q)}$ values in ohms.

The stars are the experimental points and the solid line is the parabolic G^{-1} curve calculated by a least square fit.

Figure 5 gives the measurements of a 5 MHz AT P/5 (CEPE BVA design) resonator.

Figure 6 gives the measurement of a 10 MHz SC P/5 resonator.

Figure 7 gives the measurement of a 10 MHz SC P/3 resonator.

IV - VHF CRYSTALS - S-Y PARAMETERS METHOD

For frequencies higher than 200 MHz, crystals generally do not show any zero phase response and, in addition, stray (or holder) capacitances are more and more perturbing. Equivalent circuit must then be seen as a two port network.

It is well known that S parameters allow an accurate measurement of the elements of the VHF equivalent circuit (4-8). We will show a procedure of interpretation of S measurements which is usable and accurate for VHF crystals, and which requires a minimum number of measurements and saves time.

For a given run of frequencies around f_0 , we measure the S_{ij} parameters and then, by the relations given in figure 9, compute the Y matrix. By (1) and (3) we will calculate, from Y_{12} or Y_{21} , the reciprocal of the real part of Y_Q , giving access, by the G^{-1} curve fitting to f_0 , R_1 , L_1 .

Capacitances will be calculated from Y_{11} and Y_{22} .

It can be seen that we will have two sets of points to calculate Y_Q , one given by Y_{12} and the other by Y_{21} .

This gives a mean to check the good operation of the system.

Description of the equipment

The S parameters measurement requires :

- a frequency synthesiser
- a S parameter test adapter
- a vector voltmeter
- a test circuit (50 Ω load adaptation)

The equipment is described in figure 8.

The synthesiser, S adapter and vector voltmeter are under computer control through IEEE 488 bus.

The various relays giving the different ways of RF propagation (transmission and reflexion) are shown in figure 8.

The computer is in charge of applying the frequencies, the status of the relays and of reading the informations on the vector analyser. It will also give the interpretation in terms of f_0 , R_1 , L_1 , C_0 , C_2 , C_3 ..

Calibration procedure

The minimum calibration of the system is performed on S_{11} , with an open circuit on the test circuit.

After initial reading, S_{11} is adjusted to an amplitude = 1 and a phase = 0.

For highest frequencies, it can be measured that S_{ij} parameters ($i, j \neq 1,1$) are slightly different than their theoretical values ($S_{22} = 1 : 0$ on open circuit, $S_{12} = S_{21} = 1 : 0$ on short circuit). This can be due to unequivalent electrical lengths, rotations of phase ...

Some workers have overcome this problem by introducing "error Matrix" in their computation

We have chosen an other solution which consists in measuring, at the nominal frequency, after S_{11} calibration, S_{22}^m (open circuit) and S_{12}^m and S_{21}^m (short circuit). Then we calculate S_{ij} parameters (compensation parameters) by :

$$S_{ij}^m \cdot S_{ij}^c = 1 \quad (i, j \neq 1,1) \quad (5)$$

Any further S_{ij} measurements performed on a crystal will be multiplied by these three error parameters.

An example of such a compensation is given on figure 10 (600 MHz AT fundamental quartz crystal).

Figure 10.a gives the G^{-1} fitting without taking into account the compensation. Figure 10.b gives the same curve calculated with the S_{ij} compensation. It can be seen that the fit is better and that differences between Y_{12} and Y_{21} are lower when using compensation. Nevertheless, it is shown that this compensation does not affect greatly the values of the elements, indicating a good equilibrium between electrical lengths of the S adapter.

After completion of calibration procedure, crystal is inserted in the test circuit.

Measurement procedure

As in the low frequency operation, the measurement procedure starts by a rapid determination of f_0 .

This is done by using the amplitude of S_{11} which must be minimum around f_0 . Two S_{11} readings are performed at frequencies (f_1, f_2) close to a previously estimated f_0 (starting from f_N). Amplitudes of both S_{11} measurements lead to a new determination of f . In most cases, three iterative pairs of S_{11} measurement give the location of f_0 .

A run of N measurements is performed at frequencies around f ($N=9$ in most applications). Each measurement is made of the set of four S parameters. For each frequency of the run, the calculator computes, by (5), the compensated S matrix, and then the Y matrix. Both Y_{12} and Y_{21} are used to give a pair of $\frac{1}{R_e(Y_Q)}$

experimental points at each frequency step.

The G^{-1} curve is then fitted on the 2.N points, as a function of frequency.

Accuracy and reproducibility

Here again, accuracy of the method of calculation of motional elements and frequency can be estimated, dis-regarding the influence of mismatch of characteristic impedances, or imperfections in the test circuit, which can be checked by 50 Ω tests and references.

Accuracy of the $\frac{1}{R_e(Y_Q)}$ points can be taken from the

voltmeter and S adapter characteristics. Assuming an accuracy around 1 % on each $\frac{1}{R_e(Y_Q)}$ readings, one can

do the same computational analysis than in the low frequency case. This will give the order of magnitude of the accuracy of the calculated motional elements and frequency. Table II summarizes this analysis.

	Q	$\Delta f/f$	$\Delta L/L$	$\Delta R/R$
600 MHz AT (F)	$10 \cdot 10^3$	$\pm 3.4 \cdot 10^{-7}$	$\pm 1.2 \%$	$\pm 0.8 \%$
200 MHz SC (F)	$22 \cdot 10^3$	$\pm 1.6 \cdot 10^{-7}$	$\pm 1.8 \%$	$\pm 0.8 \%$
1,06 GHz BT P/5	$28 \cdot 10^3$	$\pm 1.8 \cdot 10^{-7}$	$\pm 2 \%$	$\pm 0.8 \%$

Table II

Reproducibility has been checked by performing 50 successive measurements, after a given calibration, on a 600 MHz AT fundamental mode crystal resonator.

Figure 11 gives the variation of f_0 , L_1 and R_1 during these fifty measurements. It appears that frequency does not change by more 0,8 ppm (500 Hz) during the operation. It can be thought that most of the frequency drift come from small temperature variations during the 1 hour and half of time covered by the 50 measurements.

Accuracy and reproducibility in L and R appear to be better than 2 and 1 %, and in the range of some parts in 10^{-7} on frequency, even at 1 GHz.

Typical results

Typical measurement results are presented in figures 12 to 14.

figure 12 : SC 200 MHz fundamental

figure 10 : AT 600 MHz fundamental

figure 13 : BT 1,060 GHz P/5

figure 14 : AT 1,9 GHz P/3.

In each figure, stars indicate $\frac{1}{R_e(Y_Q)}$ calculated from

Y_{12} and Y_{21} , and the solid curve gives the parabolic curve fitted on the 2N points.

All of these crystals were manufactured by CEPE using the now well known ion beam milling technique (11) (fig. 15).

V - DISCUSSION

One can make a comparison between the two methods described above, and with the IEC W network method.

Table III gives the results of measurements on 70 MHz fundamental mode AT cut crystals.

	Zero Phase W Network	1 / G W Network	1 / G S parameters
Frequency (Hz)	69960648 7.10 ⁻⁸ (*)	69960618 5.10 ⁻⁸ (*)	69960642 3.10 ⁻⁸ (*)
Resistance (Ω)	9.89 0.5%(*)	9.91 0.2%(*)	9.67 0.2%(*)
Inductance (mH)	2.705 1.2%(*)	2.680 0.3%(*)	2.701 0.2%(*)
Q factor	120230	118880	122780

(*) : dispersion on 5 measurements

Table III

Table IV gives the industrial result of such crystals, based on thousand of units manufactured, and measured by the above methods.

	Initial tolerance	R (Ω)	L (mH)	Q
+ 4 ⁺ - 1 ⁺	+ 4 10 ⁻⁶	13,5 ± 1	2.85 ± 0.12	93000

Table IV

Table V gives a comparison between Y transmission and S-Y methods measurements performed on 181,666 MHz AT cut crystals, for 5 successive measurements (in ambient air).

	Y. transmission	S. Y
F	181 665 819,2 ± 20 Hz	181 666 128,5 ± 10 Hz
R	104,1 ± 0,25 Ω	101,9 ± 0,1 Ω
L	5.284 ± 0.02 mH	5.431 ± 0.03 mH

Table V

Between both methods, disagreement in L₁ and R₁ is smaller than 2 % and frequency errors (~2 ppm) are probably due to differences in temperature during each run of measurements.

CONCLUSION

We have described a fully programmable equipment, allowing automatic calibration and measurement procedures. Suitable programs allow to choose the best conditions in order to reach a given accuracy in the minimum time. General method is usable from KHz to GHz range.

The described procedure leads to the serie arm characterisation, independant of Co and stray capacitances. Without any compensation circuit, it allows the measurement of crystals which do not show zero phase response. It is also possible to detect anomalies in frequency-phase curve or to characterise a crystal at a given phase offset.

Typical values are reported up to GHz range.

At 1 GHz QF product as high as 30 10¹² is obtained with BT cut fifth overtones. QF products of AT cut are between 6 10¹² (fundamental) and 13 10¹²

These results seem to be the greater Q value which can be obtained, even by SAW or composite devices, in this GHz frequency range.

REFERENCES

- 1/ L. BIDART and J. CHAUVIN "Direct Frequency Crystal Oscillator", 35th AFCS, 1981 pp 365-375
- 2/ L. BIDART, "New Design of Very High Frequency Sources" Ferroelectrics 1982 vol.40 pp 231-236
- 3/ B. d'ALBARET and P. SIFFERT, "Recent Advances in VHF Crystal Filters", 36th AFCS, 1982 pp 405-418
- 4/ H.S. PUSTARFI and W.L. SMITH, "An Automatic Crystal Measurement System", 27th AFCS, 1973, pp 63-72
- 5/ IEC Publication 444 (1973)
- 6/ W.H. HORTON, T.S. PAYNE, R.C. SMYTHE and D.A. SYMONDS, "Comparison of methods for measurements of Quartz Crystal Resonators with load-capacitance", 35th AFCS, 1981, pp 271-279
- 7/ R.C. SMYTHE, "An Automated Resonator Measurement System using a Reflexion Coefficient Bridge", 35th AFCS, 1981, pp 280-285
- 8/ R.C. PEACH, A.J. DYER, A.J. BYRNE and S.P. DOKERTY, "New Method for the Measurement of Quartz Crystal Parameters", 36th AFCS, 1982, pp 297-301
- 9/ C.S. STONE and O.J. BALTZER, "A Frequency Domain Reflectometer for Quartz Resonator Investigations", 36th AFCS, 1982, pp 321-326
- 10/ R. FISCHER and L. SCHULZKE, "Extending the Frequency Range of the Transmission Line Method for the Measurement of Quartz Crystal up to 250 MHz", 31st AFCS, 1977, pp 96-101
- 11/ M. BERTÉ, "Acoustic Bulk Wave Resonators and Filters Operating in the Fundamental Mode at Frequencies Greater than 100 MHz", 31st AFCS, 1977, pp 122-125.

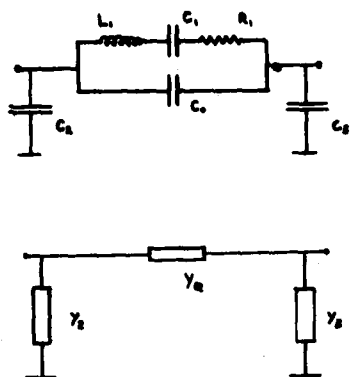


figure 1

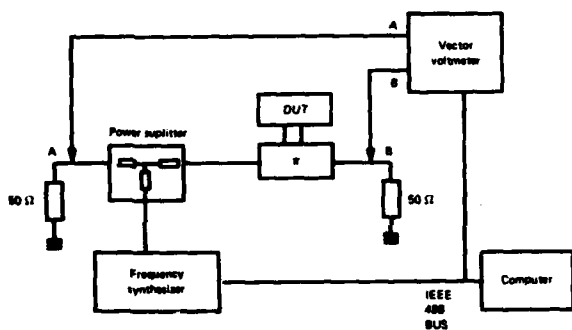


figure 3

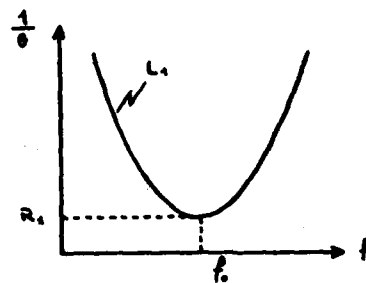


figure 2

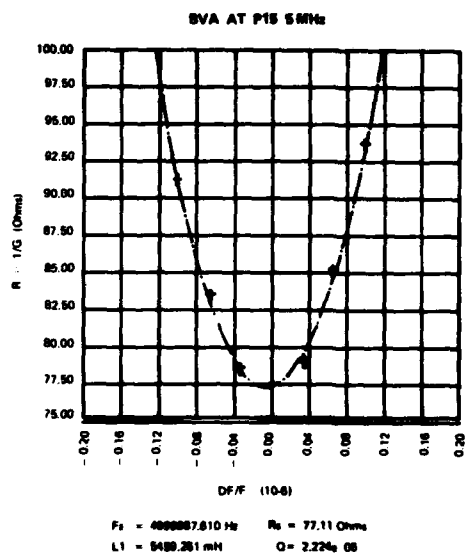


figure 5

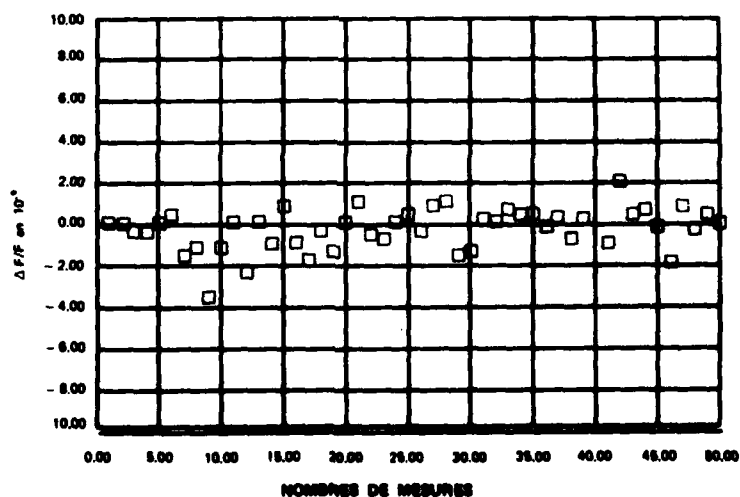


figure 4

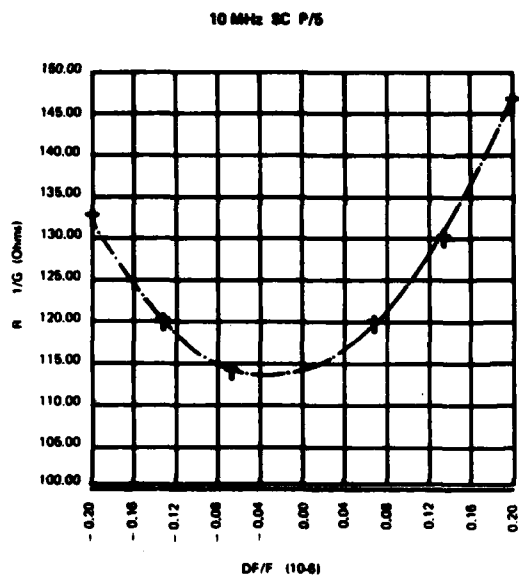


figure 6

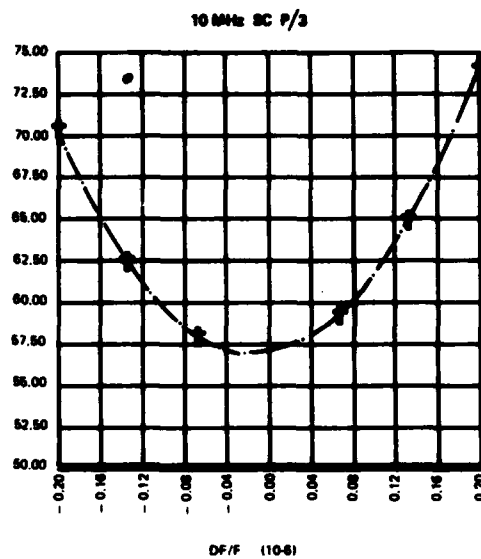


figure 7

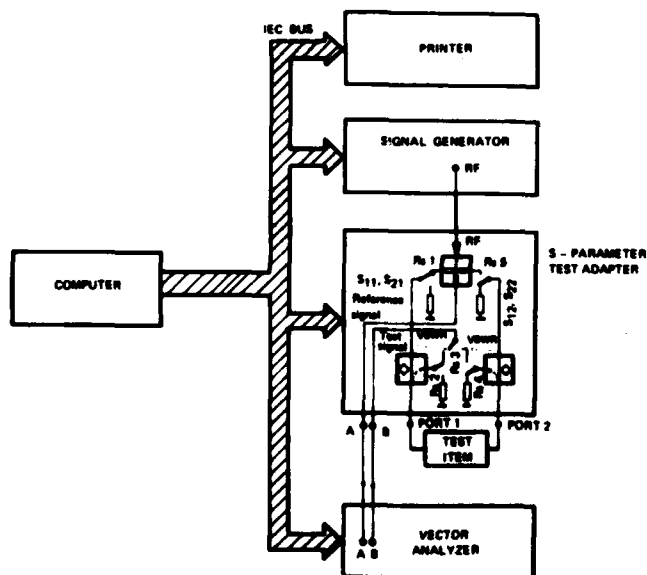


figure 8

RELATIONS BETWEEN S and Y PARAMETERS

$$Y_{11} = \frac{(1-S_{11})(1+S_{22})+S_{12}S_{21}}{[(1+S_{11})(1+S_{22})-S_{12}S_{21}] \cdot Z_L}$$

$$Y_{12} = \frac{-2S_{12}}{[(1+S_{11})(1+S_{22})-S_{12}S_{21}] \cdot Z_L}$$

$$Y_{21} = \frac{-2S_{21}}{[(1+S_{11})(1+S_{22})-S_{12}S_{21}] \cdot Z_L}$$

$$Y_{22} = \frac{(1+S_{11})(1-S_{22})+S_{12}S_{21}}{[(1+S_{11})(1+S_{22})-S_{12}S_{21}] \cdot Z_L}$$

figure 9

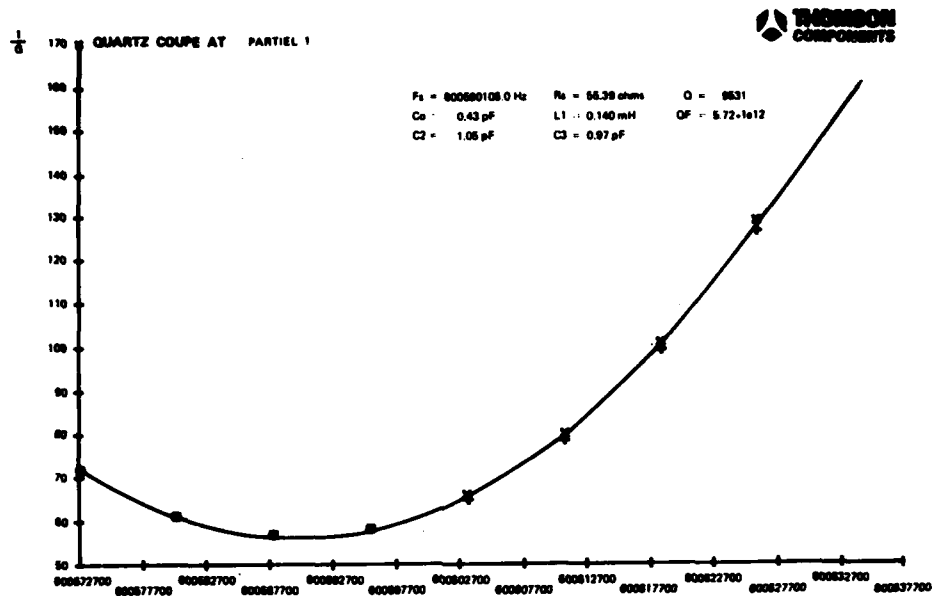


figure 10 a

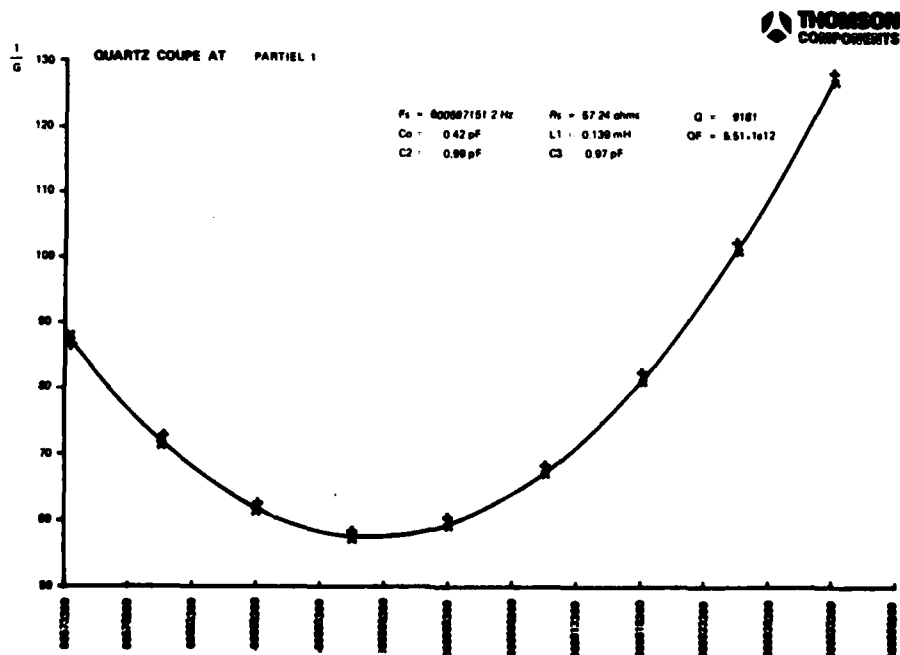


figure 10 b

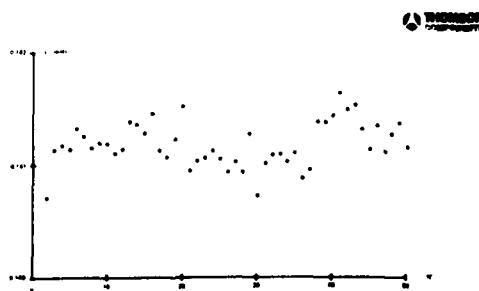
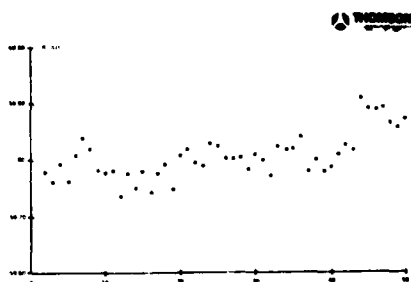
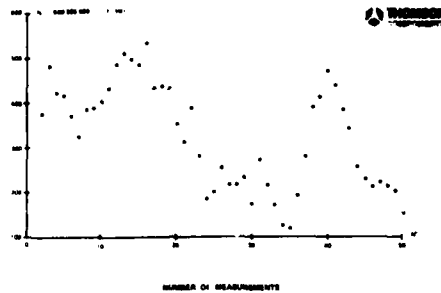


figure 11

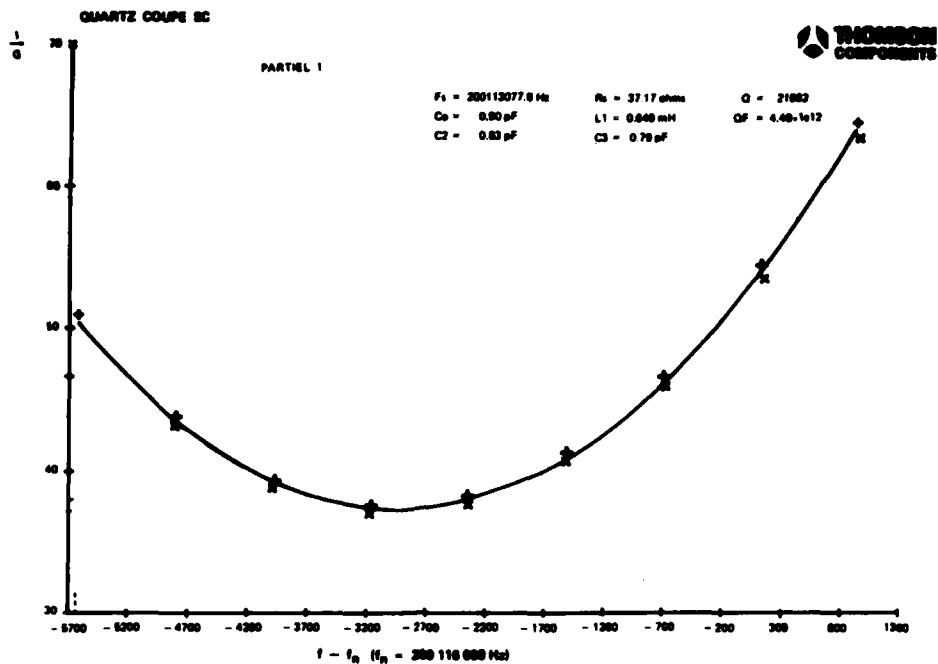


figure 12

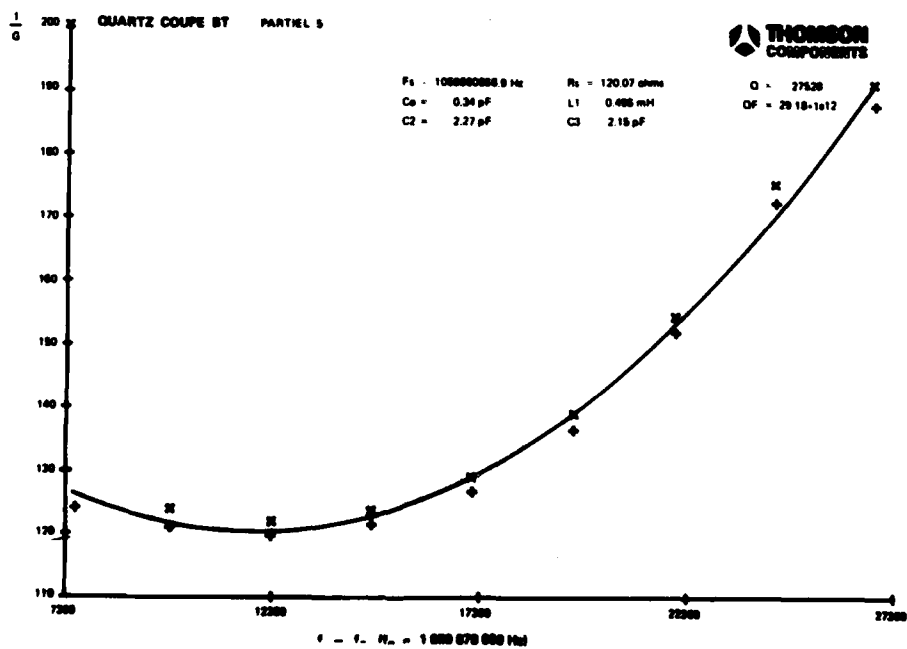


figure 13

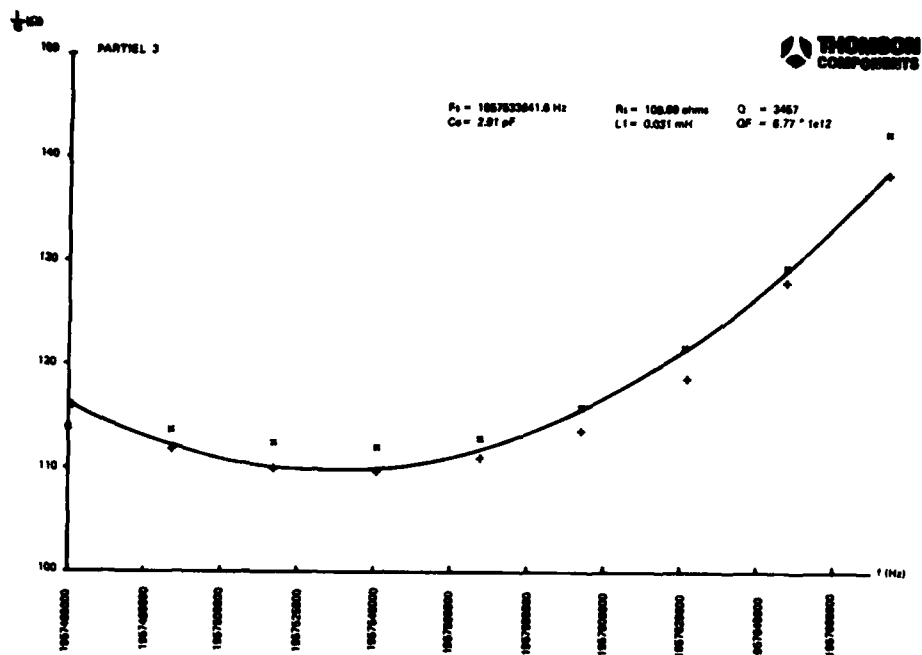


figure 14

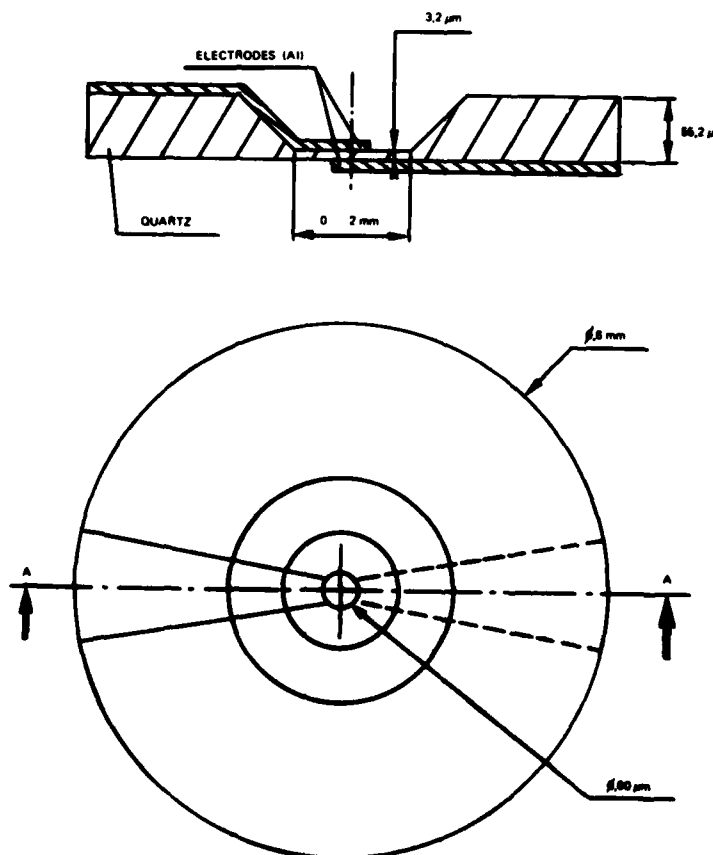


figure 15

SPURIOUS RESONANCES AND MODELLING OF COMPOSITE RESONATORS*

K. B. Yoo, H. Überall, D. Ashrafi and S. Ashrafi

Department of Physics

The Catholic University of America

Washington, DC 20064

Abstract

Several three-dimensional models have been developed for explaining the characteristics of composite resonators. The agreement between the resonance spectra predicted by the models and the experimental data is excellent, including spurious resonance frequencies. The size and geometry of the resonator is found to be directly responsible for the spurious resonances, and to be important for the principal resonances. A significant relationship between resonance frequencies and material properties has been found, and its implication is discussed.

1 Introduction

Bulk acoustic wave resonators on a silicon substrate with fundamental frequency in the VHF/UHF range have been developed by several groups.¹ Although proto-types have been successfully fabricated, they all exhibited the undesirable feature of spurious resonances. The main purpose of this work is to understand the origin of the spurious resonances, so that means for their effective control can be found.

The composite resonator¹ consists of a thin silicon membrane etched out of a block of silicon to a thickness of a few microns, with a sputter-deposited piezoelectric ZnO film of comparable thickness. At the top and bottom of the piezoelectric film, thin electrode layers are located to excite the resonator. The shape of the top electrode is typically circular or rectangular. To analyze the resonator, an equivalent circuit model¹ (which is strictly one-dimensional), or a two-dimensional model² had been employed. The one-dimensional model did predict a series of principal resonance frequencies but it failed to account for spurious resonances which were observed experimentally. The mentioned two-dimensional model, however, showed only the qualitative structure of the spectrum.

A realistic three-dimensional model³ which we have developed clearly showed good quantitative agreement with an experimental measurement. To represent the composite nature of the resonator more accurately, we have developed our earlier model further so that it now consists of two different plates joined together. With minor adjustments, we were able to obtain essentially the same spectrum as in the single-plate model, which had checked well with experiment. This agreement implies greater significance for the model since each individual material property is directly related to the resonance frequencies.

In the following section, a brief description of the single plate model is furnished, and is extended to the development of a coupled-plate model. A comparison

* Supported in part by the U.S. government, and by the Office of Naval Research.

of its predictions with experimental data, and discussions will be given in section 3.

2 Three-dimensional modelA. Rectangular plate model

This model will best represent resonators with a rectangular top electrode. As in the disc model³ which we developed previously, the single homogeneous, isotropic elastic plate is assumed to have side lengths $2a$ and $2b$, and thickness $2L$. The origin of the coordinates is at the center of the plate. The displacement vector \vec{u} can be expressed⁴ in terms of a scalar potential ϕ and a partial displacement \vec{u}' :

$$\vec{u} = \vec{\nabla}\phi + \vec{u}' \quad (1)$$

with

$$\vec{\nabla} \cdot \vec{u}' = 0.$$

Considering only symmetric displacements with respect to the $z = 0$ plane, we have

$$u_x = -A \sin k_1 x \cos k_2 y \cos k_3 z + B_1 \sin q_1 x \cos q_2 y \cos q_3 z, \quad (2a)$$

$$u_y = -A(k_2/k_1) \cos k_1 x \sin k_2 y \cos k_3 z + B_2 \cos q_1 x \sin q_2 y \cos q_3 z, \quad (2b)$$

$$u_z = -A(k_3/k_1) \cos k_1 x \cos k_2 y \sin k_3 z + B_3 \cos q_1 x \cos q_2 y \sin q_3 z, \quad (2c)$$

where

$$k_3^2 = h^2 - (k_1^2 + k_2^2), \quad (3a)$$

$$q_3^2 = k^2 - (q_1^2 + q_2^2), \quad (3b)$$

$$h = \omega / c_L, \quad (3c)$$

$$k = \omega / c_T, \quad (3d)$$

c_L and c_T being longitudinal and transverse wave speeds in the plate.

To satisfy the stress-free condition on the $z = \pm L$ surfaces for all x and y , q_3 has to be equal to k_3 and q_2 equal to k_2 . Then we obtain the characteristic equation

3 Results and Discussion

$$\frac{\tan q_3 L}{\tan k_3 L} = - \frac{4q_3 k_3 (k_1^2 + k_2^2)}{(k_1^2 + k_2^2 - q_3^2)^2}. \quad (4)$$

Here k_1 and k_2 are determined from the boundary conditions on the $x = \pm a$ and $y = \pm b$ surfaces, respectively. Satisfying $\tau_{xx} = 0$ or $u_y = u_z = 0$ at $x = \pm a$ leads to

$$k_1 = \frac{2m+1}{2a} \pi; \quad (5a)$$

the $\tau_{xx} = \tau_{xy} = 0$ or $u_x = 0$ condition leads to

$$k_1 = \frac{n}{a} \pi. \quad (5b)$$

Similar equations for k_2 are obtained by satisfying boundary conditions on the $y = \pm b$ surfaces:

$$k_2 = \frac{2n+1}{2b} \pi \quad (6a)$$

or

$$k_2 = \frac{n}{b} \pi, \quad (6b)$$

where m and n are integers.

As the frequency equation (4) depends only on $k_1^2 + k_2^2$, the condition can be summarized in the following way.

For $\tau_{xx} = \tau_{yy} = 0$:

$$k_1^2 + k_2^2 = \left\{ \left(\frac{2m+1}{2a} \right)^2 + \left(\frac{2n+1}{2b} \right)^2 \right\} \pi^2, \quad (7a)$$

and for $\tau_{xx} = \tau_{xy} = \tau_{yz} = 0$:

$$k_1^2 + k_2^2 = \left\{ \left(\frac{n}{a} \right)^2 + \left(\frac{n}{b} \right)^2 \right\} \pi^2. \quad (7b)$$

B. Coupled-plate model

This model assumes two plates of different materials, joined together like the actual composite resonator. Coupled discs and coupled rectangular plates may represent composite resonators with a circular top electrode and with a rectangular electrode, respectively. The displacement vectors are similar in form as in single disc or rectangular plate model. However there is one more boundary surface to be considered. The displacements and stresses are continuous across the junction plane. This leads to 8×8 determinant equations for the disc model and 12×12 determinant equations for the plate model. The detailed equations will be published elsewhere⁵. The boundary conditions on the side surfaces are assumed to be the same as in the single-plate model.

As we had found before³, the boundary conditions on the side surfaces, Eqs. (7a) and (7b), have little effect on the resonance frequency spectrum, giving at most a few MHz ($\sim 1\%$) shifts. We thus employed the tangential stress-free condition, Eq. (7b), throughout our calculation.

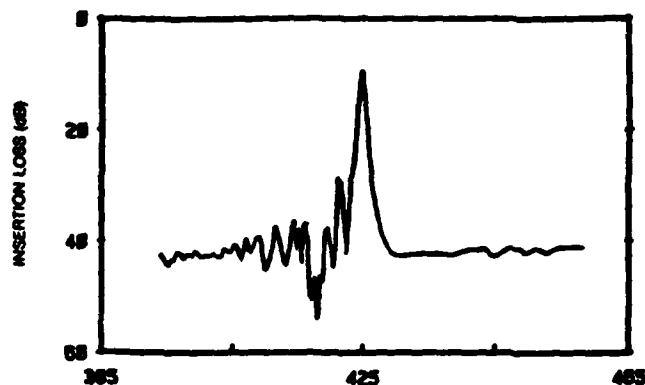


Fig. 1 Experimental frequency spectrum of bulk-acoustic wave resonator (filter).

Figure 1 shows the measured spectrum of a filter⁶ with silicon thickness of $5.6 \mu\text{m}$, zinc-oxide thickness of $2.47 \mu\text{m}$, and rectangular top electrodes ($254 \times 508 \mu\text{m}$) spaced $12 \mu\text{m}$ from each other. There appears a fundamental resonance at 425 MHz which is trailed toward the lower-frequency side by a series of smaller spurious resonances. Weak indications of a first overtone and its spurious frequencies are evident around 445 MHz; they appear far more distinctly in other, unpublished work of Grudkowski et al., Ref. 1. To model this spectrum, we first used a single rectangular plate model with the same area dimension as the electrode, and a thickness of $8.1 \mu\text{m}$. Effective velocities adopted were $\bar{c}_L = 7160$ m/sec and $\bar{c}_T = 3443$ m/sec, assuming $2\bar{c}_T < \bar{c}_L$. These values were obtained in the limiting case of Eq. (4) when the plate becomes infinite, and fixing the principal frequencies at 425 and 442 MHz. Different values of \bar{c}_L and \bar{c}_T are obtained if we assume $2\bar{c}_T > \bar{c}_L$, but the predicted spectra are the same.

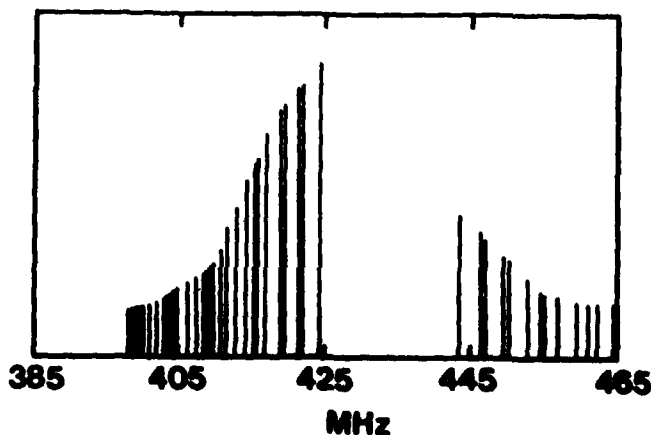


Fig. 2 Theoretical frequency spectrum from an effective single rectangular plate resonator. The dimension of the plate is $254 \times 508 \times 8.1 \mu\text{m}$.

Figure 2 is the spectrum of the rectangular plate model. The number of spurious resonance frequencies has increased considerably, comparing with that of the disc model³. However, the agreement with the data (Fig. 1) is better. Not only do the major peaks coincide with the predictions, but also the smaller wiggles on the lower frequency side are seen to agree. This was to be expected, since the rectangular plate should better represent the resonator with a rectangular electrode than the disc model.

Next, we improved the model further by joining two plates of different materials, just like the composite resonator itself. The C_T and C_L values for each material can be directly inserted for the appropriate plate with the relevant geometric dimensions. In this way the relationships between resonance frequencies and individual material properties can be explored. The plate thicknesses used in the model are a priori the same as those of the composite resonator. A value of $0.3 \mu\text{m}$, however, was added to ZnO thickness to simulate the effect of electrode thickness. The predicted resonance frequencies were somewhat higher than the observed ones. This may be due to the anisotropic property of materials used in the real resonator while the model was assumed to be isotropic. Reducing now C_T of ZnO and Si uniformly by 11.7%, we could get satisfactory agreement with the data.

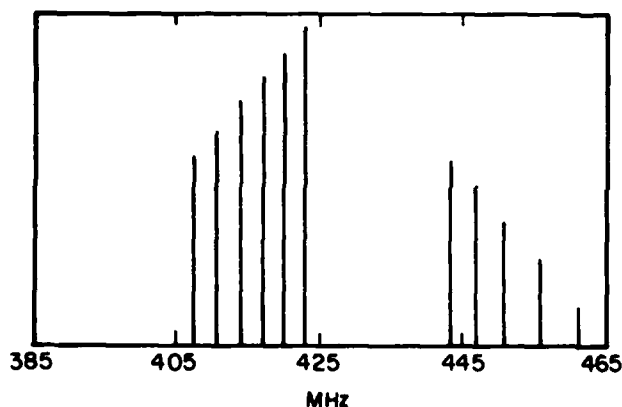


Fig. 3 Theoretical frequency spectrum from the coupled-disc model (radius = $250 \mu\text{m}$).

Figure 3 is the predicted spectrum of our coupled-disc model. This turns out to be almost identical with the results of the single disc model³.

In the mentioned fitting procedure, it was found that the fundamental frequency and its spurious frequencies are very sensitive to longitudinal velocity changes. On the other hand, the first overtone frequency and its spurious frequencies are strongly sensitive to the change of transverse velocities. This simply tells us that the effective $C_L < 2C_T$.

Physically this discovery reveals relationships important for the basic design of the composite resonator. If $C_L = 2C_T$, the fundamental and first overtone frequencies coincide, with their spurious frequencies being spread on both sides. If $C_L > 2C_T$ the opposite phenomenon (i.e. the fundamental frequency being sensitive to C_T and the first overtone frequency being sensitive to C_L) will occur.

Our coupled rectangular-plate model prediction is shown in Fig. 4.

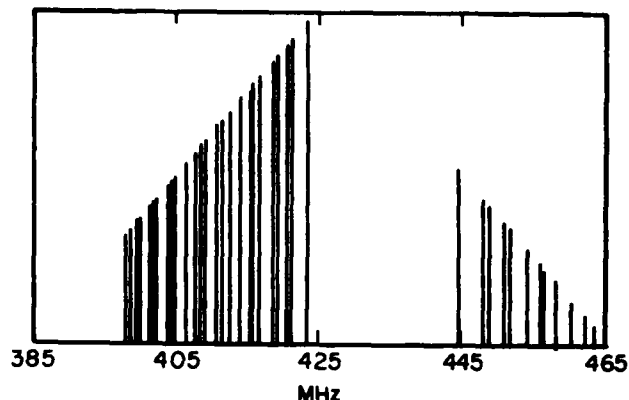


Fig. 4 Theoretical frequency spectrum from the coupled rectangular plate model ($254 \times 508 \mu\text{m}$).

The area of the rectangle is taken the same as that of the resonator, and the thicknesses were chosen in the same way as in the coupled-disc model. As we have seen in Fig. 2, a large number of spurious resonance frequencies are predicted. They are very similar to the case of the single plate model (Fig. 2). The agreement with the data is again excellent.

The effect of ZnO and Si thickness ratio variations on the values of C_L and C_T should be interesting, and will be investigated in the future since they have physical significance in controlling the spurious resonance frequencies. Also, the mode functions associated with the fundamental and the first overtone may be affected by the thickness ratio.

In conclusion, a series of models have been used to predict resonance frequencies and the agreements with experimental data are excellent. The coupled-plate model closely represents the resonator and the direct relationship of resonance frequencies with individual plate sound velocities and geometric configuration yield important information for controlling unwanted modes.

References

1. T. W. Grudkowski et al, Proc. IEEE 1980 Ultrasonics Symposium, p. 829; T. W. Grudkowski, J. F. Black, G. W. Drake, and D. E. Cullen, Proc. 36th Annual Symposium on Frequency Control, p. 537; J. S. Wang, Ph.D. thesis, Dept. of Electrical Engineering, University of Southern California (1981); K. M. Lakin and J. S. Wang, Proc. IEEE 1980 Ultrasonics Symposium p. 834.
2. R. F. Milsom, Proc. IEEE 1982 Ultrasonics Symposium p. 484.
3. K. B. Yoo, H. Überall and W. Williams, Jr., Proc. IEEE 1982 Ultrasonics Symposium p. 490.
4. R. W. Morse, J. Acoust. Soc. Am. **22**, 219 (1950).
5. K. B. Yoo, H. Überall, D. Ashrafi and S. Ashrafi, in preparation.
6. T. W. Grudkowski et al, Appl. Phys. Lett. **37**, 993, (1980).

ANALYSIS OF COMPOSITE RESONATOR GEOMETRIES

K. M. Lakin
Ames Laboratory-USDOE
Iowa State University
Ames, IA 50011

Summary

Composite resonators and other thin film piezoelectric acoustic devices are often composed of complex geometrical shapes which in turn may be made up of inhomogeneous material regions. Analytical solutions for the mode spectrum of such structures using perturbation theory or normal mode expansions must make approximations which may depart significantly from the real problem. Particularly, at microwave frequencies it may be inappropriate to neglect the finite thickness of the electrode metalization and treat it as an equivalent mass such as done for low frequency resonators. In composite resonators having p⁺ diffused silicon membranes the impurity diffusion profile and associated carrier distribution cause the material to be inhomogeneous.

In order to treat a general class of piezoelectric excitation problems a new finite difference formulation of the general two dimensional coupled piezoelectric wave equation has been derived. The formulation allows for three displacement components and the electric potential having propagation in two dimensions. Cases of obvious material symmetry reduce the complication of the formulation accordingly.

The formulation is unique in that the essential physical nature of the problem is not masked by the numerical analysis procedure. The finite difference equations are in the five point configuration having a central point and four nearest neighbor. The coupled difference equations describing the motion of the central point are in the form of the usual eigenvalue christoffel equations having nearest neighbor mesh points as apparent sources.

I Introduction

Composite resonators represent a new class of VHF-microwave piezoelectric devices that show considerable promise as fundamental mode resonators and as filter building blocks.¹ A major problem with such devices has been the complicated nature of the spurious responses exhibited by the resonators when employing c-axis normal films for wave generation. That film orientation generates a complex set of Lamb and flexure waves in the composite plate structure that is difficult to analyze analytically. However, some simple geometrical configurations have been analyzed using plate wave mode expansions to satisfy simplified boundary conditions.^{2,3}

It is the purpose of the study reported here to formulate a very general approach to solving complicated boundary value problems under conditions of arbitrary material anisotropy and boundary configurations. Typical boundary configurations that might be expected in composite resonators are shown in Fig 1. This figure is meant to represent a composite of possible boundary shapes, including abrupt termination of the piezoelectric layer, and is not suggesting a particular device design.

In addition to having complex boundaries the material may also be inhomogeneous as in the p⁺ silicon membranes. Because coupled resonator filters play such an important role in conventional filters it was felt that the analysis technique should be able to determine filter response, and network scattering parameters as well as mode distributions. The final result will be equivalent circuits to aid in device design.

Because of the complicated boundary shapes in real devices it was decided to use a numerical analysis formulation which would describe the material as being composed of a finite continuum with a superimposed mesh or grid where field values are to be calculated. For the general case of arbitrary crystal anisotropy and up to four field components, a finite difference equation approach was adopted to approximate the differential equations rather than a finite element formulation using a free energy minimization. The finite difference formulation allows the boundary conditions to be met in a clear physical manner just as they would be derived from the initial field equations.

Although one, two, and three dimensional formulations are under study, only the two dimensional formulation is discussed in detail in this paper.

II Theory

The theoretical development of the finite difference formalism starts from the general point form linear elastic equations for a region of arbitrary anisotropy,

NEWTON'S SECOND LAW

$$\frac{\partial}{\partial x_i} T_{ij} = -\omega^2 \rho u_j \quad (1)$$

GAUSS'S LAW

$$\frac{\partial}{\partial x_i} D_i = 0 \quad (2)$$

CONSTITUTIVE RELATIONS

$$T_{ij} = C_{ijkl} \frac{\partial u_k}{\partial x_l} + e_{ijk} \frac{\partial u_e}{\partial x_k} \quad (3)$$

$$D_i = e_{ijk} \frac{\partial u_k}{\partial x_l} - \epsilon_{ikl} \frac{\partial u_e}{\partial x_k} \quad (4)$$

For resonator analysis the time harmonic assumption is employed although there are numerical methods for solving the time dependent wave equation as well.⁴⁻⁶ Newton's second law in the point form and Gauss's Law form a set of four differential equation coupled through constitutive relations as indicated.

There are numerous methods used to solve this set of equations, in the presence of boundaries, using expansions in partial waves, plate waves, power series, or perturbation methods. Given the complexity of the boundary conditions anticipated with composite resonators a numerical finite difference approach has been chosen. Some simplification of the problem may be had by initially restricting the problem to two dimensions. The 2D plane will be the x₁, x₂

AD P002495

plane with x_3 normal. No spatial variations are allowed along x_3 and accordingly the linear differential equations for (1) may be written as,

$$\frac{\partial}{\partial x_1}(\tau_{ij}) - \frac{\partial}{\partial x_2}(-\tau_{ji}) = -\rho \omega^2 u_i \quad (5)$$

let

$$Q_1^{(i)} = \tau_{ij} \quad , \quad Q_2^{(i)} = -\tau_{ji} \quad (6)$$

where vector,

$$\vec{Q}^{(i)} = Q_1^{(i)} \hat{i} + Q_2^{(i)} \hat{j} \quad (7)$$

then,

$$\{\nabla \times \vec{Q}^{(i)}\}_3 = \{-\rho \omega^2 u_i\}_3 \quad (8)$$

Note that the L.H.S. of (1) has been expanded for $i=1,2$ and written so that the identification in (6-8) can be made. The result of having (6) in the form of a curl operation allows a direct integration using Stokes Theorem to reduce the L.H.S. by one order,

$$\int \nabla \times \vec{Q}^{(i)} \cdot d\vec{s} = \oint \vec{Q}^{(i)} \cdot d\vec{\ell} \quad (9)$$

let

$$d\vec{s} = \hat{k} dS = \hat{k} dx_1 dx_2$$

then,

$$\oint \vec{Q}^{(i)} \cdot d\vec{\ell} = -\omega^2 \int \rho u_i dx_1 dx_2 \quad (10)$$

The complex L.H.S. area integral is simply reduced to a line integral of the stress components around a closed path in the 1,2 plane. Of course, in order to do the integrals the fields must be known or assumed to have a given form, series expansion, etc. Physically these integrals correspond to finding the net force on the material region. One form of finite difference (FD) formulation is to break the region of interest into a set of finite points or mesh elements with each element identifying a small localized region of the material.

The rectangular grid region is shown in Fig 3 representing a small localized part of 1,2 plane. The grid spacings are generalized to h_1, h_2, h_3, h_4 as shown. The symbols ①, ②, ③, ④ denote respectively each of four material regions surrounding the central mesh point (I,J). For generality the four regions are assumed to have different material properties to allow eventual handling of complex material boundaries. The nearest neighbor field points to the central point (I,J) are (I-1,J), (I,J-1), (I+1,J), and (I,J+1). These mesh points are shaded for emphasis.

Looking ahead, such a rectangular grid will give a "five point" set of difference equations when fields along the dashed integration path are described by the central point and its nearest neighbors. A hexagonal mesh would be composed of triangular regions in a seven point difference equation. The triangular mesh allows a more accurate description of non-rectangular boundaries and can be distorted to represent rectangular boundaries as well. As will be apparent, the basic matrix

form of the difference equations does not depend on the detailed nature of the mesh.

The R.H.S. area integral of (10) is the simplest to carry out when the fields within the area element are assumed to be given approximately by the fields at (I,J). The result is,

$$-\omega^2 \left[\rho \frac{h_1 h_2}{4} + \rho \frac{h_2 h_3}{4} + \rho \frac{h_3 h_4}{4} + \rho \frac{h_4 h_1}{4} \right] u_i \quad (11)$$

where the term in square brackets is just the total mass (per unit depth). Note that (2) does not have a corresponding R.H.S. and so that area integral is zero.

The L.H.S. integrals in (10) are somewhat more complicated because the stress must be described in terms of spatial derivatives of u_k along each path. The line integrals are of the form

$$\oint_{1,2} = \int_a^b Q_2^{(i)} dx_2 + \int_b^c Q_1^{(i)} dx_1 + \dots$$

where the integration limits denote the various line segments in the four material regions Fig 3. Note that L.H.S. of (2) is treated the same way as L.H.S. (1) so that the vector $\vec{Q}^{(j)}$ takes on a more general form

$$Q_1^{(i)} = -A_{ik} \frac{\partial u_k}{\partial x_1} - B_{ik} \frac{\partial u_k}{\partial x_2} \quad j,k=1,2,3,4$$

$$Q_2^{(i)} = C_{jk} \frac{\partial u_k}{\partial x_1} + D_{jk} \frac{\partial u_k}{\partial x_2} \quad j,k=1,2,3,4$$

where u_4 is the electrical potential. The material tensors A,B,C,D are defined by

A _{jk}	k=1	2	3	4
j=1	C ₁₆	C ₆₆	C ₅₆	e ₁₆
2	C ₁₂	C ₂₆	C ₂₅	e ₁₂
3	C ₁₄	C ₄₆	C ₄₅	e ₁₄
4	e ₁₂	e ₂₆	e ₂₅	-e ₂₁

B _{jk}	k=1	2	3	4
j=1	C ₆₆	C ₂₆	C ₄₆	e ₂₆
2	C ₂₆	C ₂₂	C ₂₄	e ₂₂
3	C ₄₆	C ₂₄	C ₄₄	e ₂₄
4	e ₂₆	e ₂₂	e ₂₄	-e ₂₂

C _{jk}	k=1	2	3	4
j=1	C ₁₁	C ₁₆	C ₁₅	e ₁₁
2	C ₁₆	C ₆₆	C ₅₆	e ₁₆
3	C ₁₅	C ₅₆	C ₅₅	e ₁₅
4	e ₁₁	e ₁₆	e ₁₅	-e ₁₁

D _{jk}	k=1	2	3	4
j=1	c ₁₆	c ₁₂	c ₁₄	e ₁₂
2	c ₆₆	c ₂₆	c ₄₆	e ₂₆
3	c ₅₆	c ₂₅	c ₄₅	e ₂₅
4	e ₁₆	e ₁₂	e ₁₄	-e ₁₂

Each tensor is given in terms of the usual reduced subscript material tensors and are partitioned by the dotted lines to identify those regions associated with Laplace's equation only, the wave equation without piezoelectric coupling, and finally the piezoelectric coupling coefficients. Symmetry conditions will obviously reduce the complexity of these tensors.

The line integrals are actually composed of 16 terms for each of 4 equations in 4 regions for a total of 256 integrals. Fortunately, the integrals are of only two different forms, as may be seen by dropping most subscripts and superscripts,

$$Q_1 = -A \frac{\partial u}{\partial x_1} - B \frac{\partial u}{\partial x_2}$$

$$Q_2 = C \frac{\partial u}{\partial x_1} + D \frac{\partial u}{\partial x_2}$$

In order to carry out the integrals the fields must be known or assumed along a given path. Terms having the A and D coefficients integrate directly while the B, C terms require that the derivatives be described in terms of the local fields. For example, in the path a-b, Fig 3, the derivative is approximated by its finite difference,

$$\frac{\partial u}{\partial x_1} = \frac{u(I, J+1) - u(I, J)}{h_1}$$

while the endpoints are approximated by

$$u(a) = (u(I, J+1) + u(I, J))/2$$

$$u(b) = (u(I, J-1) + u(I-1, J))/2.$$

When the various line segments are integrated and summed, while identifying material tensors with an appropriate region superscript, the resulting set of difference equations is

$$\begin{bmatrix} \alpha_{11} - Q_0 & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} - Q_0 & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} - Q_0 & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \end{bmatrix} \begin{bmatrix} u_1(I, J) \\ u_2(I, J) \\ u_3(I, J) \\ u_4(I, J) \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix}$$

where $\alpha_{jk} = \alpha_{jk}^{(1)} + \alpha_{jk}^{(2)} + \alpha_{jk}^{(3)} + \alpha_{jk}^{(4)}$

$$\alpha_{jk}^{(1)} = \frac{C_{ijk}^{(1)} h_2 + C_{ijk}^{(1)} h_4}{2 h_1} + \frac{A_{ijk}^{(1)} - A_{ijk}^{(2)}}{2}$$

$$\alpha_{jk}^{(2)} = \frac{B_{ijk}^{(2)} h_1 + B_{ijk}^{(2)} h_3}{2 h_2} + \frac{D_{ijk}^{(2)} - D_{ijk}^{(3)}}{2}$$

$$\alpha_{jk}^{(3)} = \frac{C_{ijk}^{(3)} h_4 + C_{ijk}^{(3)} h_2}{2 h_3} + \frac{A_{ijk}^{(3)} - A_{ijk}^{(4)}}{2}$$

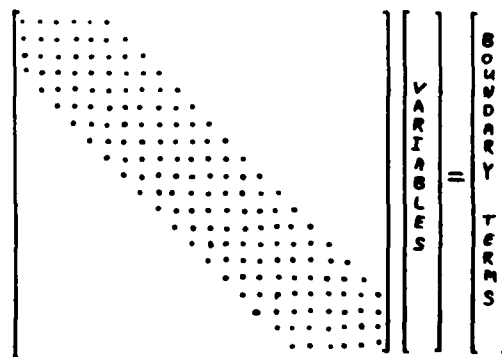
$$\alpha_{jk}^{(4)} = \frac{B_{ijk}^{(4)} h_3 + B_{ijk}^{(4)} h_1}{2 h_4} + \frac{D_{ijk}^{(4)} - D_{ijk}^{(1)}}{2}$$

$$Q_0 = -\omega^2 [\rho^{(1)} h_1 h_4 + \rho^{(2)} h_2 h_3 + \rho^{(3)} h_3 h_4 + \rho^{(4)} h_4 h_1] \frac{1}{4}$$

$$U_j = \alpha_{jk}^{(1)} U_k(I, J+1) + \alpha_{jk}^{(2)} U_k(I-1, J) + \alpha_{jk}^{(3)} U_k(I, J-1) + \alpha_{jk}^{(4)} U_k(I+1, J)$$

The above matrix form for the finite difference equations holds at each point in the mesh. Physically, the fields at a given mesh point are viewed as being driven by the nearest neighbor. This matrix description has the expected form of a driven vibration problem. The dashed lines in the matrix separate those parts coupled by piezoelectricity. Simplifying examples will be given later.

Within the mesh any point also influences its neighbors and so the vector on the R.H.S. of (11) is not a constant except for some terms at driven excitation points. Consequently, for interior points (11) must be written as a homogeneous set of equations, one set for each point in the mesh. There will then be as many coupled equations as there are unknown field quantities. Such a set typically forms the square but sparsely populated matrix having non-zero coefficients clustered along the diagonal,



The dots represent non-zero matrix elements.

There is an art to labeling the mesh with a sequence of indices such that the matrix has a small "bandwidth" (clustered tightly about the diagonal). Solution of the equations is efficiently carried out by Gaussian elimination followed by back substitution or other numerical methods.

A very simple example is for a one dim-

ensional resonator whose finite difference matrix is of the form

U(1)	V(2)	U(2)	V(3)	U(3)	V(4)	U(4)	V(5)	U(5)
a	b	1	0	0	0	0	0	0
0	-2	R	1	0	0	0	0	0
1	0	-A	0	1	0	0	0	0
0	1	0	-2	R	1	0	0	0
0	0	1	0	-A	0	1	0	0
0	0	0	1	0	-2	R	1	0
0	0	0	0	1	0	-A	0	1

where a, b, A and R are non-zero constants and U and V are the displacement and electric potential respectively. Only the upper left corner of the matrix is shown. Note that V(1) is the known boundary potential, and so would appear on the R.H.S. of the equations. The remaining unknowns are taken in pairs corresponding to each mesh point. The matrix is relatively well banded and only the non-zero elements need be stored. The solution actually is carried out by performing the Gaussian elimination (matrix all zero below the diagonal) as the matrix is built. The matrix remains zero in the upper diagonal region above the diagonal row of ones. An inappropriate labeling or ordering of the equations could result in a much larger matrix that might be too large to store in computer memory.

In Fig 4 is shown a case where a one dimensional resonator was analyzed for a frequency near the fifth overtone using 51 mesh points (50 segments of material). The sinusoidal variation of the displacement is clearly brought out by the fine mesh. Using only 11 mesh points the finite difference equations are not able to resolve the high spatial frequencies of the overtone resonance, Fig 5.

The two dimension problem proved to be too large for the desk calculator that carried out the 1D analysis results shown in Fig 4 and 5. Consequently the entire problem is being moved to a computer capable of handling the larger and complex variables matrix. Results for the two dimensional case will be reported later.

Within the bulk of a homogeneous material the difference equations greatly simplify. For example, in the case of 6mm symmetry having c-axis perpendicular to the 1,2 plane (horizontal shear) and for a square grid of spacing h, the piezoelectrically coupled difference equations reduce to

$$u_3(I, J) = \frac{u_3(I, J-1) + u_3(I, J+1) + u_3(I-1, J) + u_3(I+1, J)}{4 - (kh)^2}$$

$$u_4(I, J) = \frac{u_4(I, J-1) + u_4(I, J+1) + u_4(I-1, J) + u_4(I+1, J)}{4} + \frac{e}{\epsilon} \left(\frac{kh}{2} \right)^2 u_3(I, J)$$

where $kh = 2\pi \frac{h}{\lambda}$

The case for zero frequency, $k = 0$, reduces to the simple difference equations for Laplace's equation for both the electrostatic and mechanically static case. The grid size is normalized by the acoustic wavelength and may be chosen for as few as 20 cells per wavelength. The finite difference equations for boundary terms are more complicated and must be derived for each specific case.

III Summary

A finite difference formulation of the general anisotropic piezoelectric coupled wave equations has been derived for the analysis of complex composite resonator boundary value and excitation problems. The difference equations form a coupled set of four equations describing fields at each element in the mesh. The physical form of the equations show that a central mesh point is driven by its nearest neighbors as would be expected from physical considerations. When expanded the difference equations can be written in the form of sparse banded set of equations whose solution is readily obtained by well known means.

The formulation is much more general than that required by the resonator boundary value problem. The formulation and solution technique should be applicable to a wide class of excitation and acoustic scattering problems in arbitrarily anisotropic and inhomogeneous materials with complex boundaries.

A three dimensional version of the formulation is under investigation. Although mathematically more complicated, the 3D formulation follows the derivation of the linear elastic equations from classical mechanics and accordingly the physics is somewhat clearer than in the two dimensional case.

IV Acknowledgments

The author wishes to acknowledge the support of the Air Force Office of Scientific Research and the encouragement of numerous colleagues to initiate this task.

References

1. "Thin Film Resonator and Filters" K.M.Lakin J.S.Wang, G.R.Kline, A.R.Landin, Y.Y.Chen, and J.D.Hunt, 1982 Ultrasonics Symposium Proceedings p. 466.
2. "Two-dimensional Theory of Thin -film ZnO Resonators on Silicon" R.F.Milsom, 1982 Ultrasonics Symposium Proceedings, p. 484.
3. "Spurious Resonances in Bulk Acoustic Wave Resonators" K.B.Yoo and H.Uberall, 1982 Ultrasonics Symposium Proceedings, p. 490.
4. The Analysis of Eddy Currents R.L.Stoll Clarendon Press, Oxford, 1974 ISBN 0-19859311-2.
5. Numerical Solution of Partial Differential Equations G.D.Smith, Clarendon Press Oxford, 1978, ISBN 0-19-859625-1.

6. The Finite Difference Method in Partial Differential Equations A.R. Mitchell and D.F. Griffiths, John Wiley & Sons, Ltd. 1980, ISBN 0-471-27641-3

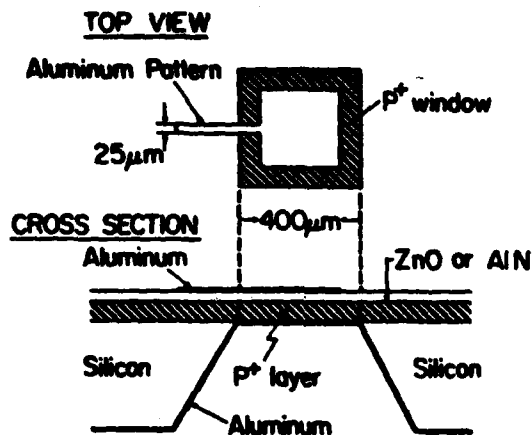


Fig 1. General configuration of a composite resonator using a silicon p⁺ membrane. Other configurations might be composed of other material combinations.

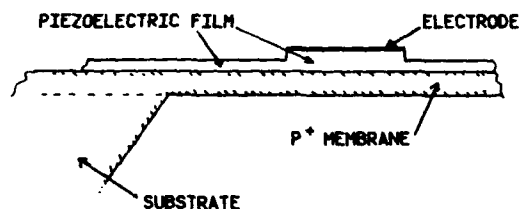


Fig 2. Illustration of some geometrical boundary configurations. Slanting and abrupt boundaries are of interest as well as in homogeneous material properties other than abrupt junctions.

DISCRETE GRID IN THE 1,2 PLANE,

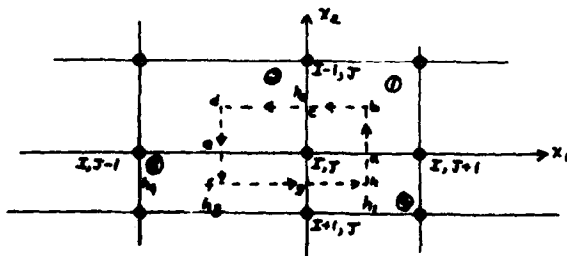


Fig 3. A rectangular sampling grid in the 1,2 plane showing the method of labeling mesh points and material regions. The line integral is carried out along the a-b-c-d-e-f-g-h-a path and the area integral within the region defined by the line. The encircled numbers denote the four material regions.

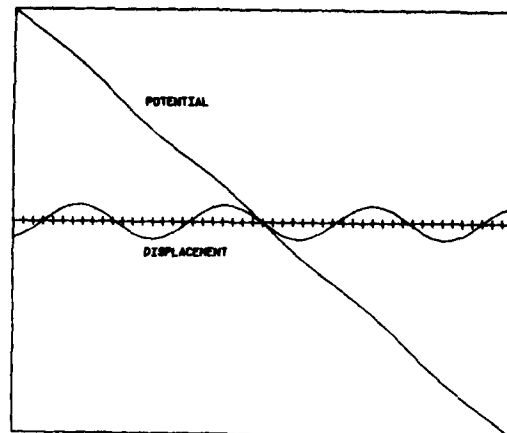


Fig 4. Numerical result for a one dimensional resonator operated near fifth overtone. The 51 point sampling grid adequately resolves the field.

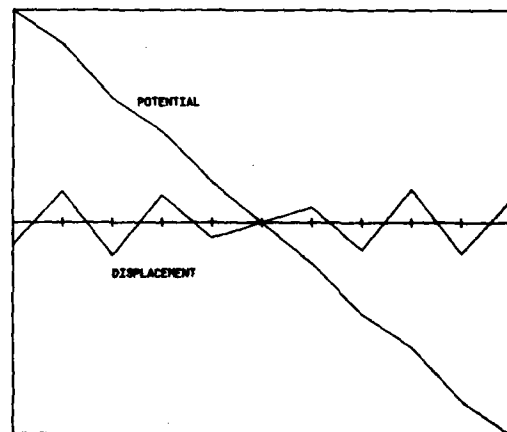


Fig 5. Some conditions as in Fig 4 except only eleven sampling points are used. Although suitable for the fundamental resonance, there are too few points for the high spatial frequencies of the overtone.

AD P 002496

AN ANALYSIS OF THICKNESS-EXTENSIONAL TRAPPED ENERGY RESONATORS WITH RECTANGULAR ELECTRODES IN THE ZINC-OXIDE THIN FILM ON SILICON CONFIGURATION

H.F. Tiersten and D.S. Stevens
Department of Mechanical Engineering,
Aeronautical Engineering & Mechanics
Rensselaer Polytechnic Institute
Troy, New York 12181

Abstract

It has recently been shown that highly uniform thin layers can be etched in a small well-defined region of a silicon wafer and that good quality thin piezoelectric films such as zinc-oxide can be deposited along with the electrodes to form high frequency trapped energy resonant device structures. The accompanying analytical work has been for pure thickness vibrations only. In this work an analysis of essentially thickness-extensional trapped energy modes in the thin piezoelectric film on silicon composite structure is performed. It is shown that for small wavenumbers along the plate the dispersion equation is isotropic in the plane of the plate even though the silicon is severely anisotropic in that plane. From the resulting dispersion relation an asymptotic differential equation describing the mode shape along the surface of the composite plate vibrating in the vicinity of a thickness-extensional resonance is obtained along with the associated edge conditions. Since the mode is essentially thickness-extensional, trapping does not usually occur in the electroded region for the fundamental mode of the flat composite plate. However, by appropriately thickening the silicon outside the electroded region trapping in the electroded region can be made to occur for the fundamental mode. Furthermore, in the case of zinc-oxide on silicon, trapping in the electroded region can be made to occur for the fundamental mode of a flat plate simply by making the zinc-oxide sufficiently thicker than the etched silicon. The aforementioned asymptotic differential equation and edge conditions are applied in the analysis of the steady-state vibrations of trapped energy resonators with rectangular electrodes and a lumped parameter representation of the admittance, which is valid in the vicinity of a resonance, is obtained. The analysis holds for all thickness-extensional modes and their accompanying lateral overtones.

1. Introduction

Recently, it has been shown¹⁻⁵ that highly uniform thin layers can be etched in a small well-defined region of a silicon wafer and that good quality thin piezoelectric films such as zinc-oxide can be deposited along with metal electrodes to form a high frequency resonant structure. Both single resonators and coupled resonator structures have been fabricated in this way and reasonably high Q 's have been obtained^{1,2} for essentially thickness-extensional modes of operation. As noted by others^{1,2}, this enables the inclusion of resonators and filters as part of the integrated circuit on the same silicon chip. The analysis that has accompanied the experimental work has used equivalent circuits, which hold for pure thickness modes only and ignore the transverse behavior of the modes of motion. However, the transverse behavior is very important for calculating the precise frequency spectra, the motional capacitance and, perhaps most importantly, the limitation on Q resulting from radiation into the thick portion of the silicon wafer.

In this work an analysis of essentially thickness-extensional trapped energy modes in the thin piezo-

electric film on silicon structure is performed. The analysis holds only when trapping is present⁶, which results in modes that have the highest possible Q 's because virtually no energy is radiated into the thick portion of the silicon. Although considerably more complicated than the earlier work because of the composite nature of the plate structure, this work parallels earlier work⁷⁻⁹ on AT-cut quartz trapped energy resonators in that the assumption of thickness dependence of all electrical variables is employed in obtaining the asymptotic expression for the fundamental and overtone essentially thickness-extensional dispersion relations in the vicinity of the cutoff frequencies, i.e., for small wave- and decay numbers along the plate. This assumption holds for small piezoelectric coupling. As in the earlier work⁷⁻⁹, simple approximate edge conditions at each junction between an electroded and unelectroded region of the plate are established. This work differs from the earlier work⁷⁻⁹ in that in the earlier work the motion is essentially thickness-shear and here it is essentially thickness-extension. Although the fundamental essentially thickness-shear mode always traps in the electroded regions of a flat plate, the fundamental essentially thickness-extensional mode usually does not trap¹⁰ in the electroded regions of a flat plate (composite or otherwise). However, by making the thin silicon layer appropriately thicker outside the electroded region (in order to make the pure thickness frequency in the unelectroded region sufficiently lower than that in the electroded region) trapping of the fundamental essentially thickness-extensional mode in the electroded regions will be caused in the usual case when it does not exist for the flat plate. Furthermore, in the case of the zinc-oxide film it is shown that for sufficiently thick zinc-oxide relative to the silicon layer trapping in the electroded regions will occur for the flat composite plate.

After the pure thickness solutions for both the unelectroded and electroded composite plates are obtained, the full solution for straight-crested waves decaying along the (010) axis of the silicon is obtained for the unelectroded composite plate. This case is treated because it is more convenient than propagating waves in the electroded plate and it has been shown in a simpler case^{7,8} that for small wavenumbers along the plate the solution for propagating waves in the electroded plate can readily be constructed from the solution for the case treated. An asymptotic expression for the fundamental and overtone essentially thickness-extensional dispersion relations valid in the vicinity of the thickness-extensional cutoff frequencies is determined from an expansion of the full solution. In addition, it is shown that the asymptotic dispersion relation is isotropic in the plane of the composite plate. As a consequence, the asymptotic dispersion relation for straight-crested waves is readily extended to waves having either trigonometric or exponential crests, as in certain earlier work¹¹. The approximate edge conditions to be satisfied at each junction between an electroded and unelectroded portion of the composite plate are obtained. As in earlier work^{7,8} the solutions obtained here for the trapped energy modes in the composite plate with rectangular electrodes satisfy the equations to second order in the small

wavenumbers along the plate in all regions except the unimportant corner type region, in which they are satisfied to first order.

The aforementioned asymptotic dispersion relations, approximate edge conditions and approach taken in corner type regions are employed in the analysis of trapped energy resonators. Lumped parameter representations of the admittance of the trapped energy resonator, which are valid in the vicinity of the resonances, are determined.

2. Basic Equations

The differential equations of motion and electrostatics and linear piezoelectric constitutive equations for the piezoelectric film may be written in the tensor form^{12,13}

$$T_{ij,i}^f = \rho^f \ddot{u}_j^f, \quad D_{i,i}^f = 0, \quad (2.1)$$

$$T_{ij}^f = c_{ijkl}^f u_{k,l}^f + e_{kij}^f \varphi_{,k}^f, \\ D_i^f = e_{ikl}^f u_{k,l}^f - \epsilon_{ik}^f \varphi_{,k}^f, \quad (2.2)$$

where the notation is conventional. The equations of motion and the linear elastic constitutive equations for the silicon may be written in the tensor form^{14,15}

$$T_{ij,i}^s = \rho^s \ddot{u}_j^s, \quad T_{ij}^s = c_{ijkl}^s u_{k,l}^s, \quad (2.3)$$

and we note that the equations are too cumbersome to write out in detail. At this point we introduce a Cartesian coordinate system x_1, x_2, x_3 with the x_3 -axis normal to the major surfaces of the zinc-oxide film and along a cube axis of the silicon. Since zinc-oxide has small piezoelectric coupling, for small wave and decay numbers along the plate we need retain only x_3 -dependence of all electrical variables and we have

$$D_{3,3}^f = 0, \quad D_3^f = e_{3kl}^f u_{k,l}^f - \epsilon_{33}^f \varphi_{,3}^f. \quad (2.4)$$

Figure 1 shows a schematic diagram of a cross section of a thin zinc-oxide film on a thin silicon layer composite trapped energy resonator. In the unelectroded regions the boundary conditions on the major surfaces are

$$T_{3j}^f = 0, \quad D_3 = 0 \quad \text{at } x_3 = h^f, \quad (2.5)$$

$$T_{3j}^f - T_{3j}^s = h^s \rho^s \ddot{u}_j^s, \quad u_j^f = u_j^s, \quad \varphi^f = 0 \quad \text{at } x_3 = 0, \quad (2.6)$$

$$T_{3j}^s = 0 \quad \text{at } x_3 = -h^s, \quad (2.7)$$

where the superscripts f and s stand for the thin film and silicon layer, respectively, and ρ^s is the mass density of the ground electrode. The electrical condition in (2.5)₂ is a consequence of the fact that the x_1 - and x_2 -dependence of all electrical variables has been left out of account and the electrical potential in space is bounded at $x_3 = \infty$. Since the electrodes are much thinner than either the thin film or the layer, we have employed approximate thin plate equations¹⁵ for the electrode plating in (2.6)₁, and we have made use of the fact that the mechanical stiffness of the very thin electrode plating is negligible for small wavenumbers along the plate. Similarly, on the major surfaces of the composite plate in the electroded region the boundary conditions are

$$T_{3j}^f = -\rho^f h^f \ddot{u}_j^f, \quad \varphi^f = V_0 e^{i\omega t} \quad \text{at } x_3 = h^f, \quad (2.8)$$

$$\varphi^f = 0 \quad \text{at } x_3 = 0, \quad (2.9)$$

along with (2.6) and (2.7), which remain unchanged. The quantity ρ^f is the mass density of the upper electrode.

3. Pure Thickness-Extensional Vibrations

In order to obtain the asymptotic dispersion relations in the vicinity of the thickness-extensional resonant frequencies, we must first obtain the solutions for the pure thickness-extensional vibrations of the composite plate. Accordingly, in this section we briefly present the solutions to the problems of the pure thickness-extensional vibrations of the composite plate without an upper electrode but with a very thin ground electrode, as shown in Figure 2, and of the completely electroded composite plate driven by a voltage, as shown in Figure 3.

For the composite plate without an upper driving electrode we have¹³

$$\varphi^f = (e_{33}^f / \epsilon_{33}^f) [u_3^f - u_3^f(0, t)], \quad (3.1)$$

along with

$$u_3^f = (A^f \cos \eta_f x_3 + B^f \sin \eta_f x_3) e^{i\omega t}, \\ u_3^s = (A^s \cos \eta_s x_3 + B^s \sin \eta_s x_3) e^{i\omega t}, \quad (3.2)$$

provided

$$\bar{\epsilon}_{33}^f \eta_f^2 = \rho^f \omega^2, \quad c_{33}^s \eta_s^2 = \rho^s \omega^2, \quad (3.3)$$

where

$$\bar{\epsilon}_{33}^f = c_{33}^f + (e_{33}^f)^2 / \epsilon_{33}^f. \quad (3.4)$$

From the boundary conditions (2.5) - (2.7) we obtain the transcendental equation¹³

$$\tan \eta_f h^f + c^r \mu \tan \mu \eta_f h^f + R^r \eta_f h^f = 0, \quad (3.5)$$

where

$$c^r = \frac{c_{33}^s}{\bar{\epsilon}_{33}^f}, \quad \mu = \sqrt{\frac{\bar{\epsilon}_{33}^f \rho^s}{c_{33}^s \rho^f}}, \quad \sigma = \frac{h^s}{h^f}, \quad R^r = \frac{\rho^s h^s}{\rho^f h^f}. \quad (3.6)$$

and

$$\eta_s = \mu \eta_f. \quad (3.7)$$

The roots $\eta_f h^f$ of Eq. (3.5) determine the eigenfrequencies ω of pure thickness-extensional vibrations of the composite plate without a driving electrode by means of (3.3)₁. Since $R^r \ll 1$, it is convenient to expand the roots $\eta_f h^f$ of (3.5) about the roots $\eta_f^0 h^f$ of (3.5) with $R^r = 0$, i.e., without a ground electrode. Accordingly, we write

$$\eta_f h^f = \eta_f^0 h^f + \Delta^f, \quad (3.8)$$

where

$$\tan \eta_f^0 h^f + c^r \mu \tan \mu \eta_f^0 h^f = 0, \quad (3.9)$$

which is the transcendental equation without a ground electrode that yields roots $\eta_f^0 h^f$, the lowest few of which are of primary interest to use. Substituting from (3.8) into (3.5), expanding and retaining terms linear in Δ^f , we obtain

$$\Delta^f = -R'' \eta_f^0 h^f / G^0, \quad (3.10)$$

where

$$G^0 = \frac{1}{\cos \eta_f^0 h^f} + \frac{c^r \mu^2}{\cos \mu \eta_f^0 h^f}. \quad (3.11)$$

The further substitution of (3.8) and (3.10) into (3.3), yields

$$\omega_e = \sqrt{\frac{\epsilon_{33}^f}{\rho^f}} \eta_f^0 \left(1 - \frac{R^f}{G^0}\right), \quad (3.12)$$

which gives the eigenfrequencies of the composite plate without a driving electrode but with a ground electrode in terms of the roots of (3.9), which is without a ground electrode.

In the case of the completely electroded plate driven by a voltage we have¹³

$$\begin{aligned} \bar{u}_3^f &= u_3^f - (e_{33}^f v x_3 / c_{33}^f h^f) e^{i\omega t}, \\ \varphi^f &= \frac{e_{33}^f}{\epsilon_{33}^f} u_3^f + \left[c x_3 + K + \frac{v x_3}{h^f} \right] e^{i\omega t}, \end{aligned} \quad (3.13)$$

where u_3^f is given in (3.2), and

$$k^2 = (\epsilon_{33}^f)^2 / \epsilon_{33}^f c_{33}^f = \hat{k}^2 / (1 + \hat{k}^2), \quad R^f = \rho^f h^f / \rho^f h^f. \quad (3.14)$$

Again it is convenient to expand the roots $\bar{\eta}_f h^f$ of (2.6) - (2.9) about the roots $\eta_f^0 h^f$ of (3.9). Accordingly, we write

$$\bar{\eta}_f h^f = \eta_f^0 h^f + \bar{\Delta}^f, \quad (3.15)$$

and obtain¹³

$$\bar{\Delta}^f = \eta_f^0 h^f P^0 / G^0, \quad (3.16)$$

where

$$\begin{aligned} P^0 &= \frac{k^2}{(\eta_f^0 h^f)^2} \left(\frac{2}{\cos \eta_f^0 h^f} - 2 + c^r \mu \tan \mu \eta_f^0 h^f \tan \eta_f^0 h^f \right) \\ &\quad - R^f - R^f (1 - c^r \mu \tan \eta_f^0 h^f \tan \mu \eta_f^0 h^f). \end{aligned} \quad (3.17)$$

The further substitution of (3.15) with (3.16) into (3.3), yields

$$\bar{\omega}_e = \sqrt{\frac{\epsilon_{33}^f}{\rho^f}} \eta_f^0 \left(1 + \frac{P^0}{G^0}\right), \quad (3.18)$$

which gives the eigenfrequencies of the fully electroded composite plate with shorted electrodes in terms of the roots of (3.9) for the unelectroded composite plate.

The difference between the resonant frequencies of pure thickness-extensional vibrations of the composite plate with and without a driving electrode, i.e., $\omega_e - \bar{\omega}_e$ for the same η_f^0 , is very important for describing and understanding the behavior of trapped energy resonators in the thin film-on-silicon structure. In fact, the trapped energy resonator can operate at a frequency ω essentially in the narrow frequency range between ω_e and $\bar{\omega}_e$.

4. Straight-Crested Dispersion Relations

In this section we obtain the straight-crested dispersion relations for the composite plate in the vicinity of the thickness-extensional resonant frequencies determined in the preceding section, i.e., for small wavenumbers along the plate. We specifically treat decaying waves in the unelectroded composite plate because these are simpler (less cumbersome) to treat than propagating waves in the electroded structure and the dispersion relations for propagating waves in the electroded composite plate can readily be obtained from the dispersion relations for decaying waves in the unelectroded case with the aid of certain results obtained for the case of pure thickness-extensional vibrations of the composite plate in Sec. 3.

As a solution of the differential equations¹³ for straight-crested waves, we take

$$\begin{aligned} u_3^f &= (A_3^f \cos \eta_f x_3 + B_3^f \sin \eta_f x_3) e^{-\xi x_1} e^{i\omega t}, \\ u_1^f &= (A_1^f \sin \eta_f x_3 + B_1^f \cos \eta_f x_3) e^{-\xi x_1} e^{i\omega t}, \end{aligned} \quad (4.1)$$

$$\begin{aligned} u_3^s &= (A_3^s \cos \eta_s x_3 + B_3^s \sin \eta_s x_3) e^{-\xi x_1} e^{i\omega t}, \\ u_1^s &= (A_1^s \sin \eta_s x_3 + B_1^s \cos \eta_s x_3) e^{-\xi x_1} e^{i\omega t}, \end{aligned} \quad (4.2)$$

the first, (4.1), of which satisfies the differential equations¹³ for the zinc-oxide provided

$$\begin{aligned} \sigma_{11}^f A_1^f + \sigma_{13}^f A_3^f &= 0, \quad -\sigma_{13}^f A_1^f + \sigma_{33}^f A_3^f = 0, \\ \sigma_{11}^f B_1^f - \sigma_{13}^f B_3^f &= 0, \quad \sigma_{13}^f B_1^f + \sigma_{33}^f B_3^f = 0, \end{aligned} \quad (4.3)$$

where

$$\begin{aligned} \sigma_{11}^f &= (c_{11}^f \xi^2 - c_{44}^f \eta_f^2 + \rho^f \omega^2), \quad \sigma_{13}^f = (c_{13}^f + c_{44}^f) \xi \eta_f, \\ \sigma_{33}^f &= (c_{44}^f \xi^2 - \epsilon_{33}^f \eta_f^2 + \rho^f \omega^2), \end{aligned} \quad (4.4)$$

and the latter, (4.2), of which satisfies the differential equations¹³ for the silicon provided

$$\begin{aligned} \sigma_{11}^s A_1^s + \sigma_{13}^s A_3^s &= 0, \quad -\sigma_{13}^s A_1^s + \sigma_{33}^s A_3^s = 0, \\ \sigma_{11}^s B_1^s - \sigma_{13}^s B_3^s &= 0, \quad \sigma_{13}^s B_1^s + \sigma_{33}^s B_3^s = 0, \end{aligned} \quad (4.5)$$

where

$$\begin{aligned} \sigma_{11}^s &= (c_{11}^s \xi^2 - c_{44}^s \eta_s^2 + \rho^s \omega^2), \quad \sigma_{13}^s = (c_{13}^s + c_{44}^s) \xi \eta_s, \\ \sigma_{33}^s &= (c_{44}^s \xi^2 - c_{33}^s \eta_s^2 + \rho^s \omega^2). \end{aligned} \quad (4.6)$$

Each of these four systems, (4.3)₁₋₂ and (4.5)₁₋₂, of two linear, homogeneous algebraic equations in two amplitudes yields nontrivial solutions when the determinant of the coefficients of the amplitudes vanishes. Both determinants for the film are identical, as are both determinants for the layer. Each of the two independent determinants is quadratic in ξ^2 , η^2 and ω^2 . Hence, for a given ξ and ω , each determinant yields two independent η ($\eta^{(1)}$, $\eta^{(2)}$) and $\eta^{(1)}$, $\eta^{(2)}$ and each $\eta^{(1)}$ yields independent amplitude ratios from either of the two equations leading to each of the four determinants. Let us denote the eight sets of amplitude ratios by

$$A_1^{fi} = \mu^{fi} A_3^{fi}, \quad B_1^{fi} = \nu^{fi} B_3^{fi}, \quad A_1^{si} = \mu^{si} A_3^{si}, \quad B_1^{si} = \nu^{si} B_3^{si}, \quad i=1,2. \quad (4.7)$$

As a solution of the boundary conditions (2.5) - (2.7), we take

$$\begin{aligned} u_3^f &= (A_3^{f1} \cos \eta_{f1} x_3 + A_3^{f2} \cos \eta_{f2} x_3 + B_3^{f1} \sin \eta_{f1} x_3 + B_3^{f2} \sin \eta_{f2} x_3) e^{-\xi x_1} e^{i\omega t}, \\ u_1^f &= (A_1^{f1} \sin \eta_{f1} x_3 + A_1^{f2} \sin \eta_{f2} x_3 + B_1^{f1} \cos \eta_{f1} x_3 + B_1^{f2} \cos \eta_{f2} x_3) e^{-\xi x_1} e^{i\omega t}, \\ u_3^s &= (A_3^{s1} \cos \eta_{s1} x_3 + A_3^{s2} \cos \eta_{s2} x_3 + B_3^{s1} \sin \eta_{s1} x_3 + B_3^{s2} \sin \eta_{s2} x_3) e^{-\xi x_1} e^{i\omega t}, \\ u_1^s &= (A_1^{s1} \sin \eta_{s1} x_3 + A_1^{s2} \sin \eta_{s2} x_3 + B_1^{s1} \cos \eta_{s1} x_3 + B_1^{s2} \cos \eta_{s2} x_3) e^{-\xi x_1} e^{i\omega t}, \end{aligned} \quad (4.8)$$

in which the amplitudes with the subscript 1 are known in terms of the amplitudes with the subscript 3 from (4.7). Consequently, there are only eight unknown amplitudes in (4.8). Substituting from (4.8) into the boundary conditions (2.5) - (2.7) and employing (4.7), we obtain¹³

$$\begin{aligned} \sum_{i=1}^2 [A_3^{fi} (\eta_{fi} \mu^{fi} - \xi) \cos \eta_{fi} h^f - B_3^{fi} (\eta_{fi} \nu^{fi} + \xi) \sin \eta_{fi} h^f] &= 0, \\ \sum_{i=1}^2 [A_3^{fi} (c_{13}^f \xi \mu^{fi} + c_{33}^f \eta_{fi}) \sin \eta_{fi} h^f + B_3^{fi} (c_{13}^f \xi \nu^{fi} - c_{33}^f \eta_{fi}) \cos \eta_{fi} h^f] &= 0, \\ \sum_{i=1}^2 [B_3^{fi} (c_{33}^f \eta_{fi} - c_{13}^f \xi \nu^{fi}) - B_3^{si} (c_{33}^s \eta_{si} - c_{13}^s \xi \nu^{si})] &= 0, \\ \sum_{i=1}^2 [A_3^{fi} - A_3^{si}] &= 0, \quad \sum_{i=1}^2 [A_3^{fi} c_{44}^f (\eta_{fi} - \xi \mu^{fi}) - A_3^{si} c_{44}^s (\eta_{si} - \xi \mu^{si})] = 0, \\ \sum_{i=1}^2 [B_3^{fi} \nu^{fi} - B_3^{si} \nu^{si}] &= 0, \quad \sum_{i=1}^2 [A_3^{si} (c_{13}^s \xi + c_{33}^s \eta_{si} \mu^{si}) \sin \eta_{si} h^s + B_3^{si} (c_{33}^s \eta_{si} - c_{13}^s \xi \nu^{si}) \cos \eta_{si} h^s] = 0, \\ \sum_{i=1}^2 [A_3^{si} (\mu^{si} \eta_{si} - \xi) \cos \eta_{si} h^s + B_3^{si} (\nu^{si} \eta_{si} + \xi) \sin \eta_{si} h^s] &= 0. \end{aligned} \quad (4.9)$$

Equations (4.9) constitute a system of eight linear homogeneous algebraic equations in A_3^{fi} , B_3^{fi} , A_3^{si} and B_3^{si} , which yield nontrivial solutions when the determinant of the coefficients vanishes¹³.

In order to obtain the asymptotic form of the solution in the limit of small decay numbers along the plate and near a thickness-extensional frequency

determined from a root of (3.9) we return to (4.3) - (4.6) and following earlier simpler work⁷ note that since the decay number along the composite plate ξ , is near zero and ω is in the vicinity of a thickness-extensional resonant frequency ω_h and the η_{fi} and η_{si} are finite, i.e., ξ is small while ω and the η are not, to zero order in ξ we have

$$\begin{aligned} (\eta_{f1}^0)^2 &= \frac{\rho^f \omega^2}{c_{33}^f}, \quad (\eta_{f2}^0)^2 = \frac{\rho^f \omega^2}{c_{44}^f}, \quad (\eta_{s1}^0)^2 = \frac{\rho^s \omega^2}{c_{33}^s}, \\ (\eta_{s2}^0)^2 &= \frac{\rho^s \omega^2}{c_{44}^s}, \end{aligned} \quad (4.10)$$

from (4.3) - (4.6), and to lowest (second) order in ξ we have

$$\begin{aligned} A_1^{f1} &= -r^f A_3^{f1} \frac{\xi}{\eta_{f1}}, \quad A_3^{f2} = -r^f A_1^{f2} \frac{\xi}{\eta_{f2}}, \\ B_1^{f1} &= r^f B_3^{f1} \frac{\xi}{\eta_{f1}}, \quad B_3^{f2} = r^f B_1^{f2} \frac{\xi}{\eta_{f2}}, \\ A_1^{s1} &= -r^s A_3^{s1} \frac{\xi}{\eta_{s1}}, \quad A_3^{s2} = -r^s A_1^{s2} \frac{\xi}{\eta_{s2}}, \\ B_1^{s1} &= r^s B_3^{s1} \frac{\xi}{\eta_{s1}}, \quad B_3^{s2} = r^s B_1^{s2} \frac{\xi}{\eta_{s2}}, \end{aligned} \quad (4.11)$$

where¹³

$$r^f = \frac{c_{13}^f + c_{44}^f}{c_{33}^f - c_{44}^f}, \quad r^s = \frac{c_{13}^s + c_{44}^s}{c_{33}^s - c_{44}^s}. \quad (4.12)$$

From Eqs. (4.10) and (3.3) it is clear that

$$\eta_{f1}^0 \equiv \eta_f^0, \quad \eta_{s1}^0 \equiv \eta_s^0, \quad (4.13)$$

which with (3.7) means that η_f^0 satisfies (3.9). Moreover, from Eqs. (4.10) we note that

$$\eta_{f2}^0 = \kappa^f \eta_f^0, \quad \eta_{s2}^0 = \kappa^s \eta_s^0, \quad (4.14)$$

where

$$\kappa^f = (c_{33}^f / c_{44}^f)^{1/2}, \quad \kappa^s = (c_{33}^s / c_{44}^s)^{1/2}. \quad (4.15)$$

Since we are interested in the solution in the vicinity of a pure thickness-extensional mode and in small deviations from that mode, we take

$$\eta_{f1} h^f = \eta_f^0 h^f + \delta_f, \quad \eta_{s1} h^s = \eta_s^0 h^s + \delta_s, \quad (4.16)$$

where δ_f and δ_s are small. From (4.3) - (4.6) and (4.16) we obtain¹³

$$\begin{aligned} s^f \xi^2 - c_{33}^f [(\eta_f^0)^2 + 2\eta_f^0 \frac{\delta_f}{h^f}] + \rho^f \omega^2 &= 0, \\ s^s \xi^2 - c_{33}^s [(\eta_s^0)^2 + 2\eta_s^0 \frac{\delta_s}{h^s}] + \rho^s \omega^2 &= 0, \end{aligned} \quad (4.17)$$

where

$$s^f = r^f (c_{13}^f + c_{44}^f) + c_{44}^f, \quad s^s = r^s (c_{13}^s + c_{44}^s) + c_{44}^s. \quad (4.18)$$

From (4.17), with the aid of (3.3), (3.6) and (3.7), we obtain¹³

$$\delta_g = K \xi^2 + \mu \sigma \delta_f, \quad (4.19)$$

where

$$K = (\rho^s h^s / 2 \eta_{33}^0 c_{33}^s) [(g^f / \rho^f) - (g^s / \rho^s)]. \quad (4.20)$$

Substituting from (4.11), (4.13), (4.14), (4.16) and (4.19) into the appropriate equations¹³, expanding, appropriately employing (3.9) and retaining terms up to ξ^2 , after much tedious algebra we obtain¹³

$$M \xi^2 - \bar{c}_{33}^f (\eta_f^0)^2 + \rho^f \omega^2 = 0, \quad (4.21)$$

in which M is a very lengthy expression, which is too cumbersome to present here¹³.

Equation (4.21) is the asymptotic form of the straight-crested dispersion relation for small decay number ξ , along the composite plate, in the vicinity of a thickness-extensional resonant frequency ω_e^0 with a ground electrode of zero thickness ($2h^s = 0$), i.e., with $R^s = 0$ in (3.12), which yields

$$\omega_e^0 = (\bar{c}_{33}^f / \rho^f)^{1/2} \eta_f^0. \quad (4.22)$$

The thin ground electrode serves only to change the thickness-extensional resonant frequency for $\xi = 0$ from ω_e^0 to ω_e by virtue of the change from η_f^0 to η_f , where

$$\eta_f = \eta_f^0 (1 - R^s / G^0), \quad (4.23)$$

as we know from Eqs. (3.8) - (3.12) of the thickness-extensional solution presented in Sec. 3. As a consequence, it is clear from (4.21) that for a thin ground electrode, the straight-crested dispersion relation takes the form

$$M \xi^2 - \bar{c}_{33}^f \eta_f^2 + \rho^f \omega^2 = 0. \quad (4.24)$$

In the electroded region with a driving electrode we are interested in the situation in which we have a propagation wavenumber ξ instead of a decay number ξ . Clearly, the relation for a propagation wavenumber ξ can readily be obtained from the equation for a decay number ξ simply by replacing ξ by $i\xi$. Furthermore, for the same reasons as in the case of the ground electrode, in the electroded composite plate with a driving electrode the thickness-extensional resonant frequency for $\xi = 0$ is changed from ω_e^0 to ω_e given in (3.18) by virtue of the change from η_f^0 to η_f , where

$$\eta_f = \eta_f^0 (1 + P^0 / G^0), \quad (4.25)$$

from Eqs. (3.15) and (3.16), and P^0 , which is negative, is given in (3.17). As a result of the foregoing, it is clear from (4.21) that for the electroded plate with a driving electrode, the straight-crested dispersion relation takes the form

$$-M \xi^2 - \bar{c}_{33}^f \eta_f^2 + \rho^f \omega^2 = 0, \quad (4.26)$$

which is in accordance with what happens in an earlier treatment⁷ of a simpler case.

If M is positive the dispersion curves predicted by each of the three equations (4.21), (4.24) and (4.26) are presented in Figure 4, which clearly shows that the influence of the very thin electrode films along with the small piezoelectric coupling simply serves to shift

the pure thickness frequencies down from ω_e^0 to $\hat{\omega}_e$ to $\bar{\omega}_e$ with exactly the same curves emanating from each of the three points. The curvatures of each of the three dispersion curves at cut-off are exactly the same because the difference between the three frequencies ω_e^0 , $\hat{\omega}_e$ and $\bar{\omega}_e$ is very small since the piezoelectric coupling k and electrode mass loading factors R^s and R^s are very small. On the other hand if M is negative, the three dispersion curves take a quite different form shown in Figure 5, with each curve having the same curvature at cutoff. For a given zinc-oxide film on silicon the sign of M in a given instance depends on the ratio of the thickness of the silicon layer to that of zinc-oxide σ . When M changes sign it does not go through zero but instead goes to plus infinity and comes back from minus infinity. At the point at which M changes sign, the asymptotic expression we have obtained for M breaks down because the η_{f2}^0 and η_{f2}^0 thickness-shear wavenumbers satisfy a relation analogous to (3.9) at the same frequency as the thickness-extensional resonance ω_e^0 and in that case an expansion about the pure thickness-shear resonant solution must be made for the $\eta_{f2} - \eta_{f2}$ solution as well as the expansion that has been made for the $\eta_{f1} - \eta_{f1}$ solution. This exceptional situation occurs whenever there is a coincidence¹⁶ of frequencies of the pure thickness solutions that are coupled in the dispersion equation and will not be treated here.

On account of the considerable amount of algebra involved in the evaluation of M , we have checked its accuracy to be sure that no error was made. To this end for a zinc-oxide thin film on a silicon layer, we have calculated the dispersion curves in the vicinity of the lowest thickness-extensional resonant frequency using both the complete solution obtained at the beginning of this section (in which the algebra is not extravagant) and the expression for M and compared the results. It can be seen from Figure 6 that the dispersion curves compare very well for $|gh^s| \ll 1$. Furthermore, for $\sigma = 1/3$, $M = 31.7761 \times 10^{10} \text{ N/M}^2$ and the value of M obtained from numerical differentiation of the dispersion curve is $M = 31.7761 \times 10^{10} \text{ N/M}^2$, which shows that they are in excellent agreement. Consequently, we can be certain that no error was made in the determination of the expression for M .

In the trapped energy resonator with high Q , an electroded region abuts an unelectroded region, as shown in Figure 1. Since only the thickness dependence of the electrical variables has been included, the electrical continuity conditions at a junction between an electroded and unelectroded region of the plate need not be satisfied here, as in earlier work⁷⁻⁹. Consequently, at such a junction we require the continuity of the mechanical quantities

$$u_3^f, T_{11}^f, u_1^f, T_{13}^f, \quad (4.27)$$

in the film and the same quantities with the superscript s in the silicon only. Since in trapped energy resonators or acoustically coupled resonator filters with trapping in the composite plate the overall section properties of each constituent, i.e., the material constants, the film and layer thicknesses and the thickness wavenumbers η_f^0 and η_f^0 , are either the same or very nearly the same in the electroded and unelectroded regions of the composite plate, we can satisfy, approximately, the four continuity conditions in (4.27) along with the four with the superscript s because they are coupled together in each region by the asymptotic solution in that region, simply by requiring the continuity of only the two quantities⁹

$$u_3^f, u_{3,1}^f, \quad (4.28)$$

at the junction, as in earlier work⁷, or equivalently

$$u_3^s, u_{3,1}^s. \quad (4.29)$$

As noted earlier the value of M , which determines the dispersion characteristics, varies with the ratio of the thicknesses σ . In order to exhibit this variation we calculate M for zinc-oxide films on silicon layers for a few values of σ and plot the dispersion curves. For a relatively thick zinc-oxide film on a silicon layer corresponding to a value of σ of 1/3, $M = 31.7761 \times 10^{10}$ N/M² and the dispersion curves take the form shown in Figure 7. From Figure 7 it is clear that if the driving frequency ω is confined to the range $\bar{\omega}_u < \omega < \bar{\omega}_u$, trapping in the unelectroded regions will occur for the configuration shown in Figure 1. On the other hand, for a zinc-oxide film somewhat thinner than the silicon layer, corresponding to a value of σ of 2.1923, $M = -30.2993 \times 10^{10}$ N/M² and the dispersion curves take the form shown in Figure 8, which shows that trapping in the unelectroded regions will not occur for the configuration shown in Figure 1 for any value of ω . However, if outside the region containing the driving electrode the silicon layer is thickened slightly as shown in Figure 9, so as to make $\bar{\omega}_u < \bar{\omega}_u$ by a small but sufficient amount, the dispersion curves take the positions shown in Figure 10 and trapping in the electroded region will occur for ω in the range $\bar{\omega}_u < \omega < \bar{\omega}_u$. For a zinc-oxide film on a silicon layer we have found that the transition in the sign of M occurs at $\sigma = 1.0654$. For a thickness ratio anywhere near the transition value of σ , no trapping occurs at all because the dispersion curves take the form shown in Figure 11. At the precise transition value of σ there is no imaginary ξ dispersion loop at all and the real dispersion curves meet at a point on the ω -axis with nonzero slopes.

5. Variable-Crested Dispersion Relations

In Sec. 4 we obtained the asymptotic limit of the straight-crested dispersion relations for small wavenumbers along the x_1 -direction of the composite plate. Since the zinc-oxide is isotropic in the plane of the plate and the silicon is cubic with the coordinate axes along the cube edges, the straight-crested dispersion relation obviously holds along the x_2 -direction also. In this section we extend the asymptotic limit of the straight-crested solutions to solutions with variable crests. Clearly, only the elastic constants of the silicon influencing the asymptotic solution obtained in Sec. 4 need be considered, i.e., c_{12}^s , c_{44}^s and c_{55}^s , because the constants of zinc-oxide are independent of the angle θ in the plane of the plate. From the equations for silicon it has been shown that the asymptotic limiting form of the equations can be written in the form¹³

$$(c_{13}^s + c_{44}^s) \nabla^2 u_{3,3}^s + c_{44}^s u_{3,3}^{sp} = \rho^s \ddot{u}_{3,3}^s, \\ c_{44}^s (\nabla^2)^2 u_{3,3}^s + (c_{13}^s + c_{44}^s) \nabla^2 \cdot u_{3,3}^s + c_{33}^s u_{3,3}^s = \rho^s \ddot{u}_{3,3}^s, \quad (5.1)$$

$$T_{33}^s = c_{13}^s \nabla^2 \cdot u_{3,3}^s + c_{33}^s u_{3,3}^s, \quad \xi^{sp} = c_{44}^s (\nabla^2 u_{3,3}^s + u_{3,3}^{sp}), \quad (5.2)$$

in which the planar displacement u^s , traction t^s and differential operator ∇^2 , respectively, are defined by

$$u^s = \xi_1 u_1^s + \xi_2 u_2^s, \quad \xi^{sp} = \xi_1 T_{31}^s + \xi_2 T_{32}^s, \\ \nabla^2 = \xi_1 \frac{\partial}{\partial x_1} + \xi_2 \frac{\partial}{\partial x_2}, \quad (5.3)$$

and where ξ_1 and ξ_2 are unit base vectors in the x_1 - and x_2 -directions, respectively. Equations (5.1) and (5.2) are clearly isotropic in the x_1 - x_2 plane and show that in the asymptotic limit of small wavenumbers along the plate the dispersion relation (4.26) is independent of direction in the plane of the plate, and, hence, is isotropic. Consequently, we can, by a superposition of two straight-crested standing waves with wave normals in the directions given by $+\theta$ and $-\theta$, respectively, obtain a variable-crested solution of the form^{1,13}

$$u_3^f = 2(A_3^{f1} \cos \eta_{f3}^0 x_3 + B_3^{f1} \sin \eta_{f3}^0 x_3) \cos \bar{\xi} x_1 \cos \bar{\nu} x_2 e^{i\omega t}, \\ u_3^s = 2(A_3^{s1} \cos \eta_{s3}^0 x_3 + B_3^{s1} \sin \eta_{s3}^0 x_3) \cos \bar{\xi} x_1 \cos \bar{\nu} x_2 e^{i\omega t}, \quad (5.4)$$

$$u_1^f = \frac{2\bar{\xi}}{\bar{\zeta}} \tau^f(x_3) \sin \bar{\xi} x_1 \cos \bar{\nu} x_2 e^{i\omega t}, \\ u_2^f = \frac{2\bar{\nu}}{\bar{\zeta}} \tau^f(x_3) \cos \bar{\xi} x_1 \sin \bar{\nu} x_2 e^{i\omega t}, \\ u_1^s = \frac{2\bar{\xi}}{\bar{\zeta}} \tau^s(x_3) \sin \bar{\xi} x_1 \cos \bar{\nu} x_2 e^{i\omega t}, \\ u_2^s = \frac{2\bar{\nu}}{\bar{\zeta}} \tau^s(x_3) \cos \bar{\xi} x_1 \sin \bar{\nu} x_2 e^{i\omega t}, \quad (5.5)$$

where

$$\tau^f = A_1^{f1} \sin \eta_{f3}^0 x_3 + A_1^{f2} \sin \kappa \eta_{f3}^0 x_3 + B_1^{f1} \cos \eta_{f3}^0 x_3 \\ + B_1^{f2} \cos \kappa \eta_{f3}^0 x_3, \\ \tau^s = A_1^{s1} \sin \eta_{s3}^0 x_3 + A_1^{s2} \sin \kappa \eta_{s3}^0 x_3 + B_1^{s1} \cos \eta_{s3}^0 x_3 \\ + B_1^{s2} \cos \kappa \eta_{s3}^0 x_3, \quad (5.6)$$

to lowest order in $\bar{\zeta}$ where

$$\bar{\zeta}^2 = \bar{\xi}^2 + \bar{\nu}^2, \quad \cot \theta = \bar{\xi}/\bar{\nu}, \quad (5.7)$$

and from (4.26) we see that the dispersion relation takes the form

$$-M(\bar{\xi}^2 + \bar{\nu}^2) - \bar{\epsilon}_{33}^f \bar{\nu}^2 + \rho^f \omega^2 = 0. \quad (5.8)$$

Moreover, it has been shown¹³ that u_1^f and u_2^f are large while u_1^s , u_2^s , u_3^f and u_3^s are small and are of no further direct interest in this work. This fact has already been used in obtaining (4.28) and (4.29). Furthermore, it has been shown that¹³

$$u_3^{fn} = (A_3^{fn} \cos \eta_{fn}^0 x_3 + B_3^{fn} \sin \eta_{fn}^0 x_3) f^n(x_1, x_2, t), \\ 0 < x_3 < h^f, \\ u_3^{sn} = (A_3^{sn} \cos \eta_{sn}^0 x_3 + B_3^{sn} \sin \eta_{sn}^0 x_3) f^n(x_1, x_2, t), \\ 0 > x_3 > -h^s, \quad (5.9)$$

very accurately in each region because the section properties are very nearly the same and the piezoelectric coupling is small, and where

$$B_3^{fn} = 1, \quad A_3^{fn} = \cot \eta_{fn}^0 h^f, \quad B_3^{sn} = \frac{\bar{\epsilon}_{33}^s \eta_{sn}^0}{c_{33}^s \eta_{sn}^0}, \\ A_3^{sn} = \cot \eta_{sn}^0 h^s, \quad (5.10)$$

provided the f^n satisfy the homogeneous differential equation

$$M_n \left(\frac{\partial^2 f^n}{\partial x_1^2} + \frac{\partial^2 f^n}{\partial x_2^2} \right) - \epsilon_{33}^f \eta_{fn}^2 f^n - \rho f^n = 0, \quad (5.11)$$

where η_{fn} is understood to be $\bar{\eta}_{fn}$ in the electroded, and η_{fn} in the unelectroded regions, respectively. In addition, from (4.28), (4.29) and (5.9) it is clear that at a junction between regions with and without driving electrodes we have the continuity conditions

$$\bar{f}^n = f^n, \quad \bar{f}_{,1}^n = f_{,1}^n, \quad (5.12)$$

when the junction line is in the x_2 -direction and

$$\bar{f}^n = f^n, \quad \bar{f}_{,2}^n = f_{,2}^n, \quad (5.13)$$

when the junction line is in the x_1 -direction.

It has been shown¹³ that the steady-state solution of the inhomogeneous problem is governed by the inhomogeneous differential equation

$$\sum_n \left[M_n \left(\frac{\partial^2 u_3^{fn}}{\partial x_1^2} + \frac{\partial^2 u_3^{fn}}{\partial x_2^2} \right) - \epsilon_{33}^f \eta_{fn}^2 u_3^{fn} - \rho u_3^{fn} \right] = \rho \frac{e}{f} \omega^2 \frac{V x_3}{c_{33} h} e^{i\omega t}, \quad (5.14)$$

for $0 < x_3 < h^f$, while for $-h^s < x_3 < 0$ we have the homogeneous equation¹³

$$\sum_n \left[M_n \left(\frac{\partial^2 u_3^{sn}}{\partial x_1^2} + \frac{\partial^2 u_3^{sn}}{\partial x_2^2} \right) - \epsilon_{33}^s \eta_{fn}^2 u_3^{sn} - \rho u_3^{sn} \right] = 0, \quad (5.15)$$

where

$$u_3^f = \sum_{n=1,2}^{\infty} u_3^{fn}, \quad u_3^s = \sum_{n=1,2}^{\infty} u_3^{sn}. \quad (5.16)$$

Substituting from (5.9) into (5.14) and (5.15), which are integrated with respect to x_3 and added and employing (5.10) and the orthogonality of the x_3 -dependent functions in the interval $-h^s < x_3 < h^f$, we obtain

$$M_n \left(\frac{\partial^2 f^n}{\partial x_1^2} + \frac{\partial^2 f^n}{\partial x_2^2} \right) - \epsilon_{33}^f \eta_{fn}^2 f^n - \rho f^n = \rho \frac{e}{f} \omega^2 \frac{c_{33}}{h} \frac{G_1}{G_2} V e^{i\omega t}, \quad (5.17)$$

where

$$G_1 = (1 - \cos \eta_{fn}^f h^f) / (\eta_{fn}^f)^2 \sin \eta_{fn}^f h^f, \\ G_2 = \frac{2\eta_{fn}^f \sin \eta_{fn}^f h^f + \sin 2\eta_{fn}^f h^f}{4\eta_{fn}^f \sin^2 \eta_{fn}^f h^f} + \left(\frac{\epsilon_{33}^f \eta_{fn}^f}{c_{33} \eta_{fn}^s} \right)^2 \frac{2\eta_{fn}^s \sin \eta_{fn}^s h^s + \sin 2\eta_{fn}^s h^s}{4\eta_{fn}^s \sin^2 \eta_{fn}^s h^s}. \quad (5.18)$$

Equation (5.17) is for the region with a driving electrode which is driven by a driving voltage V . For an undriven region, $V=0$ and η_{fn} is replaced by $\bar{\eta}_{fn}$. Finally, it should be noted that when the undriven region is thickened slightly, as shown in Figure 9,

η_{fn} in (5.17) must be for the slightly thickened section. However, in the x_3 -dependent functions which have been integrated, η_{fn}^f and η_{fn}^s can always be used. This is a result of the fact that even though all η_{fn} and $\bar{\eta}_{fn}$ are very nearly the same in each region, the small differences in the differential equations are essential¹⁷⁻⁸.

6. Application to Trapped Energy Resonators

In this section we treat the trapped energy resonator, the plan view of which is shown in Figure 12, following the procedures outlined in Refs.8 and 9. This means that the solutions presented here satisfy the equations to second order in the small wavenumbers ξ and v along the plate in all regions except the relatively unimportant corner-type regions (labeled C in Figure 12) in which they are satisfied to first order in ξ and v , in accordance with the discussions in Refs.8 and 9. In accordance with Sec.5 of this work, in the region with a driving electrode we have Eq.(5.17), and, as noted in Sec.5, in any region without a driving electrode, we have

$$M_n \left(\frac{\partial^2 f^n}{\partial x_1^2} + \frac{\partial^2 f^n}{\partial x_2^2} \right) - \epsilon_{33}^f \eta_{fn}^2 f^n - \rho f^n = 0, \quad (6.1)$$

in place of (5.17), and we note that Eq.(6.1) is not satisfied in a corner type region. As noted in Sec.5, the edge conditions to be satisfied at a junction between an electroded and unelectroded region are given by either (5.12) or (5.13), and we note that either (5.12) or (5.13) is also satisfied at a junction between a corner type region and an adjacent unelectroded region. The fact that (6.1) is not satisfied in a corner type region but either (5.12) or (5.13) is satisfied at a junction with a corner type region is the reason that the solution does not hold to second order in ξ and v in the relatively unimportant corner type region but does hold to first order in ξ and v . Since the functions antisymmetric in either x_1 or x_2 cannot be driven, as a solution of the associated homogeneous problem, i.e., with $V=0$, we take the functions symmetric in x_1 and x_2

$$\begin{aligned} \bar{f}^n &= \bar{E}_n \cos \bar{\epsilon} x_1 \cos \bar{\nu} x_2 e^{i\omega t}, \quad 0 < x_1 < l, \quad 0 < x_2 < b, \\ f^{sn} &= E_n^s e^{-\xi^s (x_1 - l)} \cos \bar{\nu} x_2 e^{i\omega t}, \quad l < x_1, \quad 0 < x_2 < b, \\ f^{tn} &= E_n^T \cos \bar{\epsilon} x_1 e^{-v^T (x_2 - b)} e^{i\omega t}, \quad 0 < x_1 < l, \quad b < x_2, \\ f^{cn} &= E_n^C e^{-\xi^s (x_1 - l)} e^{-v^T (x_2 - b)} e^{i\omega t}, \quad l < x_1, \quad b < x_3, \end{aligned} \quad (6.2)$$

the first of which satisfies the homogeneous form of (5.17) in the s -region and the second two of which satisfy (6.1) in the s and T regions, respectively, provided

$$\begin{aligned} -M_n (\bar{\epsilon}^2 + \bar{\nu}^2) - \epsilon_{33}^f \bar{\eta}_{fn}^2 + \rho \bar{\omega}^2 &= 0, \\ M_n (\xi^s{}^2 - \bar{\nu}^2) - \epsilon_{33}^f \eta_{fn}^2 + \rho \bar{\omega}^2 &= 0, \\ M_n (v^T{}^2 - \bar{\epsilon}^2) - \epsilon_{33}^f \bar{\eta}_{fn}^2 + \rho \bar{\omega}^2 &= 0. \end{aligned} \quad (6.3)$$

Substituting from (6.2) into (5.12) at $x_2 = l$ and (5.13) at $x_1 = b$ and introducing the condition for a nontrivial solution, we obtain

$$\bar{\xi}_n \tan \bar{\xi}_n l = \xi_n^s, \quad \bar{v}_n \tan \bar{v}_n b = v_n^T, \quad (6.4)$$

where

$$\begin{aligned} E_n^s &= \bar{E}_n \cos \bar{\xi}_n l, \quad E_n^T = \bar{E}_n \cos \bar{v}_n b, \\ E_n^C &= \bar{E}_n \cos \bar{\xi}_n l \cos \bar{v}_n b. \end{aligned} \quad (6.5)$$

Now, from (6.3), (3.8), (3.10), (3.15) and (3.16) and neglecting products of Δ^f and Δ^T , we obtain

$$\xi_n^s = \left[\frac{2\pi^f}{M_n} (\eta_f^0)^2 \Delta_n^2 - \bar{\xi}_n^2 \right]^{1/2}, \quad v_n^T = \left[\frac{2\pi^f}{M_n} (\eta_f^0)^2 \Delta_n^2 - \bar{v}_n^2 \right]^{1/2}, \quad (6.6)$$

where

$$\Delta_n^2 = -(P^{on} + R'')/G^{on}. \quad (6.7)$$

Equations (6.4), with (6.6) and (6.7), constitute two independent transcendental equations for $\bar{\xi}_n$ and \bar{v}_n for a given l and b , which may readily be solved for all the symmetric modes with respect to x_1 and x_2 . The eigenfrequency for that mode may then be determined from (6.3)₁. Some results for a typical calculation for the fundamental and an overtone mode are given in Table I.

It has been shown¹³ from the series representation of the steady-state solution for the linear forced vibrations of this composite trapped energy resonator that in the vicinity of a resonance, say the N th, one term in the sum dominates and the others are negligible so that in the vicinity of said resonance the solution in the electrode region of the piezoelectric film may be written

$$\begin{aligned} \bar{u}_3^f &= (A_3^{fln} \cos \eta_{fn}^0 x_3 + B_3^{fln} \sin \eta_{fn}^0 x_3) H^{N\tau\epsilon} \cos \bar{\xi}_n x_1 \\ &\quad \cdot \cos \bar{v}_n x_2 e^{i\omega t} - (e_{33}^f V / c_{33}^f h^f) x_3 e^{i\omega t}, \\ \bar{\varphi}^f &= \frac{e_{33}^f}{\epsilon_{33}} (A_3^{fln} \cos \eta_{fn}^0 x_3 + B_3^{fln} \sin \eta_{fn}^0 x_3) H^{N\tau\epsilon} \\ &\quad \cdot \cos \bar{\xi}_n x_1 \cos \bar{v}_n x_2 e^{i\omega t} + (C_{N\tau\epsilon} x_3 + K_{N\tau\epsilon}) e^{i\omega t} \\ &\quad + (V/h^f) x_3 e^{i\omega t}, \end{aligned} \quad (6.8)$$

and it should be noted that we can write the associated solution functions in all the other regions if we wish, but they are not of sufficient interest. In (6.8) $C_{N\tau\epsilon}$ and $K_{N\tau\epsilon}$ may readily be obtained from the electrical boundary conditions (2.8)₂ and (2.9) and A_3^{fln} and B_3^{fln} are given in (5.10)_{1,2} and, as usual, $\omega_{N\tau\epsilon}$ in (6.8) is to be replaced by

$$\bar{\omega}_{N\tau\epsilon} = \omega_{N\tau\epsilon} + i\omega_{N\tau\epsilon}/2Q_N, \quad (6.9)$$

in which Q_N is the unloaded quality factor of the composite trapped energy resonator in the N th thickness-extensional mode. Also, $H^{N\tau\epsilon}$ is given by¹³

$$H^{N\tau\epsilon} = \frac{4\pi^f V G_{N\tau\epsilon}^2 \sin \bar{\xi}_n l \sin \bar{v}_n b}{c_{33}^f (1 - \omega_{N\tau\epsilon}^2 / \omega^2) h^f G_{N\tau\epsilon}^2 \bar{v}_n L_{N\tau\epsilon}}, \quad (6.10)$$

where the normalization integral $L_{N\tau\epsilon}$ is given by¹³

$$\begin{aligned} L_{N\tau\epsilon} &= b l \left[\left(1 + \frac{\sin 2\bar{\xi}_n l}{2\bar{\xi}_n l} \right) \left(1 + \frac{\sin 2\bar{v}_n b}{2\bar{v}_n b} \right) \right. \\ &\quad \left. + \frac{\cos^2 \bar{v}_n b}{\bar{v}_n^2 b} \left(1 + \frac{\sin 2\bar{\xi}_n l}{2\bar{\xi}_n l} \right) \right. \\ &\quad \left. + \frac{\cos^2 \bar{\xi}_n l}{\bar{\xi}_n^2 l} \left(1 + \frac{\sin 2\bar{v}_n b}{2\bar{v}_n b} \right) + \frac{\cos^2 \bar{\xi}_n l \cos^2 \bar{v}_n b}{\bar{\xi}_n^2 \bar{v}_n^2} \right]. \end{aligned} \quad (6.11)$$

The admittance $Y_{N\tau\epsilon}$ of this piezoelectric film on silicon composite plate operating in the N th trapped energy mode is obtained by substituting from (6.8) with (6.10) into (2.4)₂ with $k = l = 3$ which is then substituted into

$$I = -4 \int_0^l \int_0^b \bar{D}_3 dx_1 dx_2, \quad (6.12)$$

with the result

$$\begin{aligned} Y_{N\tau\epsilon} &= \frac{I}{V} = \frac{4b l i \omega \epsilon_{33}^f}{h^f} (1 + k^2) \\ &\quad + \frac{16 i \omega \epsilon_{33}^f k^2 (G_{N\tau\epsilon}^2 (\eta_{fn}^0)^2 \sin^2 \bar{\xi}_n l \sin^2 \bar{v}_n b)}{[(\omega_{N\tau\epsilon}^2 / \omega^2) - 1] (h^f)^2 G_{N\tau\epsilon}^2 \bar{v}_n L_{N\tau\epsilon}}. \end{aligned} \quad (6.13)$$

The quantities \hat{C}_0 and $\hat{C}_{N\tau\epsilon}$ defined by

$$\hat{C}_0 = 4b l \epsilon_{33}^f (1 + k^2) / h^f, \quad (6.14)$$

$$\begin{aligned} \hat{C}_{N\tau\epsilon} &= 16 \epsilon_{33}^f k^2 (G_{N\tau\epsilon}^2 (\eta_{fn}^0)^2 \sin^2 \bar{\xi}_n l \\ &\quad \cdot \sin^2 \bar{v}_n b / (h^f)^2 G_{N\tau\epsilon}^2 \bar{v}_n L_{N\tau\epsilon}), \end{aligned} \quad (6.15)$$

are called the static capacitance and motional capacitance, respectively, and have been calculated for a few typical cases and are given in Table I.

Acknowledgements

We wish to thank Dr. T.W. Grudkowski of United Technologies Research Center for bringing this problem to our attention and for many valuable discussions.

This work was supported in part by the Army Research Office under Contract No. DAAG 29-82-K-0130 and the United Technologies Research Center.

References

1. T.W. Grudkowski, J.F. Black, T.M. Reader, D.E. Cullen and R.A. Wagner, "Fundamental-Mode VHF/UHF Miniature Acoustic Resonators and Filters on Silicon," *Appl. Phys. Lett.*, **37**, 993 (1980).
2. K.M. Lakin and J.S. Wang, "Acoustic Bulk Wave Composite Resonators," *Appl. Phys. Lett.*, **38**, 125 (1981).
3. K.H. Nakamura, H. Sasaki and H. Slumser, "ZnO/SiO₂ Diaphragm Composite Resonator on a Silicon Wafer," *Electronics Lett.*, **17**, 507 (1981).

4. T.W. Grudkowski, J.F. Black, G.W. Drake and D.E. Cullen, "Progress in the Development of Miniature Thin Film BAW Resonator and Filter Technology," Proceedings of the 36th Annual Symposium on Frequency Control, U.S. Army Electronics Research and Development Command, Fort Monmouth, New Jersey, 539 (1982).
5. J.S. Wang and K.M. Lakin, "Low Temperature Coefficient Bulk Acoustic Wave Composite Resonators," Appl. Phys. Lett., **40**, 308 (1982).
6. An analysis of this type of structure for the case of strip electrodes when trapping is not present is given in R.F. Milson, "Two-Dimensional Theory of Thin Film ZnO Resonators on Silicon," 1982 Ultrasonics Symposium Proceedings, IEEE Cat. No.82 CH 1823-4, Institute of Electrical and Electronics Engineers, New York, 484 (1982).
7. H.F. Tiersten, "Analysis of Intermodulation in Thickness-Shear and Trapped Energy Resonators," J. Acoust. Soc. Am., **57**, 667 (1975).
8. H.F. Tiersten, "Analysis of Trapped Energy Resonators Operating in Overtones of Coupled Thickness-Shear and Thickness-Twist," J. Acoust. Soc. Am., **59**, 879 (1976).
9. H.F. Tiersten, "Analysis of Overtone Modes in Monolithic Crystal Filters," J. Acoust. Soc. Am., **62**, 142 (1977).
10. The very severe requirement for the fundamental essentially thickness-extensional mode to trap in the electroded regions of a flat piezoelectric plate is discussed in H.F. Tiersten, J.F. McDonald and P.K. Das, "Monolithic Mosaic Transducer Utilizing Trapped Energy Modes," Appl. Phys. Lett., **29**, 767 (1976).
11. H.F. Tiersten, "Elastic Surface Waves Guided by Thin Films," J. Appl. Phys., **40**, 770 (1969).
12. H.F. Tiersten, Linear Piezoelectric Plate Vibrations (Plenum, New York, 1969), Eqs.(7.28) and (7.29).
13. For more detail see H.F. Tiersten and D.S. Stevens, "An Analysis of Thickness-Extensional Trapped Energy Resonant Device Structures with Rectangular Electrodes in the Piezoelectric Thin Film on Silicon Configuration," to be published in the Journal of Applied Physics (1983).
14. B.A. Auld, Acoustic Fields and Waves in Solids (Wiley, New York, 1973), Vol.I, p.210.
15. Ref.12, Chap.14, Sec.4.
16. This was first noted by Mindlin in the case of isotropic plates and is discussed in R.D. Mindlin, An Introduction to the Mathematical Theory of the Vibrations of Elastic Plates, (U.S. Army Signal Corps Eng. Lab., Fort Monmouth, New Jersey, 1955); Signal Corps Contract DA-36-0395C-56772, pp.2.44-2.51.

TABLE I

N	R'	$\omega_{N00}/2\pi$ (MHz)	\hat{c}_{N00} (pF)
1	0.01	170.276	137.2
1	0.02	168.732	135.1
1	0.03	167.188	134.3
2	0.01	344.781	2.427
2	0.02	342.277	2.390
2	0.03	339.774	2.372
1*	0.01	212.510	50.62

$$2\ell = 2b = .76 \text{ mm} \quad h^f = 15 \text{ } \mu\text{m} \quad h^s = 5 \text{ } \mu\text{m}$$

$$C_0 = 3.77 \text{ pf} \quad h'' = 2000 \text{ } \text{\AA} \text{ (Gold)}$$

* Notched configuration (see Figure 9)
 $h^f = 5.2 \text{ } \mu\text{m} \quad h^s = 11.4 \text{ } \mu\text{m} \quad 2b = 2\ell = 0.76 \text{ mm}$
 $\hat{h}^s = 11.97 \text{ } \mu\text{m} \quad h'' = 2000 \text{ } \text{\AA}$
 $C_0 = 10.87 \text{ pf}$

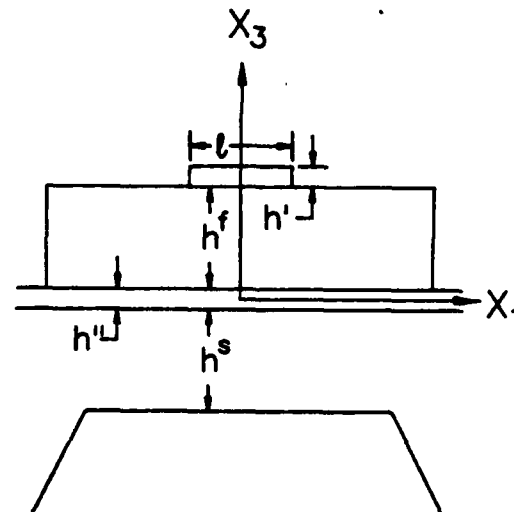


Figure 1 Schematic Diagram of a Cross-Section of a Composite Trapped Energy Resonator Consisting of a Thin Piezoelectric Film on a Silicon Layer

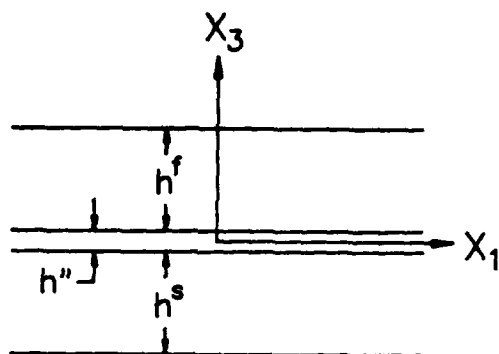


Figure 2 Cross-Section of an Infinite Composite Plate with a Ground Electrode but without a Driving Electrode

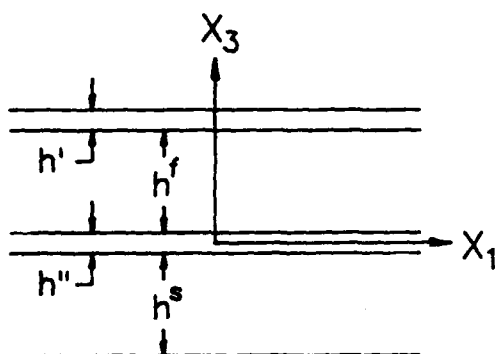


Figure 3 Cross-Section of an Infinite Completely Electroded Composite Plate

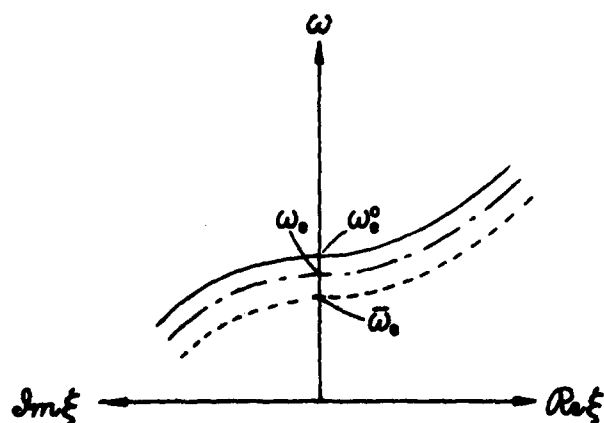


Figure 4 Dispersion Curves in the Vicinity of the Fundamental Thickness-Extensional Resonant Frequency of the ZnO Thin Film on Silicon Composite Plate when M is Positive. The solid curve is for the totally unelectroded composite plate, the dashed curve is for the composite plate with a ground electrode and the dotted curve is for the fully electroded composite plate with shorted electrodes.

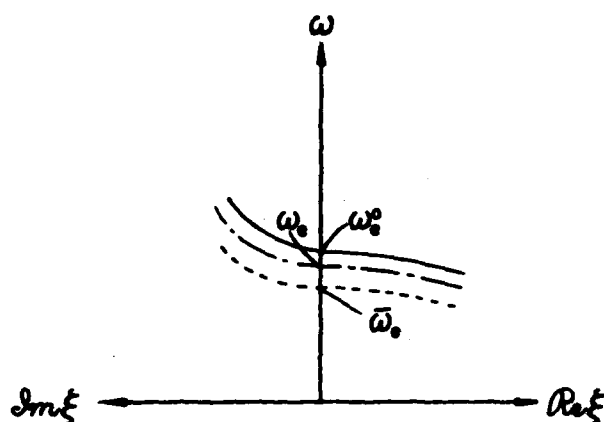


Figure 5 Dispersion Curves in the Vicinity of the Fundamental Thickness-Extensional Resonant Frequency of the Composite Plate when M is Negative. The conventions are the same as in Figure 4.

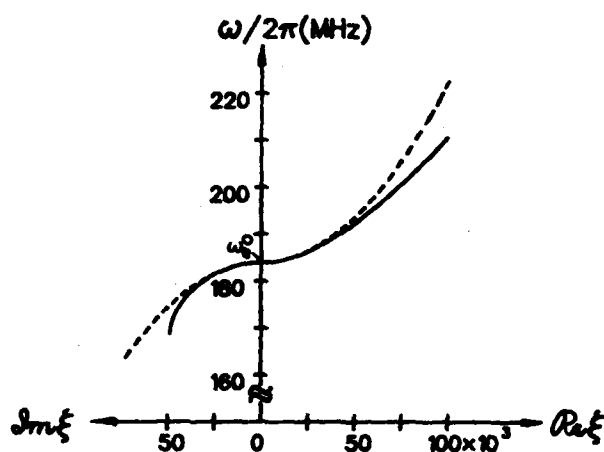


Figure 6 Comparison of the Dispersion Curve Obtained from Eq. (4.43) with that Obtained from the Full Transcendental System Obtained Earlier in Sec. 4 in the Vicinity of the Fundamental Thickness-Extensional Resonant Frequency of the Unelectroded ZnO on Silicon Composite Plate. The solid curve is obtained from the transcendental system and the dotted curve is obtained from Eq. (4.43).

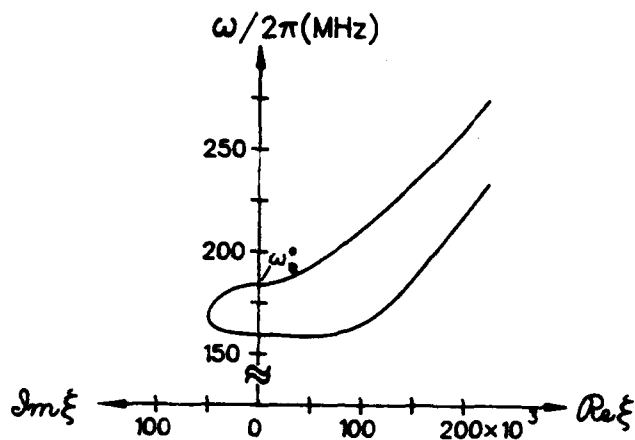


Figure 7 Dispersion Curve Obtained from the Full Transcendental System for a Typical Case in which the Fundamental Thickness-Extensional Resonant Frequency is Above the Second Thickness-Shear Resonant Frequency. The ratio of thickness of the silicon to the ZnO $\sigma = 1/3$.

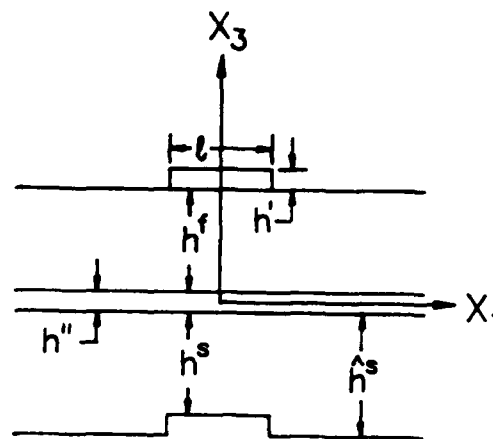


Figure 9 Schematic Diagram of a Cross-Section of a Composite Trapped Energy Resonator when the Fundamental Thickness-Extensional Resonant Frequency is Below the Second Thickness-Shear Resonant Frequency

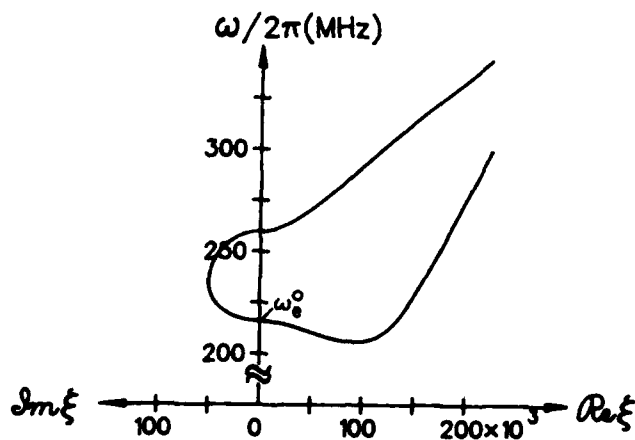


Figure 8 Dispersion Curve Obtained from the Full Transcendental System for a Typical Case in which the Fundamental Thickness-Extensional Resonant Frequency is Below the Second Thickness-Shear Resonant Frequency. The ratio of thickness of the silicon to the ZnO $\sigma = 2.1923$.

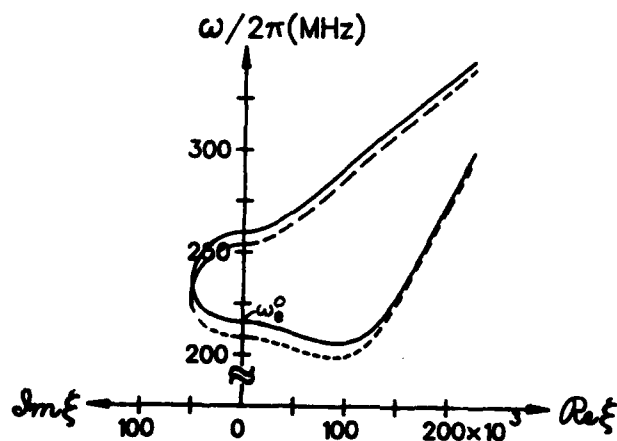


Figure 10 Dispersion Curves for the Composite Trapped Energy Resonator Shown in Figure 9. The solid curve is for the electroded region and the dotted curve is for the slightly thickened unelectroded region. The thickness ratio in the electroded region $\sigma = 2.1923$ and in the unelectroded region $\hat{\sigma} = 2.3019$.

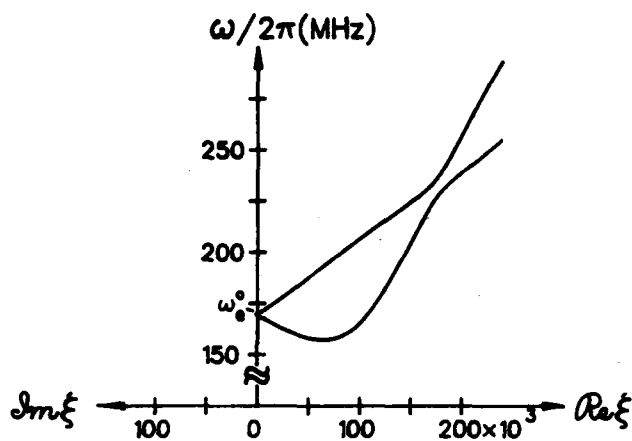


Figure 11 Dispersion Curves for the Unelectroded ZnO on Silicon Composite Plate when the Thickness-Extensional and Thickness-Shear Resonant Frequencies are Near Coincidence. The thickness ratio $\sigma = 1.06542$.

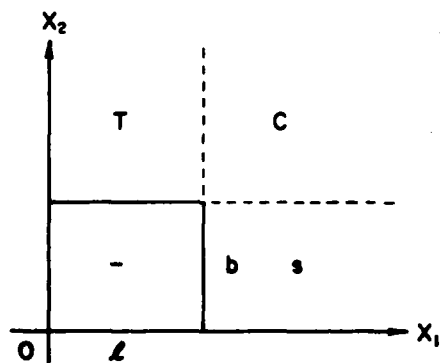


Figure 12 Plan View of One Quadrant of the Trapped Energy Resonator



Miniature LiTaO₃ X-Cut Strip Resonator

Masanobu OKZAKI and Shigenori WATANABE

N.D.K. (Nihon Dempa Kogyo Co., LTD.)
Shibuya-ku, Tokyo, JapanSummary

Because of its excellent Q and electromechanical coupling, Lithium Tantalate (LiTaO₃) is an attractive material for new piezoelectric resonators which can not be realized by Quartz or Ceramics. Lithium Niobate (LiNbO₃), which is one of the family of LiTaO₃, has not as good frequency temperature characteristics as LiTaO₃ resonators. The area of industrial application for LiTaO₃ is expanding rapidly. Examples of this are voltage controlled oscillators (VCO) with a broad range of pullability, wide passband filters and SAW filters. Especially, SAW devices for IF filter of the TV set are industrialized widely. Besides the basic research for the study of LiTaO₃ resonators were reported. (1) Also, practical application of LiTaO₃ for monolithic filter was reported. (2)

This paper reports the development of a miniature LiTaO₃ X-Cut strip resonator enclosed in cylindrical can, which is presently used for wrist watch crystals. An X-Cut resonator of LiTaO₃ has two shear modes. Displacement direction of these modes are perpendicular to each other on the Y-Z plane. One of these modes is so strongly excited due to its high electromechanical coupling and is referred as "fast mode". This mode is applied to a main resonant of a LiTaO₃ X-Cut resonator. The other mode is not excited as strongly, but interacts with the main resonant, and is called the "slow mode". It is necessary to suppress the slow mode as much as possible to permit the fast mode to oscillate in a stable fashion.

This displacement direction of the fast mode is obtained theoretically and is confirmed experimentally. The slow mode is suppressed without affecting the main mode (fast mode) by cutting the strip resonator elongated along the displacement direction of the fast mode. The shape of the wafer is chosen as a strip type elongated along the displacement direction of the fast mode. An X-cut resonator of LiTaO₃ has great many spurious responses. In particular, spurious responses of contour modes are strongly excited. To suppress these unwanted modes, the length and width of the strip are determined by cut and try method. And electrode shape is also chosen by the same way. Further more, in order to eliminate the mechanical influence to the fast mode, one of the short edges of the strip is mounted.

Finally, the chip size of the 3.5 MHz resonator is decreased to 6mm x 1.5mm and 0.58mm in thickness. This resonator is enclosed in the cylindrical can, which is 3mm x 8mm, the standard enclosure for wristwatch crystals. The method of cutting the wafer is very simple and the fabrication technique is the same as the mass production technique of the tuning fork resonators for wristwatches. Therefore, this resonator has a high reliability and has the possibility of mass production. This resonator has good electrical characteristics without beveling which is required in the case of AT-Cut Quartz crystal.

Key Words

LiTaO₃, X-Cut resonator, fast mode, Strip resonator, Mounting the one of the short edges, Cylindrical can, VCO, High reliability

Theoretical background

The secular equation of a LiTaO₃ X-Cut resonator is given as follows.

$$\begin{vmatrix} \bar{C}_{11}-C & 0 & 0 \\ 0 & \bar{C}_{66}-C & \bar{C}_{56} \\ 0 & \bar{C}_{56} & \bar{C}_{55}-C \end{vmatrix} = 0 \quad (1)$$

Where \bar{C}_{11} , \bar{C}_{55} , \bar{C}_{56} and \bar{C}_{66} are the stiffened stiffness. Three independent solutions of C are obtained from this equation. The other side, C is given by the following equation and has the same dimension as stiffness constant.

$$C = \rho \cdot v^2 \quad (2)$$

Where ρ is the density of LiTaO₃ and is 7450 Kg/m³, v is the propagation speed of the thickness shear vibration. One of the solutions has not a term of piezoelectric constants. Other two solutions have a term of piezoelectric constants. Therefore, only two piezoelectric vibration modes are vibrated. Displacement direction of these modes are perpendicular to each other on the Y-Z plane, as shown in Fig 1. One of these modes is called as a "fast mode". Frequency thickness constants, electromechanical coupling and displacement direction are calculated as shown in Table 1. The elastic and piezoelectric constants are referred from a paper by A.W. Warner. (3)

The fast mode is applied to a main resonant of a LiTaO₃ X-Cut resonator, and the slow mode is a unwanted resonant. The slow mode is not excited as strongly, because its electromechanical coupling is about 1/8 of fast mode. But it appears at the frequency to 80% of main resonant. Therefore, it is necessary to suppress the slow mode as much as possible to permit the fast mode to oscillate in a stable fashion.

Displacement direction of these modes are confirmed experimentally by rotating the direction of the mounting. (4) In this experiment, 16 MHz round shaped plate of LiTaO₃ X-Cut resonator is used. The plate is mounted as shown in Fig 2(a). It is considered that slow mode may be suppressed when mounted the position perpendicular to the displacement direction of slow mode. 12 different mounting position are experimented to obtain the best suppression direction of slow mode. At each position, difference of response level between fast mode (f_1) and slow mode (f_2) is measured and plotted as shown in Fig 2(b) and (c). The peak point of d is obtained. This direction of mounting is same as the direction of fast mode which is derived by theoretical calculation.

Strip resonator

Based on the theoretical background confirmed and for the purpose of miniaturization, the shape of the wafer is chosen as a strip type. This strip is elongated along the displacement direction of the fast mode, as shown in Fig 3. It is assumed that miniaturization of the LiTaO_3 X-Cut resonator may be realized adapting this strip without affecting the fast mode.

A characteristic of the frequency responses is related to the combination of the length and the width of the strip. The 3.5 MHz wafers are used for all experiments. Fig 4 shows the frequency responses for four kinds of strip shape. The length and width of strip on Fig 4 (a) are $4.1 \times 1.58\text{mm}$, (b) are $5.0 \times 1.1\text{mm}$, (c) are $6.4 \times 1.58\text{mm}$ and (d) are $6.4 \times 1.3\text{mm}$. From this experiment, the following results are obtained. The length of the strip gives influence on the equivalent resistance of this resonator, and width affects the location and the intensity of the spurious responses. A fine frequency responses and low equivalent resistance are achieved by the suitable length and width of the strip. The length and the width of Fig 4(c) yields the fine characteristics, so this shape is used to next step experiment.

As same as the width of the strip has influence on spurious responses, the shape of electrode has influence on spurious responses. Fig 5 shows the frequency responses for two kinds of electrode shape. The area of electrode of (a) and (b) are the same, though the location and the intensity of the spurious responses, and the intensity of the 3rd overtone mode are different.

Mounting

In order to eliminate the mechanical influence to the fast mode, one of the short edges of the strip is mounted. If both edges of the strip are mounted, the crystal impedance of the fast mode will increase and the resonator can not vibrate constantly at the fast mode only. Fig 6 shows an example of mounting. One of the short edges of the strip is mounted rigidly on the hermetic base with adhesive epoxy. Therefore, electrical connection between electrodes on the wafer and leads are obtained by conductive epoxy.

Because of its rigid mounting, this resonator has a excellent resistance against shock and vibration. This way of mounting is very simple and is the same as the mounting of the tuning fork resonators for wristwatches.

Production process

The fabrication technique is the same as the mass production technique of the tuning fork resonators for wristwatches. Fig 7 shows the production process of this LiTaO_3 X-Cut strip resonators.

At the 1st cutting process, 3 inches wafer of X-Cut LiTaO_3 is broken into $7 \times 10\text{mm}$ wafers. This rectangular wafer is lapped to specified thickness by monitoring its frequency. At the 2nd cutting process, rectangular wafer is broken into 6 pieces of strip chips. About 200 pieces of strip chips are obtained from a 3 inches wafer. It is very difficult to adjust the thickness of the 3 inches wafer with the narrow distribution by current lapping technique. Then the surface of the strip chip is cleaned with water and chemicals. The electrode of silver is attached to the strip chip

by the vacuum evaporation technique. Thickness of the electrode is 2000\AA . The strip chip with electrode is mounted as shown in Fig 6, and is connected electrically by conductive epoxy. A small quantity of the electrode material is additionally evaporated on the electrode and adjusted the frequency within a desired range. Finally, the can containing the resonator is filled with non-active gas, and then sealed so that the wafer will vibrate stably.

Characteristics

The LiTaO_3 X-Cut strip resonator is enclosed in the cylindrical can, which is $3\text{mm} \times 8\text{mm}$, the standard enclosure for wristwatch crystals. Fig 8 shows the picture of this resonator. This resonator has a clean frequency spectrum as shown in Fig 5(b). Typical equivalent constants of the 3.5 MHz strip resonator are give in Table 2. The crystal Impedance (R_s) of 30Ω , which is similar to HC-18/u type AT-Cut 3.5 MHz Quartz crystal, is low enough to oscillate stably. The capacitance ratio (c_0/c_1) is 9.9. This value is $1/250$ to the fundamental AT-Cut Quartz crystal resonators, which means that the wide frequency difference between resonant frequency and antiresonant frequency is obtainable. Therefore VCXO with a broad range of pullability, or a wide passband filter are considered as application. Fig 9 shows the frequency-voltage characteristics of 3.5 MHz voltage controlled oscillator with good linearity. This oscillator offers ± 2000 ppm of pullability, which could not be realized by Quartz crystals.

A LiTaO_3 X-Cut resonator has the parabolic frequency temperature curve. The turnover temperature changes depending on cutting angle, thickness, load capacitance of oscillator and others.(5) Fig 10 shows the frequency temperature characteristics of the 3.5 MHz LiTaO_3 strip resonator. The turnover temperature is more than 100°C . Typical frequency change between -10°C and $+60^\circ\text{C}$ is about 400 ppm. But this value is negligible when applied to VCXO due to its wide pullability. Furthermore, if it is necessary, the compensation of frequency temperature characteristics is easy achieved by circuit design technique. The crystal impedance is stable within a wide temperature range, too.

As shown in Fig 6, the strip chip is mounted rigidly. And this resonator has high resistance against shock and vibration. This resonator withstand against the shock of $3000G$, and the vibration of 500 Hz with $5G$ without any change of electrical and mechanical characteristics. The construction of this resonator is very simple. So, this leads to a stable frequency aging characteristics. In this measurement the resonator is exposed in the high temperature atmosphere at 85°C . The frequency drifts of the resonator are about 15 ppm within a period of 30 days at 85°C . This period and temperature are equivalent to more than 1 year at room temperature. This frequency is negligible also as a VCXO application.

Very good electrical and mechanical characteristics are obtained. And the enclosure is miniature, which has the same dimension as wristwatch crystals. Consequently, this miniature strip resonator of X-Cut LiTaO_3 has wide possibility for industrial application.

Conclusion

A miniature LiTaO_3 X-Cut resonator has been developed by good choice for the shape of the wafer, as a strip type elongated along the displacement direction of the fast mode. One of the short edges of the

strip is mounted rigidly. This resonator is enclosed in the cylindrical can of $3\text{mm} \times 8\text{mm}^2$, the standard enclosure for wristwatch crystals. The excellent electrical characteristics, especially low equivalent resistance and high Q are obtained. Depending on this simple structure, the resonator has a good frequency aging characteristics, and has high resistance against severe shock and vibration. The fabrication technique is the same as the mass production technique of the tuning fork resonators for wristwatches. So, the resonator has a high reliability and possibility of mass production. Applying the resonator, a voltage controlled oscillator with a broad range of pullability is available.

References

- (1) John J. Hannon, P. Lloyd and Robert T. Smith
"Lithium Tantalate and Lithium Niobate Piezoelectric Resonators in the Medium Frequency Range with Low Ratios of Capacitance and Low Temperature Coefficients of Frequency".
IEEE Trans., SU-17, No.4, p239, 1970
- (2) T. Uno
"200 MHz Thickness Extensional Mode LiTaO₃ Monolithic Crystal Filter".
IEEE Trans., SU-22, No.3, p168, 1975
- (3) A. W. Warner, M. Onoe and G. A. Coquin
"Determination of Elastic and Piezoelectric Constants for Crystals in Class (3m)".
J.A.S.A., Vol.42, No.6, p1223, 1967
- (4) M. Okazaki and S. Watanabe (Japanese)
"A Suppression Method of Unwanted Mode on LiTaO₃ X Plate by Mounting".
Proc. of annual meeting of A.S.J., 1977
- (5) A. Fukumoto and A. Watanabe
"Temperature Dependence of Resonant Frequencies of LiNbO₃ Plate Resonators".
Proc. IEEE, Vol 56, p1751, 1968

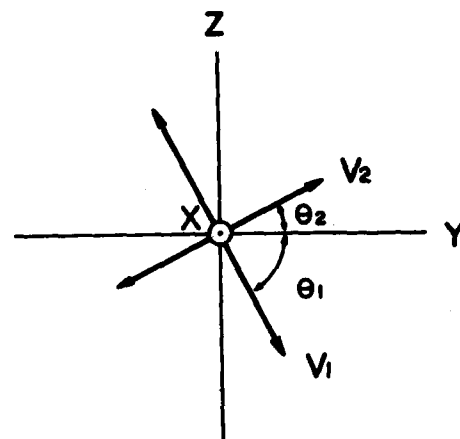


Fig.1 Displacement direction of the fast mode (V_1) and slow mode (V_2) at the X-Cut LiTaO₃ resonator.

	fast mode V_1	slow mode V_2
K_f	2.086	1.669
k	0.443	0.054
θ (deg)	θ_1 52.9	θ_2 37.1

Table 1 Constants of V_1 and V_2 mode.
Kf--- frequency thickness constant(MHz-mm)
k --- electromechanical coupling
 θ --- displacement direction

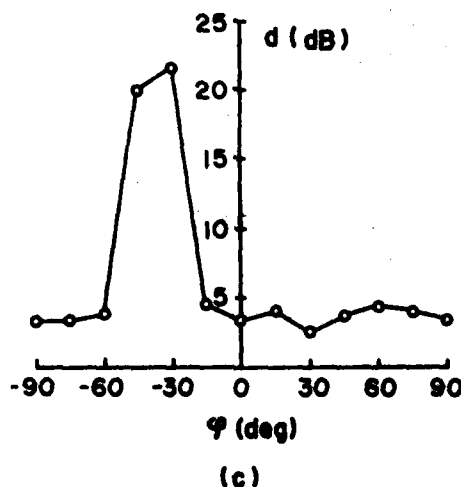
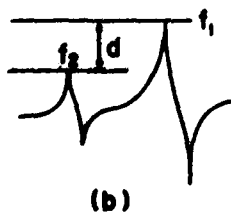
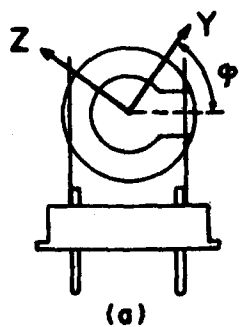


Fig.2 Spurious suppression by mounting.
(a) mounting method
(b) frequency responses
(c) difference of the response level between f_1 and f_2

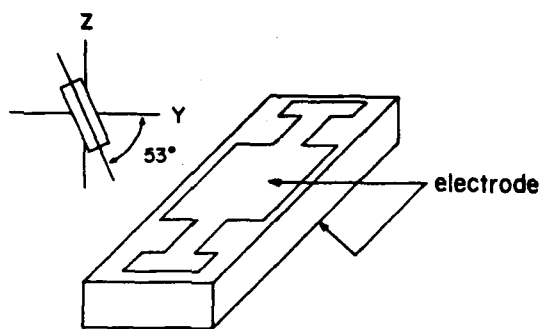


Fig.3 LiTaO₃ X-Cut strip resonator.

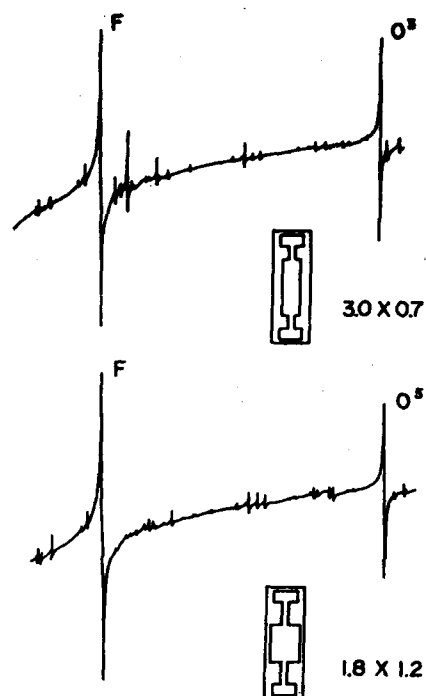


Fig.5 Resonant characteristics for each electrode pattern.



Fig.4 Resonant characteristics for each shape of the wafer.

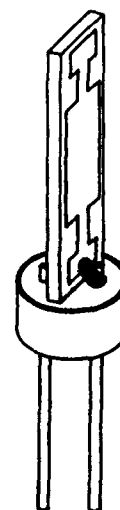


Fig.6 An example of mountin structure.

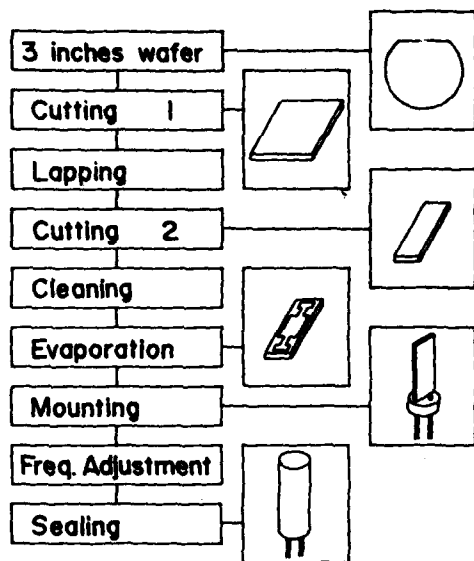


Fig.7 Production process.

	TYP. VALUE	UNIT
f	3.579545	MHz
R_s	30	Ω
C_0	3.8	pF
C_1	0.39	pF
C_0/C_1	9.9	
L_1	5.2	mH
Q	4300	

Table 2 Typical equivalent constants of the LiTaO_3 X-Cut strip resonators.

R_s --- Series equivalent resistance
 C_0 --- Shunt capacitance
 C --- Motional capacitance
 L --- Motional inductance
 Q --- Quality factor

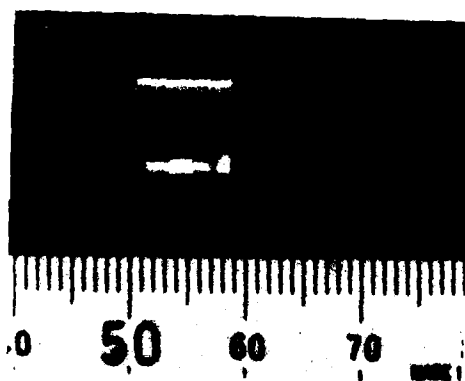


Fig.8 Picture of the LiTaO_3 X-Cut strip resonator.

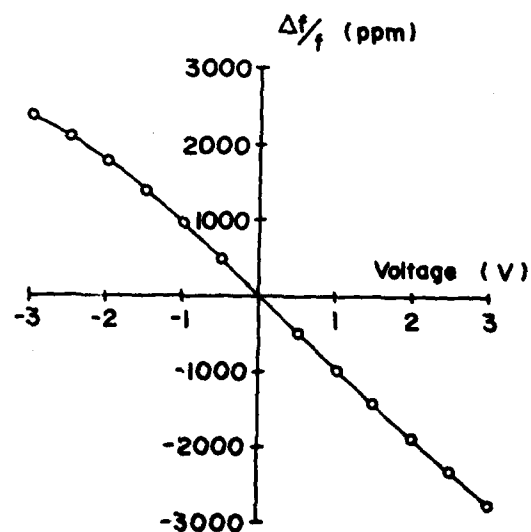


Fig.9 Frequency pullability of 3.5 MHz VCXO.

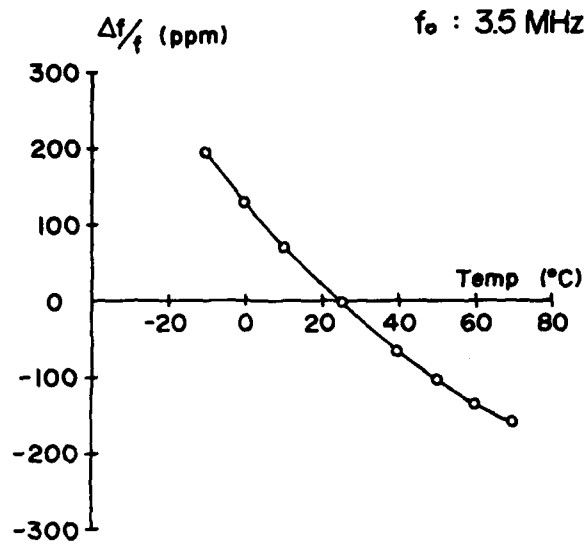


Fig.10 Measured frequency temperature characteristics over -10°C to +70°C.

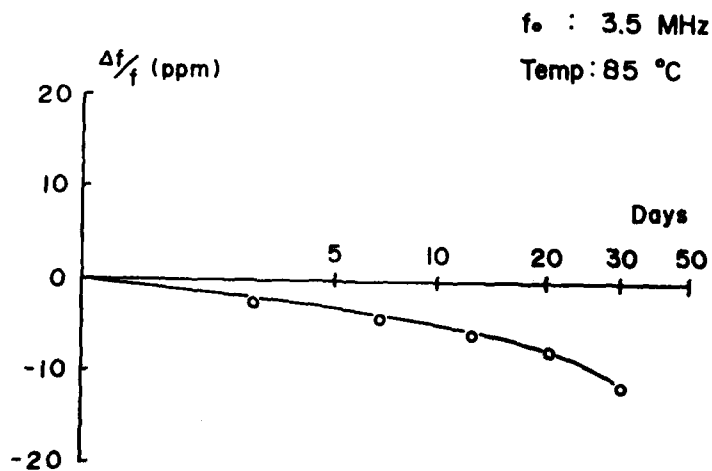


Fig.11 Accelerating frequency aging characteristics exposed at 85°C.



MINIATURIZED LiTaO_3 STRIP RESONATOR

Yoshiro Fujiwara, Sumio Yamada, and Noboru Wakatsuki

FUJITSU LIMITED

1015, Kamikodanaka Nakahara-ku Kawasaki, 211, JAPAN

ABSTRACT

The piezoactive shear wave in a LiTaO_3 X-cut plate has a velocity of 4330 m/s for the fast mode and attractively high electro-mechanical coupling (47%). But the temperature characteristics of the X-cut plate resonator shows a 500 ppm change in the resonant frequency between -10°C and 60°C . We developed a miniaturized strip plate type LiTaO_3 resonator with stable temperature characteristics.

In this paper, the displacement orientation of the thickness shear mode was studied in order to decide the crystal orientation of the longitudinal direction of the strip. Measurement of the Q factor with different rotation angles showed that displacement orientation of the fast mode was -48° from the +Y axis. The width of the strip plate was chosen to separate the main mode from spurious modes. The experimental study shows that the optimum ratio of width to thickness is 1.5 to 2.0 or 2.8 to 3.0. The length of the strip was studied to realize small size while keeping a high Q factor and gaining freedom from edge reflection. The minimum ratio of length to thickness is about 10.

To improve the temperature characteristic, the most suitable relationship between crystal orientation and energy trapping was studied. The experimental study of plates with different crystal orientation revealed the necessary condition to set the turning point around room temperature for optimum size of driving electrodes. A resonator of 3° rotated X-cut plate (along Y axis) with 1.3 mm \times 0.8 mm electrodes has a temperature characteristic that is almost a quadratic curve, and the turning point is near 25°C , in contrast to the -40°C of the X-cut plate, and has no spurious response due to inharmonic overtones. Therefore, the temperature stability is within 100 ppm from -10°C to 60°C .

These techniques can be applied to resonators with a resonance frequency from 3 MHz to 30 MHz. For 6 MHz, the size is 5.5 mm \times 0.8 mm \times 0.32 mm. The ratio of capacitances is 16; 1/20 that of AT-quartz. The mechanical Q factor is more than 10,000. The mechanical energy is so strongly concentrated around the driving electrodes, that the strip can be mounted on a ceramic substrate without reducing the Q factor. Therefore, the chip type resonators were realized, and can be used as "chip components".

1. Introduction

Lithium Tantalate (LiTaO_3) has been studied for electro-mechanical devices, because it has a high piezoelectric activity that cannot be realized by quartz. The fast mode of thickness shear wave in LiTaO_3 X-cut plate has a large electro-mechanical coupling factor (47%)[1]. This makes LiTaO_3 resonators

attractive because the ratio of capacitances is up to 20 times smaller than AT-quartz. These resonators are suitable for use in voltage controlled oscillators with a wide range of variable frequency, and have the possibility to be miniaturized in chip type due to heavy energy trapping under the electrodes. But the temperature coefficient and the crystal cost are not acceptable for industry and consumer use.

Even if the curve of resonance frequency versus temperature of LiTaO_3 X-cut resonator is parabolic, turnaround point often occurs at temperatures other than room temperature [2]. When the electrode size decreases, the turnaround point shifts up to room temperature, but the ratio of capacitances increases [3]. While the zero temperature coefficient was calculated for any orientation, the model did not take energy trapping into account [4]. This paper reports on a single mode resonator with a small capacitance ratio, and with a turnaround point at room temperature by studying the relationship between the value of energy trapping and crystal orientation.

In order to miniaturize the resonator for chip component and wafer cost reduction, the theory of piezoelectric strip [5] is examined for the LiTaO_3 crystal. First, we calculated the direction of particle displacement and confirmed it experimentally. Next, we studied the minimum dimension of the width and length of the strip to eliminate spurious responses from the resonant frequency range.

2. Optimum Crystal Orientation

2-1 Longitudinal direction of strip

One of the most important material factors for resonator is the electro-mechanical coupling factor. The relationship between the coupling factor of LiTaO_3 thickness shear fast mode and the direction of the normal to the plate is shown in Fig.1, where the two angles φ and θ are defined as Fig.2. The plate around X-cut has a large electro-mechanical coupling factor for the thickness shear mode ($\approx 47\%$), but the factors for the thickness shear slow mode and thickness extension mode are small ($\approx 6\%$ and $\approx 0\%$).

In order to have a miniaturized strip resonator, the displacement direction of the thickness shear fast mode must be parallel to the longitudinal direction of the strip. The driving electrodes for the resonator on the top and bottom surfaces then reach throughout the lateral direction, so that the spurious response caused by the thickness twist overtone cannot be generated.

The construction of the piezoelectric strip is defined in Fig.3, where $2W$ is the width of the strip, $2H$ is the thickness, $2L$ is the dimension along the longitudinal direction of the strip, $2l$ is the dimension of the driving electrode, and ϕ is the rotation angle from the Y axis in the X-cut plate.

The direction is calculated as shown in Table 1. Angle ϕ differs according to the material constants

used in the calculation [1][6][7]. To confirm the direction of resonance displacement, we measured the admittance characteristics of resonators having two different rotation angles ϕ (-48° and -54°) as shown in Fig. 4. When the angle ϕ is -54° , spurious responses are generated, but the other has no spurious response. Figure 5 shows the result of the series resistance (R_1) plotted against the angle ϕ . R_1 is minimum at -48° . From Fig. 4 and Fig. 5, the most suitable angle ϕ is -48° .

2-2 Improvement of temperature characteristics

To produce resonators having a turnaround point at room temperature, we studied the most suitable relationship between crystal orientation and energy trapping due to electrode size.

The distribution of the first order temperature coefficient of resonance frequency (RTFC) was calculated as shown in Fig. 6 using the infinite model. Zero RTFC means that the turnaround temperature is room temperature. The solid line shows that RTFC is zero and the dotted line shows that RTFC is -10 ppm/ $^\circ\text{C}$. RTFC is plus around the X-cut plate and minus out of the solid line. The temperature coefficient of antiresonance frequency is always minus for any direction; for example -40 ppm/ $^\circ\text{C}$ at the X-cut plate. Therefore, for the actual finite electrodes, one may obtain a zero RTFC only inside the solid line for zero RTFC as shown in Fig. 6.

The X-cut plate with infinite electrode has a parabolic temperature characteristics curve [3] and a turnaround point at about -40°C . When the electrode size is decreased, the turnaround temperature shifts up to room temperature and RTFC shifts to zero as shown in Fig. 7. In Fig. 7, parameter ζ shows the trapping energy value defined as the following;

$$\zeta = \frac{1}{H} \sqrt{\Delta}$$

$$\Delta = \frac{f_a - f_e}{f_a}$$

where f_a is the cutoff frequency for the non-electrode region and f_e is that for the electrode region. Figure 7 shows that the turnaround point is at room temperature when $\zeta=0.8$, but the ratio of capacitances increases to 22. Therefore, in order to offset the increase, the rotation of θ and φ around X-cut plate was examined. Figure 8 shows the result of the zero temperature coefficient plotted against the crystal orientation for $\zeta=1.0, 1.25$, and 1.5 . If the value ζ is larger than 1.25, an inharmonic overtone is generated [8]. Therefore, the plate with $\theta=4.5^\circ$ and $\varphi=0^\circ$ is the optimum crystal orientation to have zero temperature coefficient, a small capacitance ratio, and no inharmonic overtone. Figure 9 shows the turnaround temperature plotted as function of the ratio of capacitances. The ratio of capacitances of 4.5° rotated plate is 70% of the X-cut plate for zero temperature coefficient.

3. Miniaturization of strip

3-1 Width of strip

On the piezoelectric strip, spurious responses depending on the thickness twist overtone cannot be generated [5]. Spurious responses due to the width of the strip cannot be avoided, as shown in Fig. 10. These spurious responses shift up to a higher frequency when the ratio of width to thickness of strip (W/H) becomes small. The value of W/H comes to the suitable ratio ($W/H=2.0$), the spurious responses

are separated by a large frequency from the main mode. Figure 11 shows the experimental frequency spectrum of the spurious responses plotted against the W/H ratio. The f_r indicates the resonant frequency constant, and f_a indicates the antiresonant frequency constant. When the ratio W/H is 1.5 to 2.0, or 2.8 to 3.1 (the hatched region in Fig. 11), we can obtain a piezoelectric strip resonator with a single response separated by an adequately large frequency from the spurious responses.

3-2 Length of strip

The displacement amplitude along the longitudinal direction is the highest at the center of the driving electrodes and decreases on either side of the electrodes region. The support on the edge of the strip therefore does not influence the resonance energy when the strip is long enough to dampen the displacement amplitude. To miniaturize the strip, we studied the length of the strip. Figure 12 shows the result of R_1 plotted as a function of the ratio of the strip length to the strip thickness. When the ratio of the strip length to strip thickness becomes large, R_1 decreases. When the ratio is over 10, R_1 is the lowest. The minimum length of the strip is 10 times of the strip thickness for 30Ω resonant resistance.

3-3 Application of wafer processing

The size of strip is so small and the mechanical energy so strongly concentrated around the driving electrodes, that many strips can be confirmed on one wafer as shown in Fig. 13; for example about 300 strips are made from a 2" wafer for 6 MHz. The driving electrodes are made by evaporated thin metal film and confirmed using photo-lithography. The terminal electrodes are pried on the wafer with Ag conductive resin. Each strip is separated by the cutting. The strips are mounted on a ceramic substrates printed with an electrical terminal and packaged by the ceramic cap.

4. Electrical characteristics

Figure 14 shows the admittance characteristics of the 6 MHz resonator. The size of strip is small-- $5.5 \text{ mm} \times 0.8 \text{ mm} \times 0.32 \text{ mm}$. The ratio of capacitances is 16. The strip is 20 times smaller than that of the AT-quartz. The mechanical Q factor is 16,000. The strip is mounted on a ceramic base and packaged as a "chip component", so that this LiTaO_3 resonator can be used for a hybrid integrated circuit, as shown in Fig. 15 which is an illustration of a voltage controlled oscillator using a variable capacitance diode. The variable frequency range is larger than 0.4% when an FC-53 variable capacitance diode.

5. Conclusion

This paper reported on a miniaturized LiTaO_3 strip resonator with stable temperature characteristics. We studied the most suitable relationship between the value of the trapped energy and crystal orientation so as to realize a zero temperature coefficient at room temperature and a small capacitance ratio. We next applied the piezoelectric strip to the LiTaO_3 resonator and studied the dimension of the strip to separate the spurious responses by a large enough frequency from the main mode and to miniaturize the strip.

These technique can be applied to resonators with resonance frequencies from 3 MHz to 30 MHz. The size of strip is small (for example, $5.5\text{mm} \times 0.8\text{mm} \times 0.32\text{mm}$ for 6 MHz). The ratio of capacitances is small (16) and the mechanical Q factor is large (16,000). These characteristics are suitable for a high stable voltage controlled oscillator with a wide range of variable frequencies. The mechanical energy is so strongly concentrated around the electrode that the strip can be mounted on a ceramic substrate without reducing the Q factor. Therefore, chip type resonators were realized which can be used as chip components.

Acknowledgements

The authors wish to thank Mr. Endo and Dr. Hara for their guidance and suggestions during this study and to thank Mr. Funayama for preparation of hybrid circuit.

References

- [1] A.W. Warner, M. Onoe, and G.A. Coquin, "Determination of Elastic and piezoelectric Constants for Crystals in Class (3m)", J. Acoust. Soc. Am., Vol.42, Nov, 1967, 1223-1231
- [2] A.W. Warner and A.A. Ballman, "Low temperature coefficient of frequency in a lithium tantalate resonator", Proc. IEEE (Letters), Vol.55, p.450, March 1967
- [3] M. Onoe and T. Ashida, "Zero temperature coefficient of resonant frequency in an X-cut lithium tantalate at room temperature", Proc. IEEE (Letters), Vol.57, No.8, p.1446, August 1969
- [4] J. Detainst and R. Lancon, "Temperature characteristics of high frequency lithium tantalate plates", Proc. 30th Annu. Symp. Freq. Cont., p.132-140, 1976
- [5] K. Nakamura and H. Shimizu, "On the equivalent circuits for thickness twist modes and thickness shear modes in piezoelectric plates", Paper of technical group US, IECE Japan 69-21, October 1969
- [6] T. Yamada, H. Iwasaki, and N. Niizeki, "Piezoelectric and elastic properties of LiTaO_3 : temperature characteristics", Vol.8, No.9, p.1127, September 1969
- [7] R.T. Smith and F.S. Welsh, "Temperature dependence of the elastic, piezoelectric constants of lithium tantalate and lithium niobate", J. Appl. Phys., Vol.42, p.2219, 1971
- [8] M. Onoe and H. Jumonji, "Analysis of piezoelectric resonators vibrating trapped energy modes", National Convention Record of IECE of Japan, Vol.48, No.9, p.1574, 1965
- [9] Y. Fujiwara, S. Yamada, and N. Wakatsuki, "Strip type LiTaO_3 chip resonator with zero temperature coefficient", Paper of technical group US, IECE Japan 82-46, November 1982

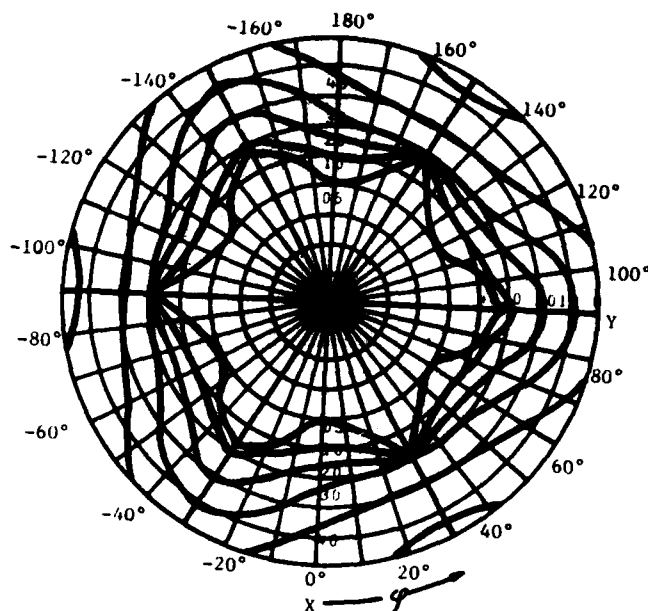


Fig.1. Coupling coefficient
(Thickness shear fast mode)

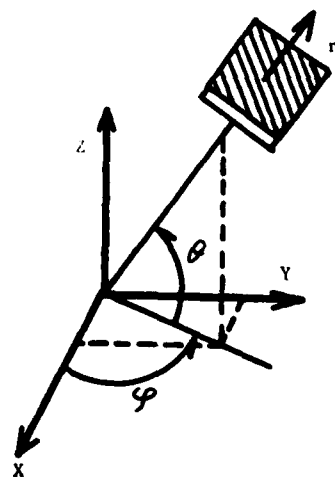


Fig.2. Plate orientation

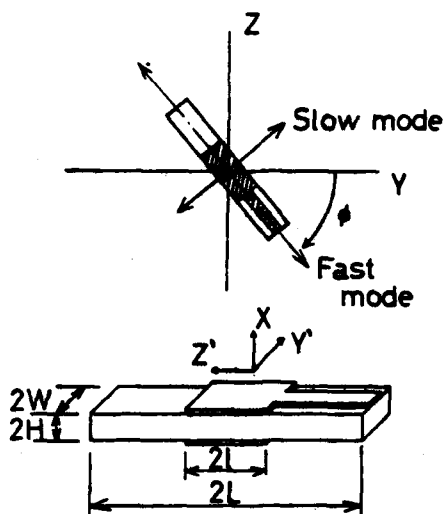


Fig.3. Piezoelectric strip construction

Table 1. Constants of LiTaO_3 X-cut plate

	Thickness extension	Thickness shear(fast)	Thickness shear(slow)
Coupling	0	0.47	0.06
Coefficient	0	0.45	0.04
Frequency	Hz·m 2865	Hz·m 2165	Hz·m 1692
Constant	Hz·m 2776	Hz·m 2106	Hz·m 1684
Diceplacement	—	-48.7°	+41.2°
Angle	—	-54.8°	+35.1°

Upper; Material constants by Yamada et. al.
Lower; Material constants by Smith et. al.

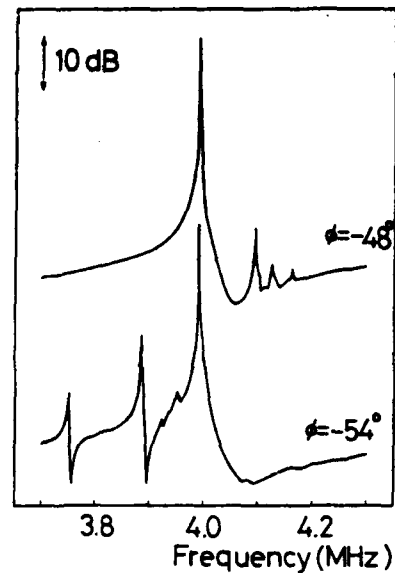


Fig.4. Admittance characteristics of resonators $\phi = -48^\circ$ and $\phi = -54^\circ$

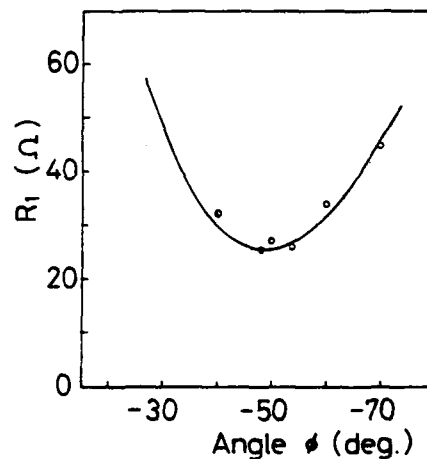


Fig.5. Relationship between angle ϕ and R_1

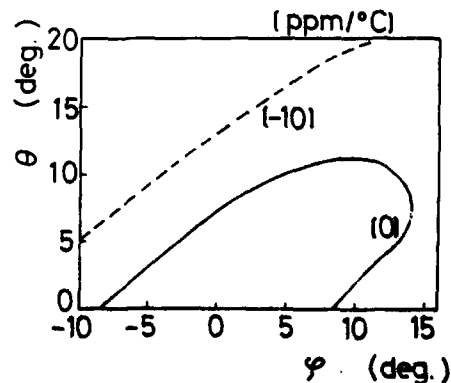


Fig.6. Temperature coefficient of resonance frequency for infinite electrode

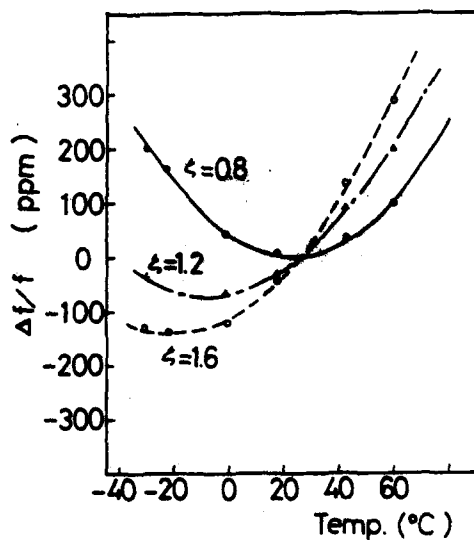


Fig. 7. Temperature characteristics of X-cut plate resonator

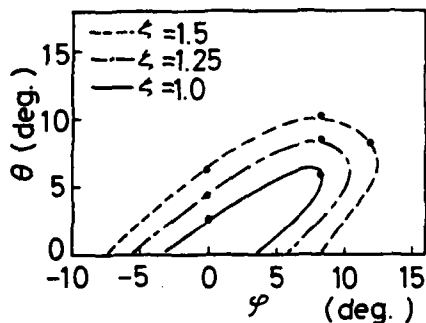


Fig. 8. Zero temperature coefficient for different values of ζ

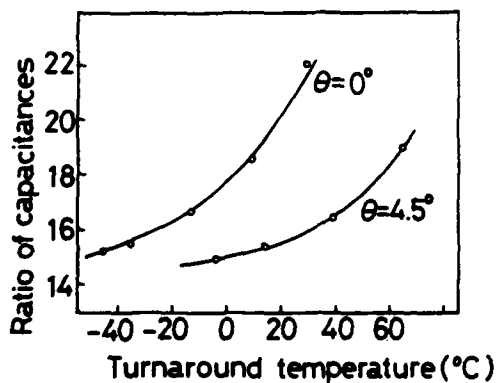


Fig. 9. Turnaround temperature as a function of the ratio of capacitances

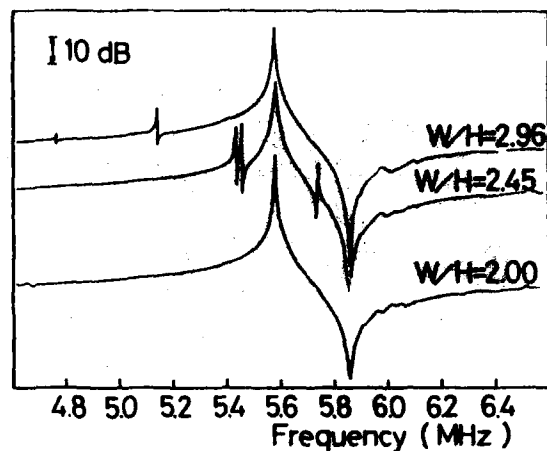


Fig. 10. Transfer of spurious responses due to the width of the strip

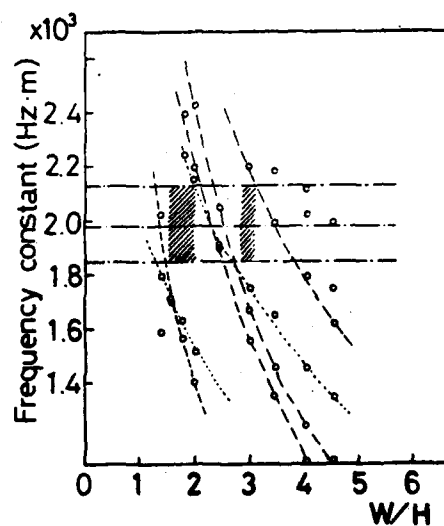


Fig. 11. Frequency spectrum of the spurious responses

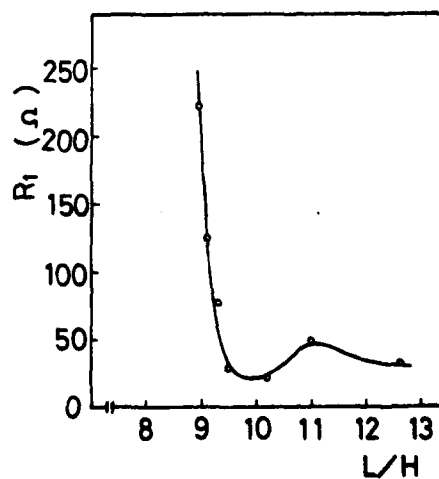


Fig. 12. Relationship between the ratio L/H and R_1

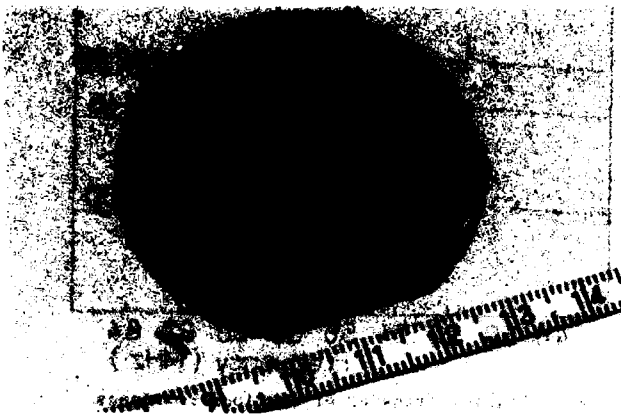


Fig.13. One LiTaO_3 wafer forms
about 300 strips

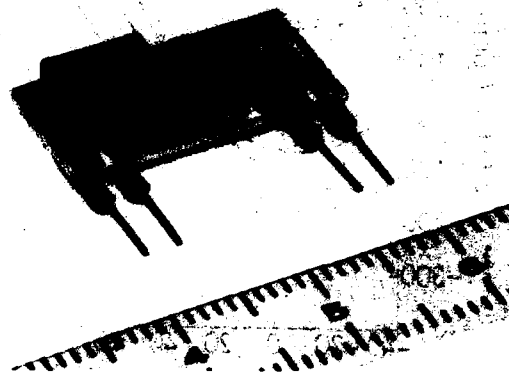
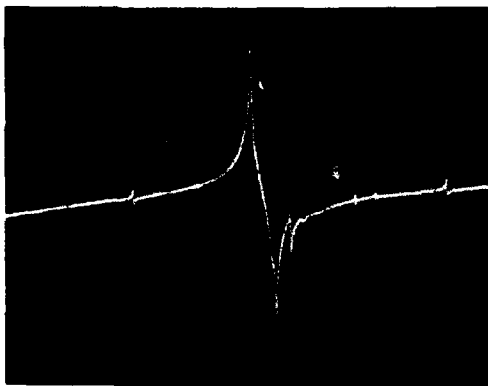


Fig.15. Voltage controlled
oscillator with
 LiTaO_3 resonator chip



Center; 6 MHz
Horizontal Scale; 0.5 MHz/div.
Vertical Scale; 10 dB/div.

Fig.14. Frequency response for
6 MHz (5.5mm x 0.8mm x 0.32mm)
resonator, $R_1=16\Omega$, $C_1=0.1\text{pF}$,
 $L_1=6.7\text{mH}$ and $Q=16,000$.



CURRENT TRENDS IN CRYSTAL FILTERS

R.C. Smythe & M.D. Howard

PIEZO TECHNOLOGY, INC.
ORLANDO, FLORIDA 32804

SUMMARY

Recent and projected developments in crystal filters are reviewed. Major growth areas are in VHF/UHF filters, data transmission filters, linearity, packaging, and cost reduction.

1. INTRODUCTION

This is a particularly appropriate time for a review of crystal filters. Continuing expansion of present applications and the opening up of new ones are providing incentives for major advances in the technology. The rapid growth of the field during the past 10-15 years, associated in large part with the advent of monolithic filters, has been outlined in some detail elsewhere [1]. The present paper will touch briefly on the most recent developments and (at some risk) likely future progress.

In the future, as in the past, crystal filters will be used primarily in communications and navigation systems. Particular growth areas are:

- Military
 - Spread-spectrum communications
 - Satellite-based navigation systems (NAVSTAR/GPS)
- Commercial
 - Cellular radio
 - Conventional/Trunked mobile radio
 - Amplitude-Companded sideband radio

The growth of these and other applications is creating new requirements in the following areas which are briefly discussed in section 2-6:

- VHF/UHF Filters
- Data Transmission Filters
- Linearity
- Packaging
- Cost Reduction

2. VHF/UHF FILTERS

Advances in the performance and frequency range of VHF and UHF crystal filters are motivated by a number of applications and are being obtained in at least three different ways.

First, the use of multi-resonator monolithic filters continues to expand. Figure 1 shows the attenuation characteristic of a 162 MHz, four-pole fully monolithic filter; it measures 11.0 x 4.7 x 11.4 mm., and has a terminating impedance of 3000 ohms. A somewhat similar filter has recently been introduced in a highly miniaturized paging receiver. Figure 2 shows the response of an eight-pole 45 MHz filter having two, four-pole monolithic sections. Filters of this type may make single conversion economical in 800 MHz conventional and trunked mobile radio systems.

Second, advances in wafer fabrication techniques allow the use of higher fundamental frequencies, making both higher frequencies and wider bandwidths available. For example, the development of cellular radio for the 800-900 MHz band has created requirements for VHF filters having bandwidths of about 30 kHz for use as first-i.f. filters. Figure 3 shows a four-pole filter now in production using two 45 MHz fundamental-mode monolithic two-poles. It seems likely that the next generation of equipment will use an even higher i.f. such as 90 MHz, figure 4.

The use of unconventional wafer fabrication techniques is extending the maximum fundamental frequency range still further. Using ion-milling, Berte and co-workers [2] have made AT-cut resonators at frequencies to 500 MHz. These resonators have been applied to filters [3] as well as oscillators. A drawback of the ion-milling process is its cost and low through-put. One alternative is chemical etching, using the ideas of Vig [4]. Figure 5 shows a 136 MHz fundamental-mode one-pole filter having a 3 dB bandwidth of 350 kHz made by a wet etching process.

Third, frequency and bandwidth may be extended upward by the use of high-coupling materials rather than quartz. Some use has been made of X-cut lithium tantalate [3]. Synthetic berlinite [5] appears even more promising than lithium tantalate if growth problems can be solved. Lithium tetraborate is a very recent candidate [6].

3. DATA TRANSMISSION FILTERS

Increasing use of digitally-encoded signals in such diverse applications as secure communications and paging systems is creating requirements for crystal filters having good pulse response as well as high selectivity. Conventionally, gaussian-approximation filters have offered the former, Tchebycheff and elliptic-function filters the latter. Neither is adequate for many new requirements.

For such applications the crystal filter designer needs to have a range of approaches at his disposal. Two are illustrated here -- one minimum-phase and the other non-minimum-phase. Both examples use tandem monolithic topology. Figure 6 shows a six-pole transitional Butterworth-Gaussian filter with added transmission zeros which offers a 2:1 delay improvement over a Butterworth filter having the same 60 dB/3 dB selectivity, at the expense of one resonator. The six-pole filter of figure 7 uses a Rhodes function [7] approximation of the type first applied to crystal filters by Herzog and Swanson [8]. For this filter, the group delay is almost constant over 65% of the passband. Though not yet optimum, the filter illustrates the power of the non-minimum-phase approach.

4. LINEARITY

Over the past decade there has been increasing awareness and understanding of nonlinear effects in crystal filters [9, 10, 11]. The most important forms of nonlinear behavior in the applications are intermodulation (of both in-band and out-of-band signals), increased phase noise (especially in frequency synthesis) and amplitude nonlinearity (especially in instrumentation.) Each new generation of systems results in increased restrictions on one or another of these, as the linearity of other system components is improved. Improvement in crystal filter linearity is achieved by process refinements as well as by improved device design. Figure 8 shows the amplitude linearity of a 20 MHz, four-pole monolithic filter. Passband attenuation varies less than .005 dB over a 30 dB range of signal level.

5. PACKAGING

Several present and future needs will influence the evolution of crystal filter packaging techniques: a need for further miniaturization, a need for cost reduction in high-volume applications, and needs for surface-mounting and hybrid-compatible filters. Only the first of these will be addressed here.

Conventionally, each resonator or monolithic filter unit is contained in its own HC-series or TO-series enclosure. This approach is easy to implement and can be used over the entire crystal filter frequency range. At high frequencies quite small filters can be made. Above 20 MHz HC-45-derived packages allow a high degree of miniaturization to be achieved, Fig. 9. An eight-pole tandem monolithic filter measures only 13.6x8.8x11.5 mm.

Single-wafer packages present a limit to achievable miniaturization which, potentially, can be circumvented by multi-wafer packages. The most notable example is the stacked-ring construction introduced by Sheahan [12]. A simpler approach, limited to two wafers, is shown in figure 10 [13].

6. COST REDUCTION

The widespread use of monolithic filters beginning around 1970 came about in part because of the economic advantage which they offered -- not only over discrete-resonator crystal filters but also over other forms of filtering. In system design there is ordinarily flexibility in choosing the frequency or frequencies at which the filtering function is to be performed, thus allowing different filter technologies to be used. The choice almost always includes economic considerations.

Hence, future growth of crystal filters, both discrete and monolithic, in many applications depends upon continued cost reduction. This will be achieved in part through economies of scale, but is heavily dependent on introduction of process automation and batch process techniques. Limitations in this regard are the present lack of product standardization and the fragmented character of the market.

REFERENCES

- 1) Smythe, R.C., "Precision Frequency Control," (E.A. Gerber & A. Ballato, Eng.), sec. 5.2.; in Press.
- 2) Berte, M., "Acoustic Bulk Wave Resonators & Filters Operating in the Fundamental Mode at Frequencies Greater than 100 MHz," Proc. 31st AFCS, pp. 122-125; 1977.
- 3) B. D'Albaret & P. Siffert, "Recent Advances in UHF Crystal Filters," Proc. 37th AFCS, pp. 405-418; 1982.
- 4) Vig, John R., John W. LeBus, & Raymond L. Filter, "Chemically Polished Quartz," Proc. 31st AFCS, pp. 131-143; 1977.
- 5) Chang, Z.P. & Barsch, G.R., "Elastic Constants & Thermal Expansion of Berlinite," IEEE Trans., vol. SU-23, no. 2, pp. 127-135; 1976.
- 6) Emin, C.D.J., & J.F. Werner, "The Bulk Acoustic Wave Properties of Lithium Tetraborate," These Proceedings.
- 7) Rhodes, J.D., "A Low-Phase Prototype Network for Microwave Linear-Phase Filters," IEEE Trans., vol. MTT-18, no. 6, pp. 290-301; 1970.
- 8) Herzig, P.A. & T.W. Swanson, "A Polyolithic Crystal Filter Employing a Rhodes Transfer Function," Proc. 32nd AFCS, pp. 233-42; 1978.
- 9) Horton, W.H. & R.C. Smythe, "Experimental Investigations in Monolithic Crystal Filters," Proc. 27th AFCS, pp. 243-5; 1973.
- 10) Tiersten, H.F., "Analysis of Intermodulation in Rotated Y-Cut Quartz Thickness-Shear Resonators," Proc. 28th AFCS, pp. 1-4; 1974.
- 11) Smythe, R.C., "Intermodulation in Thickness-Shear Resonator," Proc. 28th AFCS, pp. 5-7; 1974.
- 12) Sheahan, D.F., "Polyolithic Crystal Filters," Proc. 28th AFCS, pp. 120-127; 1975.
- 13) Smythe, R.C., "Some Recent Advances in Integrated Crystal Filters," Proc. 32nd AFCS, pp. 220-232; 1978.

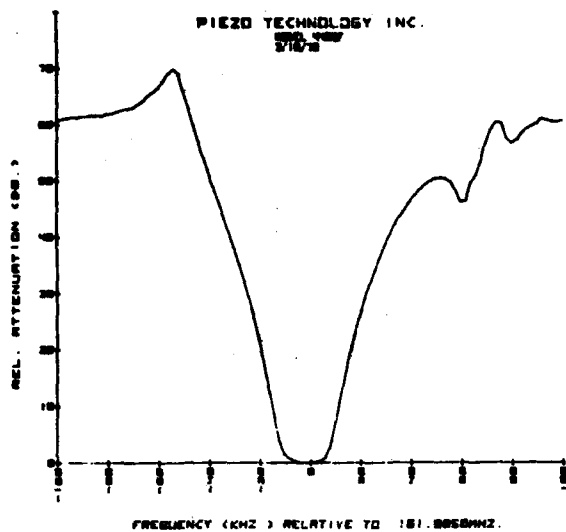


FIGURE 1 Four-Pole, 162 kHz Resonant Filter, Attenuation Characteristics

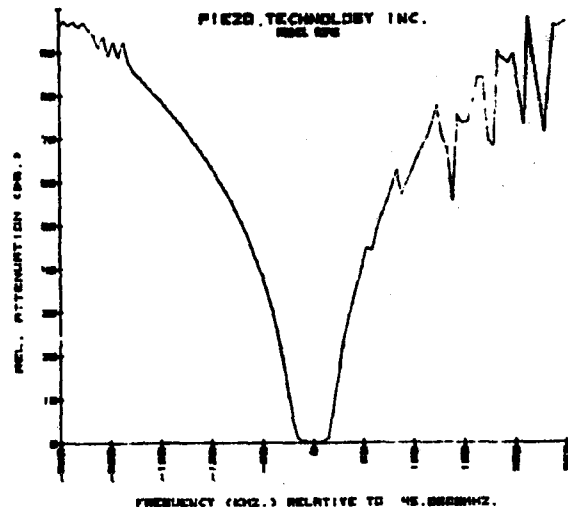


FIGURE 3 Four-Pole, 46 kHz Tapered Resonant Filter, Attenuation Characteristics

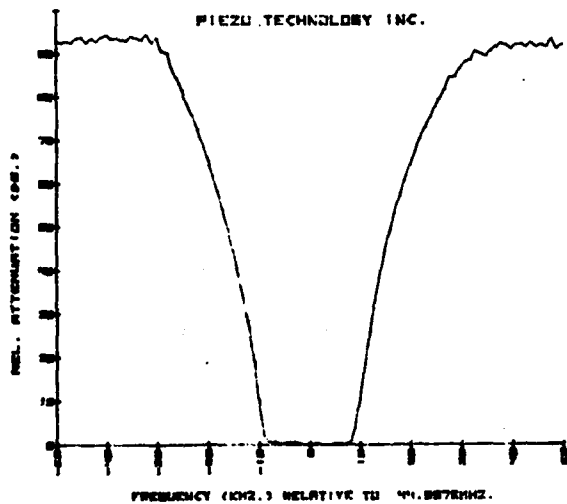


FIGURE 2 Eight-Pole, 46 kHz Tapered Resonant Filter, Attenuation Characteristics

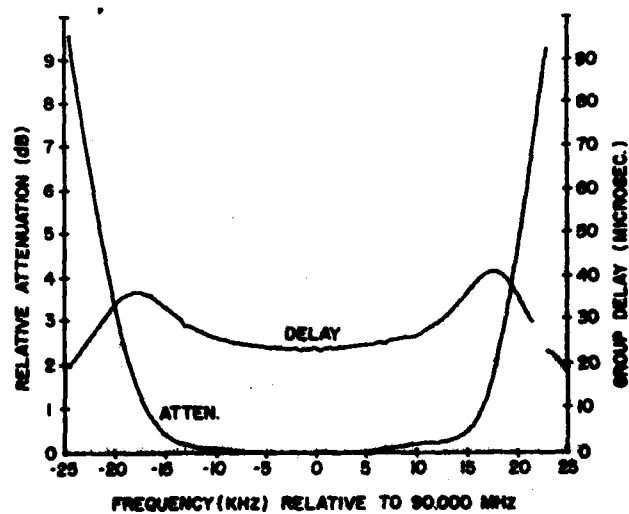


Figure 4: Four-Pole, 90 kHz Resonant Filter, Passband Attenuation & Group Delay

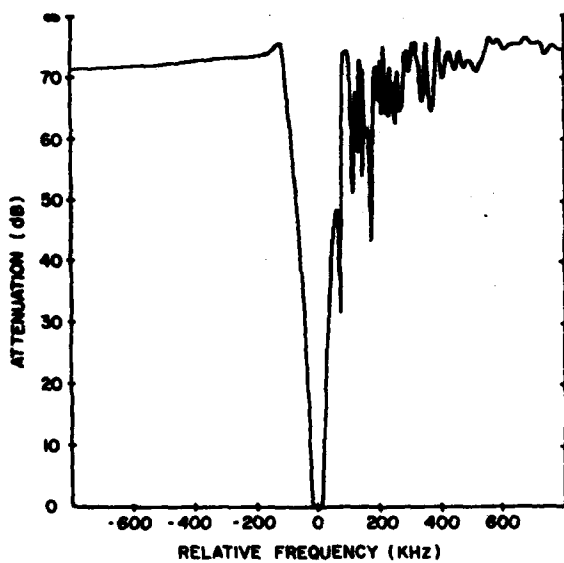


Figure 4. Four-Pole, 90 MHz Resonant Filter, Stopband Attenuation

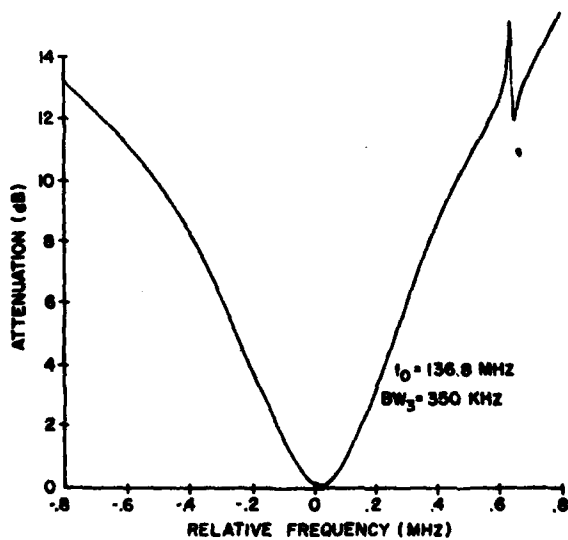


Figure 5. Single-Pole, 100 MHz Crystal Filter, Attenuation Characteristics

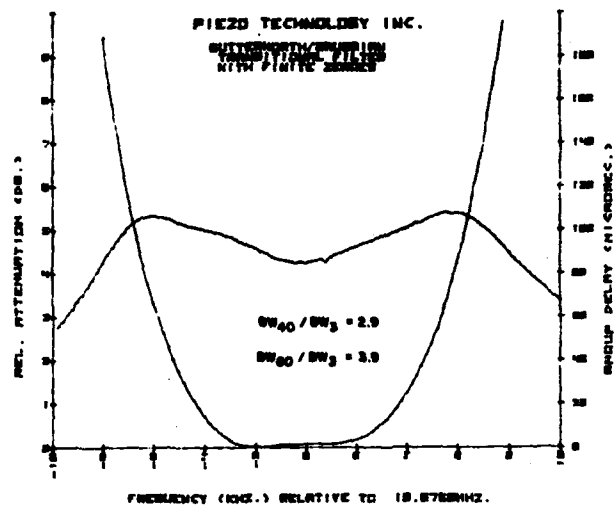


Figure 6a. Six-Pole Butterworth/Chebyshev Transitional Filter, Passband Attenuation & Group Delay

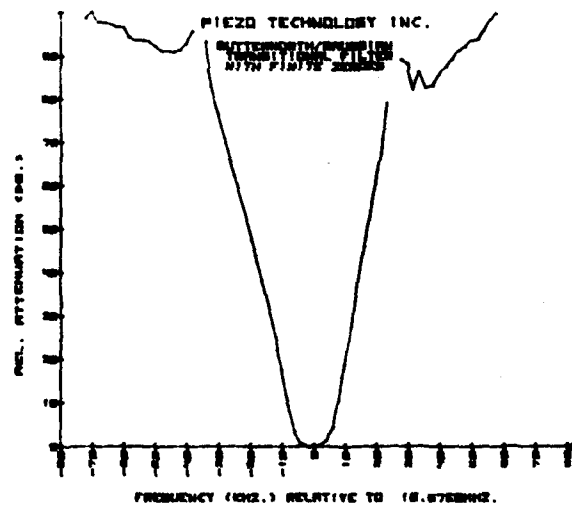


Figure 6b. Six-Pole Butterworth/Chebyshev Transitional Filter, Stopband Attenuation

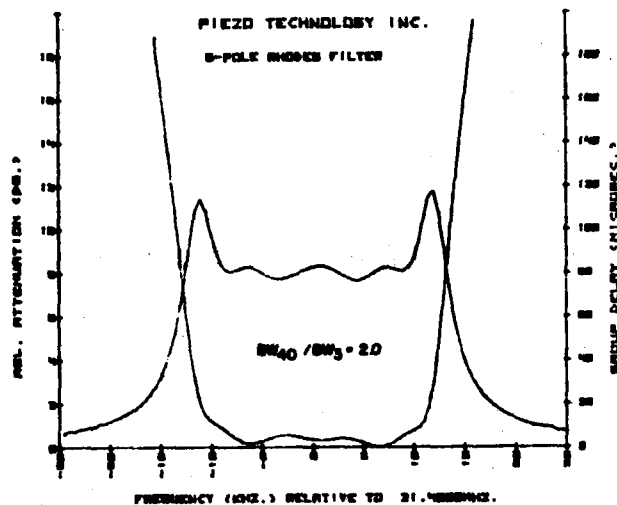


Figure 7a 5-Pole Rhodes Function Linear Phase Filter, Parabolic Attenuation & Group Delay

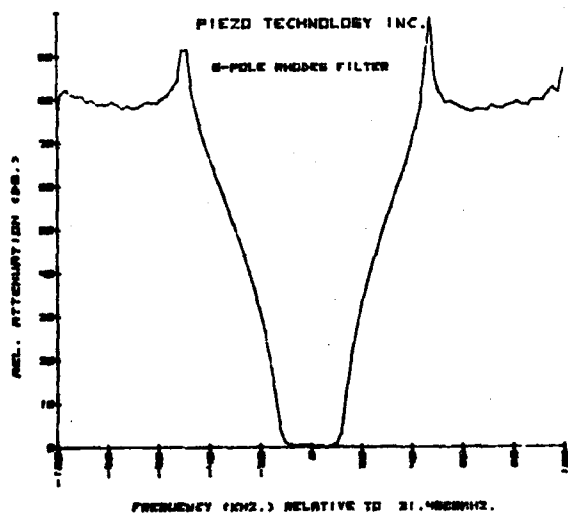


Figure 7b 5-Pole Rhodes Function Linear Phase Filter, Stepped Attenuation

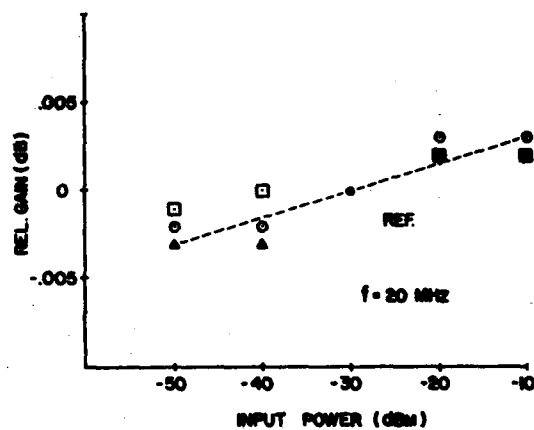


Figure 8 Amplitude Linearity, 20 MHz, Four-Pole Tannin Resonator Filter

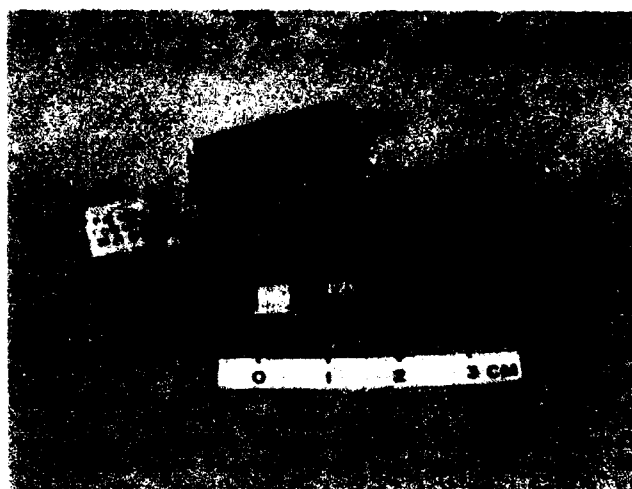


Figure 9 Ultraminiature "A" Series Resonator & Tannin Resonator Filters.

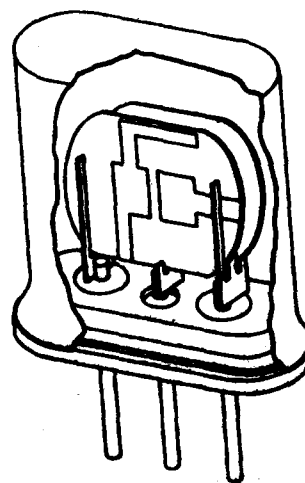


Figure 10 Step-Up Two-Stage Microinterpolation Scheme

SURFACE ACOUSTIC WAVE BANDPASS FILTERS

Clinton Hartmann and Stephen Wilkus
R.F. Monolithics, Inc.
Dallas, Texas 75234

ABSTRACT

The physics of SAW bandpass filters will be reviewed emphasizing those aspects which directly impact filter performance. Next, some of the design trade-offs in these filters will be discussed including insertion loss versus fractional bandwidth, skirt steepness versus device size and cost, and the characteristics of various substrates which are commonly used. Examples of various filters will be given including filters for television IF, satellite receivers, and data communications systems. Generalized performance curves achievable with SAW bandpass filters will be presented. Projections of future developments including SAW resonator bandpass filters and single-phase unidirectional transducer filters will be discussed.

INTRODUCTION

The Surface Acoustic Wave (SAW) field has matured to the point that at least two distinct generations of bandpass filter devices are apparent. This paper will review the principles of operation of these two generations of devices with special emphasis on the design trade-offs that should be considered in specifying SAW filters. In addition, future developments and the beginnings of the third generation of SAW filters will be discussed.

Section I will review the principles of operation and the various implementations of filters. It will discuss the design approaches and modeling techniques and derive the equations for the input impedances of a SAW filter. Section I will conclude with a discussion of the effects which degrade filter performance, how designers deal with them, and what to expect in terms of filter performance.

Section II will illustrate SAW filter capabilities through examples of four production devices. Limitations and capabilities will be pointed out, with particular attention given to how consideration of these limitations and capabilities at the system design stage has made best use of SAWs and has provided improved system performance.

Section III will present some early results of work underway to develop the third generation of SAW filters. Most notably, single-phase unidirectional transducers (SPUDT), and coupled resonator filters will be shown.

1. PRINCIPLES OF OPERATION

Figure 1 illustrates the configuration of a typical first generation SAW filter. It consists of a piezoelectric substrate such as quartz or LiNbO_3 , on which a metalized electrode pattern (typically an Aluminum film a few thousand Angstroms thick) is photolithographically produced. The electrode pattern allows an electric field to be applied to the substrate, which, being piezoelectric develops a strain which manifests itself as an acoustic wave propagating away from the transducer. In selected cuts of materials, a surface wave (as opposed to a bulk wave) is excited which propagates along the surface of the substrate.

FIRST GENERATION SAW DEVICE

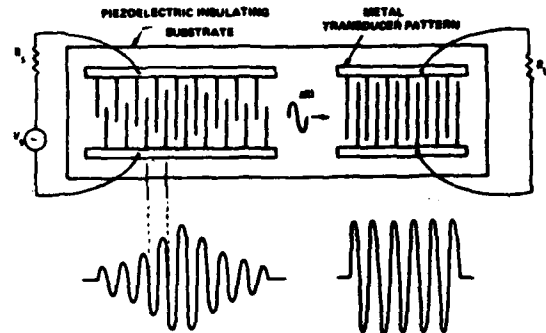


Fig 1: Illustration of first generation SAW device showing the relationship between electrode structure and impulse response.

The filtering properties of SAW devices arise from the fact that the degree of coupling to waves of a desired frequency depends upon the size of the overlap integral of the desired wave's surface potential and the charge distribution of the electrode pattern. Consequently a transducer can selectively excite waves of a controlled frequency band by controlling the pattern of the electrodes. If the electrodes are closely spaced, the frequency goes up, broader spacing lowers the frequency. Use of a large number of electrodes generally narrows the bandwidth, whereas decreasing the number of electrodes generally widens it.

The surface wave is essentially non-dispersive, and has little propagation attenuation. Since it propagates along the surface it can be transversely tapped, hence it can be modeled as a transversal filter in the same way as tapped delay lines, CCD filters, and digital finite impulse response filters (collectively known as Finite Impulse Response Filters or FIR filters.)

Unlike the classical FIR filters, the surface wave generated by one transducer is sampled and thus filtered by a second, output, transducer. Utilizing the fact that the propagating wave has accompanying it a surface potential arising from the piezoelectric effect. As a result, a SAW filter is essentially a cascade of two FIR filters where the two filters are the two transducers.

Impulse Response Model

Some comments are in order here concerning FIR filter design, which will make clear what drives the length and cost of a filter and will show that inband ripple arises not only from time spurs but also from the theoretical limitations on finite length.

A periodically sampled transversal filter such as a SAW transducer has a frequency response given by[1]:

$$H(f) = \sum_{i=1}^N h_i e^{-j2\pi i f / f_{\text{samp}}} \quad (1)$$

Where there are N periodically sampled taps at time $t = i/f_{\text{samp}}$ (f_{samp} is the so-called sampling frequency). $H(f)$ is the frequency response of the transducer. It goes without saying that the previous discussion holds true for the second transducer as well, and that the combined response consists of the product of the two transducers' responses.

Several things are worth noting at this point: $H(f)$ is periodic with a period of f_{samp} . Thus, if a passband is located at frequency f_p , then a passband is also located at $f_p + f_{\text{samp}}$. So, if we have sampled at 4 times the center frequency of the passband (a sample rate commonly used in first generation devices because 4 electrodes per wavelength cancel reflections off the electrode edges) then we can expect a fifth harmonic response. In addition, the image frequency at $-f_p$ will generate another response at $-f_p + f_{\text{samp}} = 3f_p$, which is a third harmonic response. In general, a transducer with 4 electrodes per wavelength will generate all odd harmonics. Three electrodes per wavelength will generate the 2nd, 4th, 5th, etc., harmonics.

Another point to make about equation (1) is that if we impose the restriction that the samples have even symmetry ($h_i = h_{N-i+1}$) then we have for the case of odd N :

$$H(f) = e^{-j2\pi f(N+1)/2f_{\text{samp}}} h_{1+N/2} + \sum_{i=1}^{N/2} 2h_i \cos \frac{(2\pi f(i - \frac{N+1}{2}))}{f_{\text{samp}}}$$

That is, that $H(f)$ has linear phase. Thus, SAW devices can be designed with linear phase. In fact, the phase and amplitude characteristics of SAW filters can (largely) be specified independently if the time symmetry restriction is dropped.

If in equation (1) we replace the exponential term with Z , we have the well known Z-transform:

$$H(f) = \sum_{i=1}^N h_i Z^i \quad (2)$$

$$\text{Where } Z = e^{-j2\pi f / f_{\text{samp}}}$$

which is simply a polynomial in Z of order $N-1$. Because Z is a trigonometric function of f , this expression is also referred to as a trigonometric polynomial in f . As with any polynomial, the roots of this trigonometric polynomial can be found. Stopband nulls correspond to zeros on the unit circle in the Z -plane and passband ripples correspond to zeros lying near the unit circle in the vicinity of the passband response. A desired time response can be broken into two time responses by factoring the polynomial and simply assigning the zeros to one or the other of the transducers. In this way, it is possible to arrive at a

¹Later we will introduce the frequency contribution of the "element response", a Green's function accounting for the fact that the non-zero width of the electrodes are not precisely modeled by an impulse response. This "element factor" along with the effects of matching circuits can strongly effect and even null out a harmonic response. Nevertheless, it is common for SAW filters to exhibit harmonic responses.

pair of transducers which when working together in cascade yield an optimized response.

Filter Length

So much for the easy part, getting $H(f)$ given h_i is no problem. Getting the optimal time response that will meet certain frequency constraints is. Fortunately, many tools have been developed by the digital signal processing field for doing just that. The most successful being an algorithm developed by Remez and extended by Parks and McClellan[2] which will take a frequency specification of desired values and relative ripple or rejection levels and will produce an impulse response of a given length which will minimize the absolute ripple levels. The designer then modifies the filter length until the program gives the absolute ripple level desired.

One might guess that the filter length would simply go as the inverse of the transition bandwidth, but a more accurate relationship has been determined by Rabiner[3]. By compiling hundreds of results of the trial and error approach for low pass prototypes Rabiner was able to empirically fit the following equation:

$$N = \frac{D(\delta_1, \delta_2)}{\Delta F} - \xi(\delta_1, \delta_2) \Delta F + 1$$

Where:

$$D(\delta_1, \delta_2) = \{a_1 (\log_{10} \delta_1)^2 + a_2 \log_{10} \delta_1 + a_3\} \log_{10} \delta_2 + a_4 (\log_{10} \delta_1)^2 + a_5 \log_{10} \delta_1 + a_6$$

$$\xi(\delta_1, \delta_2) = b_1 + b_2 (\log_{10} \delta_1 - \log_{10} \delta_2)$$

δ = The ripple level in the passband and stopband.

$$\begin{aligned} a_1 &= 5.309 \times 10^{-3} \\ a_2 &= 7.114 \times 10^{-2} \\ a_3 &= -4.761 \times 10^{-1} \\ a_4 &= -2.66 \times 10^{-3} \\ a_5 &= -5.941 \times 10^{-1} \\ a_6 &= -4.278 \times 10^{-1} \\ b_1 &= 11.01217 \\ b_2 &= 0.51244 \end{aligned}$$

This expression is quite accurate for symmetrical band-pass filters and can be used to give the physical length of a filter given the SAW velocity, V , for that material:

$$L = N V / f_{\text{samp}}$$

In actual practice, about 2 extra millimeters are added for dicing, handling, and acoustic absorber. Figure 2 plots this length as a function of the transition bandwidth for a narrowband filter centered at 100MHz and operating on a quartz substrate with a velocity of about 3.14 mm/usec. The parameter is the rejection level for a fixed inband ripple of 0.5dB peak to peak. As the length increases, the packaging and handling costs increase which are reflected in a very approximate way illustrated on the right side of the figure. It goes without saying, that actual pricing is highly dependent upon the size of the market and the non-recurring costs, but the cost factors should provide an insight into the effects of the design parameters on the final costs.

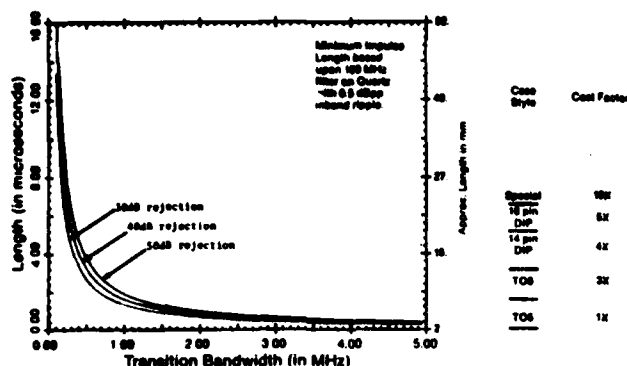


Fig 2: An idealization of the problem of filter length and cost to illustrate the effects of transition bandwidth and rejection level. The paradigm is based upon a 100MHz filter with 0.5dB peak to peak inband ripple on quartz. with 2mm of extra length for absorber and handeling. The family of curves are for different rejection levels. The curves are based upon the length equation discussed in the text and should be used with caution.

Element Factor

The previous discussion is quite general and does not distinguish SAWs from pairs of CCD filters or digital FIR filters in any way. The first and most obvious distinction is that a SAW device is an analog device which does not have perfect impulses as samples. The charge distribution on the electrode peaks at the two electrode edges instead of having an impulse in the electrode center. Consequently, the impulse response must be modified to represent the actual driving function of an individual electrode (in a uniform array). At center frequency, with 4 electrodes per wavelength, this amounts to a slight tilt to the passband, but harmonic responses are greatly modified. At these higher frequencies, the element factor also becomes sensitive to the metalization ratio. As a result, harmonic designs only slightly reduce the photolithographic tolerances over fundamental designs.

To summarize to this point, then; the overall filter response can be represented by a product of the the frequency responses of each transducer, which in turn is the product of the element factor and it's array factor calculated from the impulse response model. In addition, there is an impedance matching circuit response which is dependent upon the system impedance level and the equivalent circuit of each transducer. We consider this next.

Equivalent Circuit

SAW transducers appear principally as a somewhat lossy capacitor. The capacitance increases with the number of interdigitated electrodes, the overlap of the electrodes, the metalization ratio, and the effective dielectric constant of the substrate cut. The total capacitance is given by:

$$C_T = N W C_s$$

where W is the beamwidth and C_s is dependent upon the substrate cut, the metalization ratio, the effective dielectric constant, and the electrode connections (there is no capacitance between electrodes of the same potential.)

Connected in parallel with the capacitance, and accounting for the acoustic radiation, is a conductance

term, G_a , which for a uniform beamwidth transducer, can be found from energy conservation to be [4]:

$$G_a(f) = 8K^2 f_0 C_s W H(f)^2 F^2(f/f_0, u)$$

Where: k is the material dependent coupling constant, and $F(f/f_0, u)$ is the element factor which is dependent upon both the frequency and the metalization ratio, u .

For a transducer that does not have a uniform beamwidth, the actual G_a is decreased somewhat and can be evaluated by analyzing the transducer as many "tracks" of uniform transducers working in parallel.

For a uniform transducer with N electrode pairs, the radiation conductance is approximately given by:

$$G_a(f) = 8K^2 f_0 C_s W N^2 \frac{\sin(X)}{X}^2$$

$$\text{Where } X = 2\pi N(f - f_0)/f_0$$

and f_0 is the center frequency of operation.

Notice that as the material coupling constant, k , decreases (say, as we go from LiNbO_3 to Quartz) then G_a

decreases as k^2 . G_a also goes as N^2 , and linearly with the width.

In an unweighted transducer where all tap weights have the same strength, the bandwidth goes inversely with the number of taps, N . Consequently, G_a goes inversely with the square of the fractional bandwidth. Recalling that the transducer capacitance goes as N , or inversely with the fractional bandwidth, we know that the transducer Q will go as the fractional bandwidth. A more carefull analysis where the frequency response function is approximated by a rectangular function whose energy must equal the energy in a time function which is also a rectangular function, yields an expression for Q of:

$$Q = \frac{\pi}{4K^2} \frac{BW}{f_0}$$

where BW is the noise bandwidth of the transducer. This expression is a good approximation even for apodized transducers. This expression makes clear that as the fractional bandwidth increases, the Q increases, possibly to the point that it becomes the limiting factor of the bandwidth. That is, when:

$$\frac{BW}{f_0} = \sqrt{\frac{4K^2}{\pi}}$$

the bandwidth is limited by the electrical Q of the transducer and must be compensated for by using resistive loading, thus introducing losses. These losses and the critical bandwidth at which the losses arise are illustrated in figure 3. Notice how higher coupling materials such as LiNbO_3 can provide low loss over a wider fractional bandwidth than can Quartz. This particular figure illustrates loading losses only and not apodisation loss, bulk wave, or other losses. Also, for bidirectional devices there is an additional 6dB of loss. As a result of these loading losses, wide bandwidth devices are best suited for high center frequencies of operation.

One further complication must be added to the equivalent circuit, a reactance B_a must be included to insure causality. As it stands, a unit impulse of voltage would generate a current given by the inverse Fourier transform of the complex acoustic admittance.

UNIVERSAL INSERTION LOSS CURVE

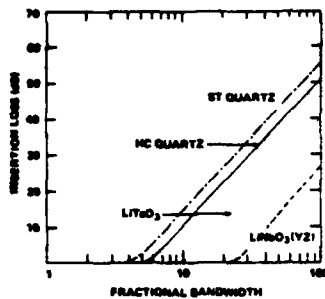


Fig 3: Insertion loss versus fractional bandwidth curve based upon the Q loading required to maintain a given bandwidth as discussed in the text. No bidirectional or other losses are considered.

To insure that it only have positive time components we need for $Y_{in}(f) = G_a(f) + j2\pi f C_T + jB_a(f)$. The

condition that $F^{-1}(Y_{in})$ be causal requires that $B_a(f)$ be the Hilbert transform of $G_a(f)$.

For the uniform transducer case:

$$B_a(f) = G_a(f_0) \frac{\sin(2X) - 2X}{2X^2}$$

$$\text{Where: } X = N\pi(f - f_0)/f_0$$

$B_a(f)$ peaks at $f - f_0 = f_0/2N$ (half the distance to the first null of G_a about 4dB down from the peak) with a value of $G_a(f_0)2/\pi$ on the low frequency side and the negative of that on the high frequency side. For all symmetric responses $B_a(f_0) = 0$.

Ignoring bulk, parasitic, and apodization losses an untuned transducer will have an insertion loss of:

$$IL = -10 \log_{10} \frac{2G_a G_L}{(G_L + G_a)^2 + (2\pi f C_T + B_a)^2}$$

for a bidirectional transducer and a load inductance of G_L . (The total insertion loss of the device is the sum of the losses for each transducer.)

TABLE I

SAW BANDPASS FILTER DEGRADATION EFFECTS

Effect	Time Signature	Freq Signature	Remedy
• Triple Transit	Trailing Pulses @ 2T	Amp. & Phase Ripple @ 1/2T	High Loss (Don't!) Good Match (Don't!)
• Spurious Reflections	Trailing Pulses @ Various Times	Amp & Phase Ripple @ 1/T time	Careful Design & Fabrication
• Diffraction	Slight Pulse Smearing	Degraded Skirt Shape on High Freq Side	Compensation
• Bulk Modes	Leading and Trailing Pulses	Extra Responses from $-1.5 f_0$ Up	Substrate Choice, Multi Strip Coupler, Backside Reengineering
• Matching Network Distortion	Slight Pulse Smearing	Slight Amp & Phase Distortion, Reduced Rejection	Compensation

In actual practice several difficulties limit the applicability of the impulse response model. No account has yet been made of diffraction, bulk waves, regeneration, matching circuit, or triple travel. Table 1 lists the more common of these spurious effects and their impact in both the time and frequency domains. Illustrating selected effects is figure 4. An idealized time response to a narrowly gated RF burst is shown in figure 4a. Notice a small RF burst leaking through the device at $t=0$. This feed through or cross-talk arises from limited isolation and can be solved with improved packaging and mounting. In the frequency domain, cross-talk is seen as a sloping limit to the ultimate rejection in figure 4b. As time increases, the desired time response is observed, in this case a 10% bandwidth filter centered at 100MHz with 50dB of rejection and 1dB peak to peak inband ripple. The skirts are 1MHz wide yielding a shape factor of 1.2.

At various times bulk waves appear, here shown idealized and centered at 2.25uSec. Bulk waves have a higher velocity than the surface waves but (in the case illustrated) can bounce off the bottom of the chip taking a longer path than the SAW does. In the frequency domain, the bulk waves appear on the high frequency side about 10-20% higher in frequency than the lowest frequency in the passband and at an amplitude commonly in the 40 to 50dB range.

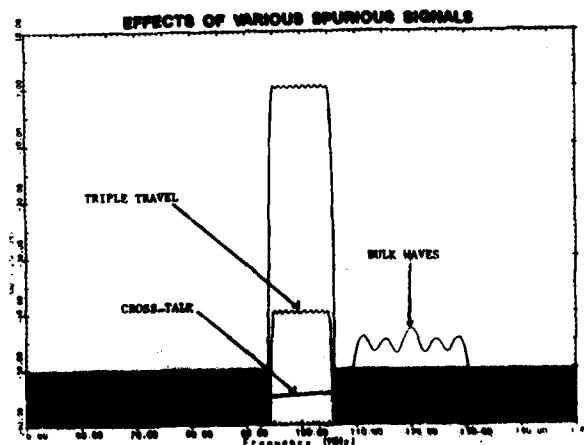
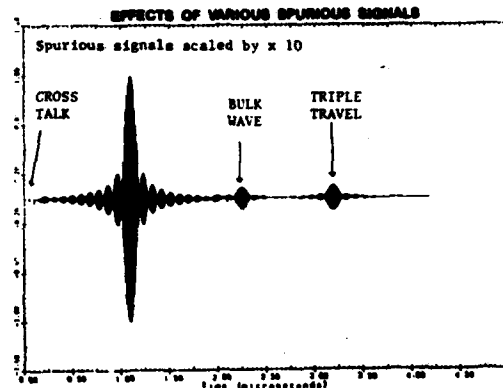


Fig 4: Idealized effects of spurious signals. a) shows the effects of bulk waves in the time domain where the spurious signals have been scaled by a factor of 10 for visibility. b) illustrates the effects the spurious signals have in the frequency domain.

At a time three times the group delay, the triple travel signal can be seen. As a wave encounters a transducer some of its energy is absorbed in the electrical load but some is also regenerated as a reflected SAW wave. For the untuned case, the acoustic reflection off of a transducer is given by :

$$R = 10 \log_{10} \frac{G_a^2}{(G_L + G_a)^2 + (2\pi f C_T + B_a)^2}$$

The triple travel arises when the desired signal reflects from the output transducer and reflects a second time from the input transducer to reappear at the output after having thus traveled the propagation distance three times. Thus the triple travel strength relative to the main response will be the sum of the reflections off the two transducers. In many cases of interest, the triple travel suppression relative to the main response is approximately given by: $TT = IL + 6dB$.

Similarly, there are fifth, seventh, ninth etc. transit signals that can contribute in and out of phase with each other. In the idealized frequency response shown, the device was assumed to have 15dB of insertion loss so that the triple travel signal was suppressed about 36dB. This contributes to a 0.28dB pp inband ripple and a 2° phase ripple.

Notice that such ripple due to time spurs are of a wholly different origin than the amplitude ripple designed in the main response. An FIR filter can have amplitude ripple (sometimes called Gibbs ripple) arising from the finite length of the time response. But notice that the Gibbs ripple does not accompany phase ripple nor does it accompany time spurs. Consequently, it can be important to specify amplitude and phase ripple separately or to specify time spurious suppression.

Second Generation

One way to reduce the insertion loss is to introduce a third phase in the electrode pattern which will excite a SAW wave in only one direction. Just as in CCD filters or even theatre marquees, three separate phases are the minimum needed for unidirectionality and good unidirectional performance is achievable over moderate bandwidths.

Figure 5 demonstrates an implementation of the three-phase air gap crossover unidirectional transducer. The air gap is needed to drive the three independent phases on a planar structure.

SECOND GENERATION SAW DEVICE

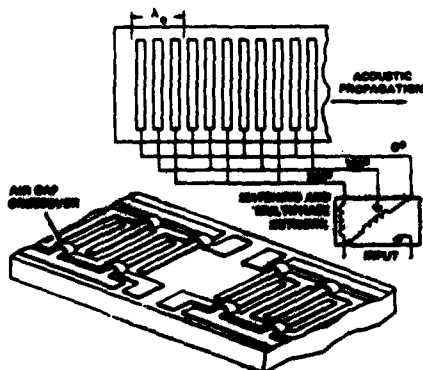


Fig 5: Air-gap crossover unidirectional device structure representing the second generation of SAW filters.

Application of this unidirectional scheme allows the 6dB bidirectional losses to be eliminated, but more importantly, the triple travel signal can be well suppressed and as a result, the device need not be mismatched to suppress the triple travel, in fact, minimizing the insertion loss maximizes the triple travel suppression.

II. FILTER EXAMPLES

As an illustration of the application of second-generation filters, figure 6 shows a superposition of spectrum analyzer data for some 10 channel 4 bandpass filters. In a "simplified" matching circuit which has a capacitor, two fixed coils and one variable coil this device exhibits an insertion loss of less than 8dB. In a fully matched circuit less than 6dB insertion loss is possible. This filter and others similarly designed, for the channel 2 and channel 3 bands are over 5.5MHz wide at the 1dB points and have over 25dB rejection of the adjacent audio and picture carriers 1.25MHz away from the 1dB points.

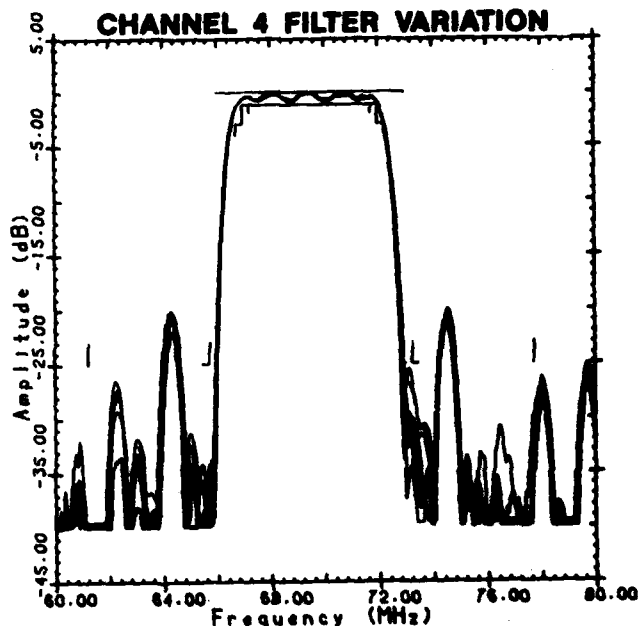


Fig 6: Overlay of some 10 channel-4 bandpass filters.

In order to meet stringent cost constraints, this filter was made to fit on a LiNbO₃ substrate measuring 10.4 x 2.8 mm on a side. Since the two transducers had to be in-line no multistrip coupler could be used. Also, conventional withdrawal weighting could not provide the performance needed. Instead, a finer gradation of withdrawable tape were generated by oversampling.

A more conventional, first generation, filter is presented in figure 7 which shows the reproducibility of 15 digital modem filters incorporated in the Scientific Atlanta model 6402 broadband modem. Using a QASK-16 modulation scheme, this IF filter has a very tight phase linearity spec. of $\pm 2^\circ$ and an amplitude ripple of 0.3dB. As shown on the expanded scale of figure 7a, it can be seen to meet these very tight specs on a medium scale production basis. This filter has a 1dB bandwidth of 720KHz centered at 190.125MHz and with an insertion loss of 23dB.

A third example is an IF filter for SATV receiver application, shown in figure 8. This second generation filter achieves an 8dB insertion loss and a 1dB bandwidth of 30KHz. Notice that at a center frequency

REPEATABILITY OF TYPICAL FILTER RESPONSE

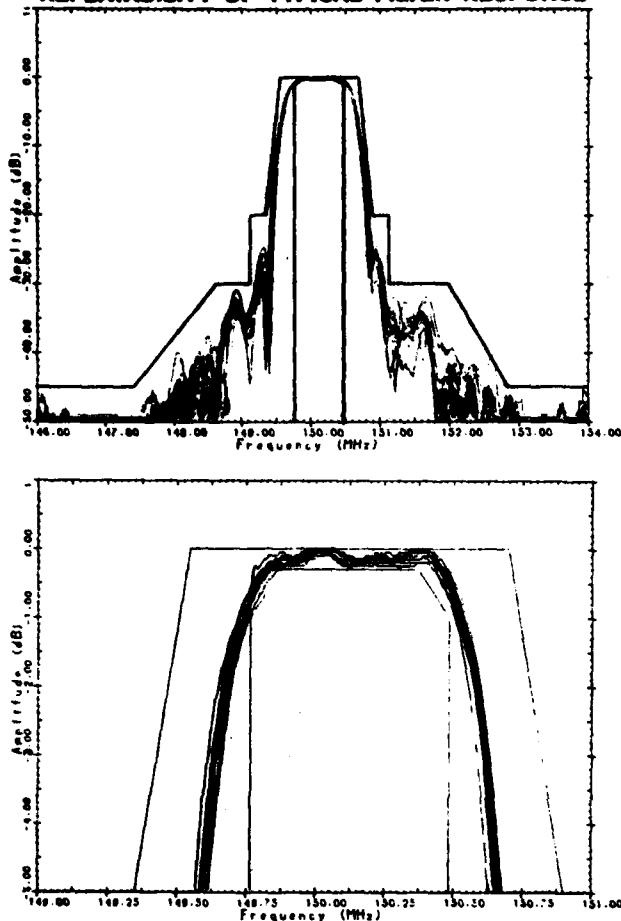


Fig 7: The repeatability of a first generation filter is illustrated for a Digital Modem application. The device has 23dB of loss at 150.125MHz and a 1dB bandwidth of 720KHz.

of 596MHz this filter has a 5% fractional bandwidth, if this filter had been designed to operate at, say, 70MHz, it would have a 44% fractional bandwidth. Such large fractional bandwidths are problematic because of restricted flexibility of weighting techniques, interference from bulk mode and diffraction distortion, difficult impedance matching, and increased insertion loss. Hence, for this and other reasons, SAW filter performance is often improved with higher center frequencies.

Figure 9 illustrates a filter designed under contract with the FCC as an alternative TVIF stage. Operating at 448.725MHz, and with a 3dB bandwidth of 4MHz, this filter exhibits a loss of 6.5dB and 50dB of rejection on Quartz. This filter also illustrates the ability of SAW filters to have a complicated and independent amplitude and phase characteristic, this filter has a designed-in phase response which compensates for over 60° of phase distortion introduced at the television transmitter and other parts of the TV receiver. In addition, it is vestigially weighted near

the picture carrier and has a "sound shelf" designed on the low frequency side to prevent cross-modulation with the sound carrier. Built on ST-Quartz the 2.7 x 10.4 mm size filter fits in a standard TO-8 transistor header.

The development of this filter and the related TV receiver was under-taken for the FCC to study the feasibility of adjacent channel allocation of the television bands. In meeting all major design goals, this filter exemplifies the significant performance improvements available with SAWs, particularly when the system and component design is done hand in hand.

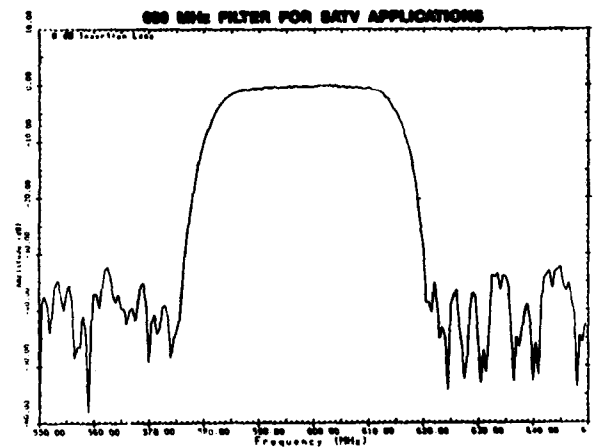


Fig 8: An SATV IF filter centered at 596MHz has an insertion loss of 8dB. The filter has a 30MHz 1dB bandwidth.

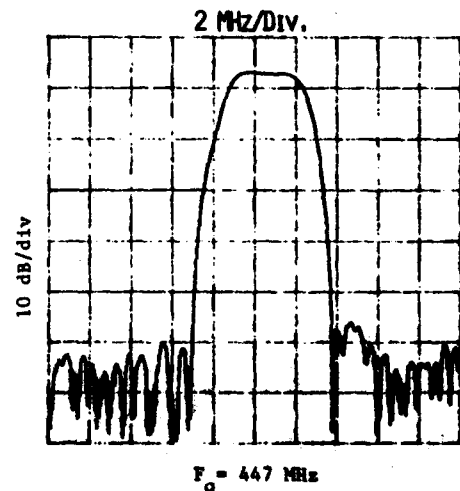


Fig 9: A filter designed for the FCC for an alternative TVIF, has an insertion loss of 6.5dB at a center frequency of 447MHz. The exhibits the ability of SAW filters to implement complex amplitude and phase characteristics, the filter has a sound shelf and phase compensation over a 4MHz bandwidth.

III OUTLOOK FOR THE FUTURE

A joining of SAW resonator and filter technologies are becoming possible with the advent of coupled resonator filters. As seen in figure 10 the peaks of two resonances are apparent, showing how resonators with bandwidths typically too small for filter applications can be combined to yield a larger bandwidth but one still too small to implement with a low cost conventional SAW filter.[5]

Another advance is underway in making a Single Phase Unidirectional Transducer (SPUDT) which essentially incorporates a reflecting array inside a transducer structure to reflect the backward traveling surface wave forward, without the need for external phasing networks.[6]

Diagramed in figure 11, the SPUDT has not yet been used in a production device, but it, along with the coupled resonator filter appear to be the beginning of a third generation of SAW filters that hold great promise for both low cost and low insertion loss.

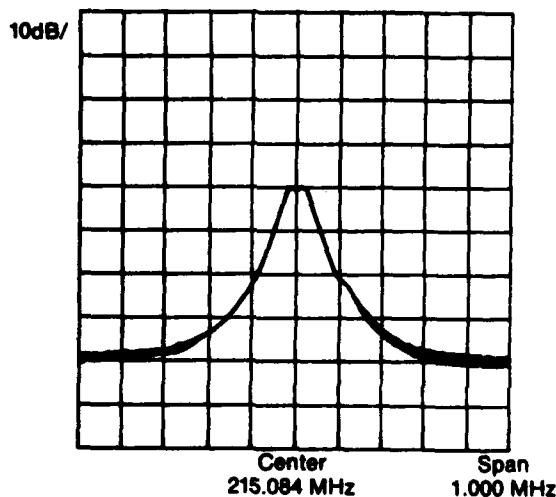


Fig 10: An early coupled resonator filter under development at RFM shows two distinct resonances that, as combined, yield a broader bandwidth than a single resonator can implement and a narrower bandwidth than can be economically made from a conventional SAW filter.

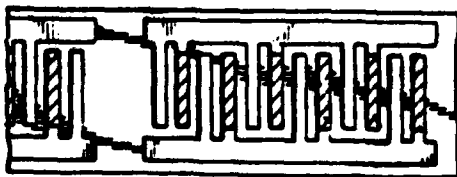


Fig 11: Schematic of a Single Phase Unidirectional filter which incorporates a thicker layer of metal on every other electrode to provide reflecting array inside the transducer. The reflections can be made to cancel in one direction, making the transducer unidirectional.

IV SUMMARY

The major advantages of SAW filters are:

- o Mass producible filter from HF to UHF
- o Reproducible high performance
- o Independent amplitude and phase control
- o High intrinsic Q
- o Small packaging
- o Immunity from vibration
- o Low cost in volume

Some of the factors limiting SAW filter usage are:

- o Design and tooling costs are usually costly and time consuming
- o High precision only available currently for medium to high bandwidth filters
- o Only bandpass functions can be implemented.

TABLE II

TYPICAL SPECIFICATION RANGE

Center Frequency :	35 to 1000 MHz
Bandwidth (1dB) :	Up to 20% of center frequency
Insertion Loss :	4 to 10 dB (Second Generation) 18 to 30 dB (First Generation)
Skirt Steepness :	300 KHz to 3 MHz
Amplitude Ripple :	0.1 dB to 2 dB peak to peak
Rejection :	Up to 50 dB
Package :	TO-8, TO-4, 14 and 16 pin DIP
Triple Transit :	45 DB minimum rejection

More specifically, table II enumerates the current types of specifications applicable to SAW filters. It goes without saying that the extremes of these ranges are not consistent and cannot all be incorporated in one filter, but the values do give a good indication of current filter capabilities.

The fact that these devices provide highly reliable filters at low cost, and are now available from several manufacturers make them a key technology applicable to a wide variety of filter problems.

REFERENCES

- [1] A. Oppenheim and R. Schaffer, "Digital Signal Processing," Prentice Hall, 1975, chapter 5.
- [2] J.H. McClellan, T.W. Parks, L.R. Rabiner, "A Computer Program for Designing Optimum FIR Linear Phase Digital Filters," IEEE Transactions on Audio and Electroacoustics, Vol. Au-21, No. 6, December 1973, pp. 506-526.
- [3] L.R. Rabiner, J.H. McClellan, and T.W. Parks, "FIR Digital Filter Design Techniques Using Weighted Chebyshev Approximation," Proceedings of the IEEE, Vol. 63, No. 4, 1975, pp. 595-610.
- [4] Supriyo Datta, B.J. Hunsinger, and D.C. Malocha, "A Generalized Model for Periodic Transducers with Arbitrary Voltages," Transactions on Sonics and Ultrasonics, Vol. SU-26, No. 3, May 1979, equation 10.
- [5] L.A. Coldren and R.L. Rosenberg, "Acoustically Coupled SAW Resonator Filters with Enhanced out of band Rejection," Transactions on Sonics and Ultrasonics, Vol. SU-26, No. 6, Nov. 1979, page 394.
- [6] C.S. Hartmann, F.V. Wright, R.J. Kanay, and E.M. Garber, "An Analysis of SAW Interdigital Transducers with Internal Reflections and the Application to the Design of Single-Phase Unidirectional Transducers," Proceedings of the 1982 Ultrasonics Symposium, page 40.

Jorg E. Raue

As is the case for microwave frequencies and below, filters are required in virtually every practical EHF system. They provide for separation or combining of different frequencies as with frequency up- and down-converters and multipliers. Even a relatively simple receive-only application typically requires a bandpass filter in front of the mixer for image rejection and channelizing. Combinations of bandpass/bandreject filters are typically employed in EHF transceiver applications, both for channel separation and receiver protection.

This paper will present a synopsis of current capabilities in EHF waveguide filters. Emphasis will be on standard type filters rather than non-fundamental mode and dual-mode filters. The latter, although offering the potential for providing the ultimate in high-Q performance, have a rather narrow application range at lower frequencies and no current practical realization above Ku-band. Included will be a comparison between theoretical and experimental Q's achievable for typical waveguide filters in the frequency range from 15 to 100 GHz, as well as a discussion of the differences between these. Specifically, bandpass, highpass and bandreject filters will be discussed. In addition, comparisons will be made between standard rectangular waveguide filters and other approaches such as microstrip, suspended stripline and dielectric image guide. In particular, experimental results obtained at 94 and 140 GHz will be presented and compared.

DIGITAL FILTERING: AN OVERVIEW

Lyman Shantz
Raytheon Company
Missile Systems Division
Bedford, Massachusetts 01730

Introduction

To introduce the topic of digital filtering, we first consider how conventional passive filters are constructed. The configuration is a succession of resonant sections connected in cascade, with the resonance contributing either an emphasis (pole) or an attenuation (transmission zero) at a particular frequency. Each section is a second order linear system, characterized by a resonant frequency and a Q , or quality factor (Figure 1).

Such a filter would be specified by a frequency response, and one must determine proper L and C values to implement the design. To convert such a design to digital form, however, one must specify the poles and zeros for the individual second order sections, rather than the L and C values.

Since the input must be in digital form, we digress momentarily to consider two important requirements to be observed in the A/D conversion of the input signal to the filter. The first is the band-limiting of both the spectrum and the characteristic function of the signal; the second is to choose both a sampling rate and a sampling step size that will preserve respectively the spectral and the statistical properties of the signal. The first of these requirements is the familiar sampling theorem of Nyquist, the second is the less familiar quantization theorem of Widrow. An analog filter is commonly used in the receiver preceding the A/D to band-limit the signal, and a suitable conversion rate is then chosen.

Likewise, since signals are seldom totally noise-free, the characteristic function is band limited by "front end" noise as illustrated in Figure 2. A quantizer step-size is then chosen at twice

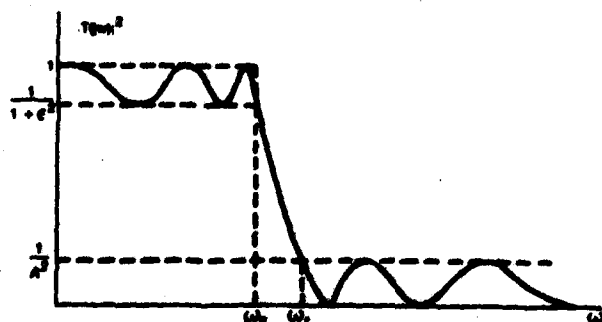
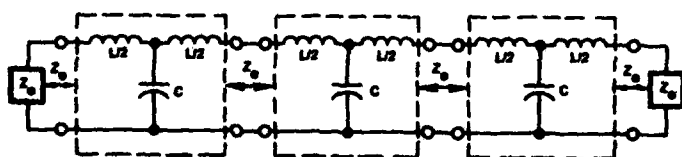


Figure 1. Typical Elliptic Characteristic

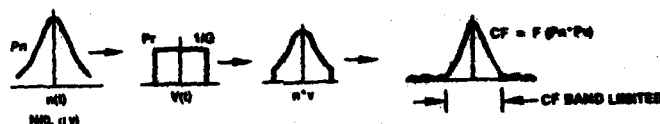
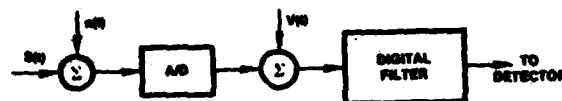


Figure 2. Quantization Theorem

the "fineness" of the background noise. This will allow the statistics of the digital system output to be determined.

Classical Filters

A digital filter, such as might be used as a tone filter in a modem, could be one of the "classical" types (Figure 3) ... Butterworth, Chebyshev, or Elliptic. The first two of these have only poles, which are arranged respectively on a circle or ellipse; the latter has both poles and zeros, and as a consequence has a superior selectivity. The poles of the elliptic filter are located as are those of the Chebyshev, but the zeros are not as simply defined. A slight digression is helpful...

A great deal of insight is gained in designing a filter if advantage is taken of the duality of the phase and amplitude characteristics of the network and the equipotential and streamlines of the electric field. Thus, to design a charge configuration giving a specified potential function is to design a filter with a specified attenuation function (Figure 4). In the case of the elliptic filter, it is convenient to first locate the plus and minus charges of the potential, or the poles and zeros of the filter in a unit rectangle (See Figure 5), where the coordinates are simple (e.g., $1/6$, $2/6$, $3/6$, etc. for a 6-pole design). The rectangle is then mapped back to the conventional s -plane with the elliptic sine function (which maps the rectangle into the half-plane) and an s -plane design results. The s -plane function $H(s)/D(s)$ becomes a z -plane function $H(z)/D(z)$ by a bilinear transformation $s=(z-1)/(z+1)$, and as a last step the filter is realized as a concatenation of cascade second order two-pole two-zero networks as illustrated in Figure 6, with single poles or zeros added as needed (odd function).

Thus a digital filter may be designed as all pole, all zero, or both. The response is periodic with symmetry about the half sampling rate, but care is always taken to insure that no frequencies are present at the unwanted response bands.

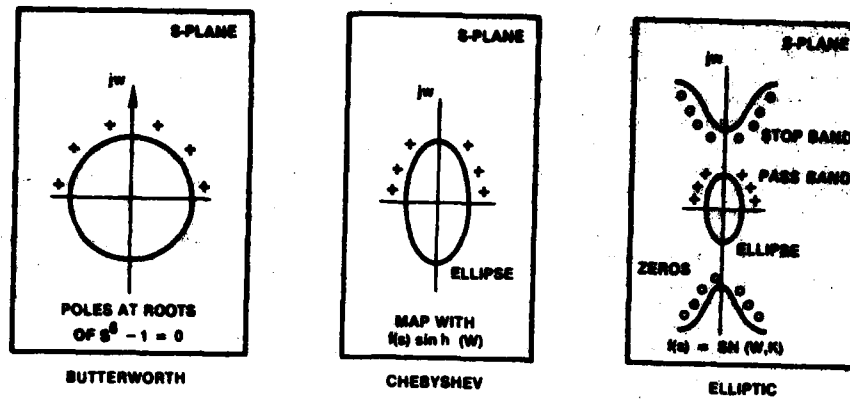


Figure 3. Classical Filters

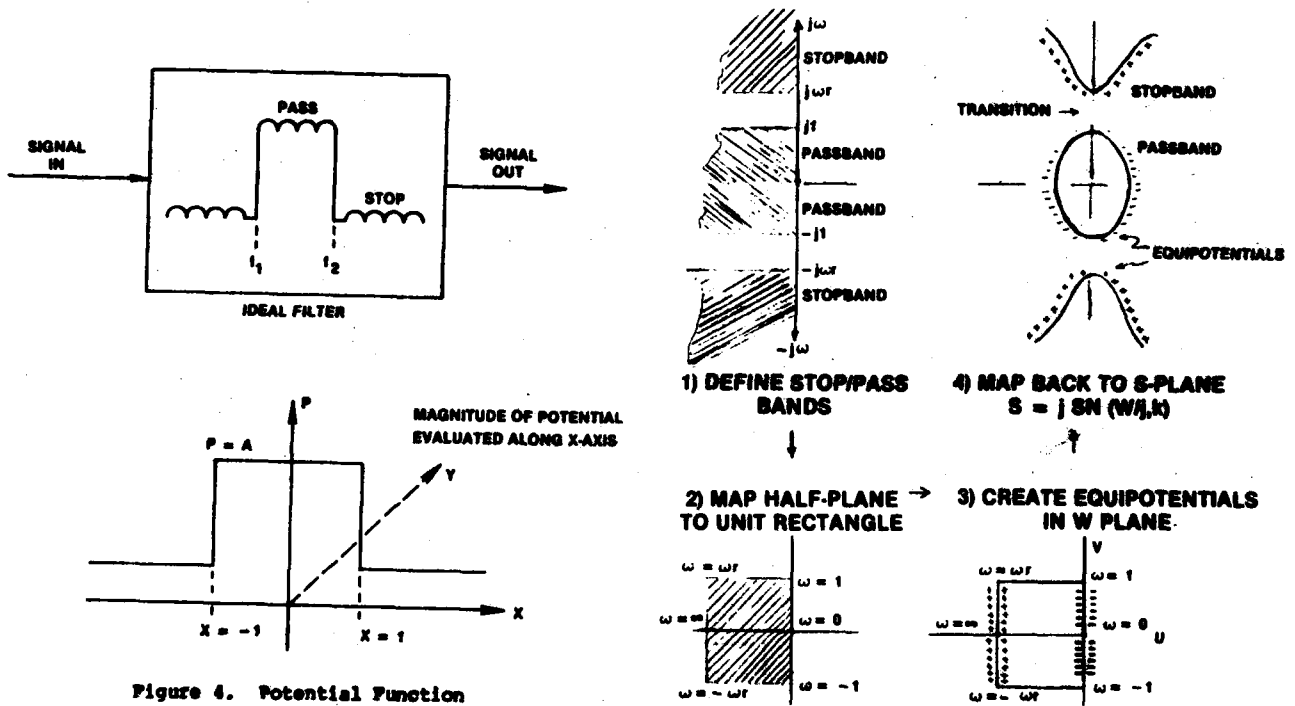


Figure 4. Potential Function

Figure 5. Potential Analogy

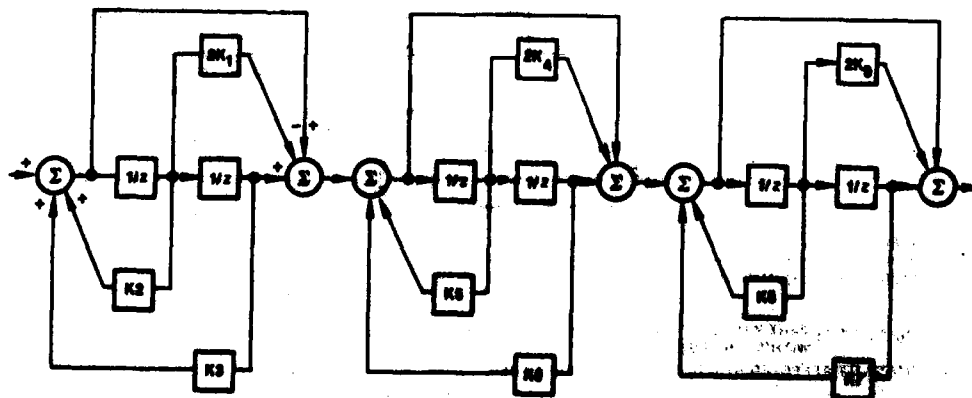


Figure 6. Sin-Pole Sin-Zero Filter

The most remarkable feature of a digital filter of this design is its ability to completely change in characteristics simply by inputting a new set of coefficients. An example of how this feature dramatically improves a system can be seen in the case of a digital autopilot or control system for an air/air missile, illustrated in Figure 7. Such a system, when coupled with the dynamic response of the airstream during flight, forms a fourth order system that is characterized by a set of coefficients that vary nonlinearly with Mach number (velocity) and attack angle (incidence angle to flight path). This over-simplified model will not exhibit stability over normal trajectories without compensation, which is usually added in the form of shaping filters for phase advance of the gyro rate or damping of body resonance. These filters must be programmed to change continuously with M and α , and that is accomplished by moving the time constants (or poles and zeros) by inserting new digital coefficients and gains.

Radar Filters

The three principal types of filters used in radar are the MTI, Doppler, or Pulse Compression filter, all of which are a form of a matched filter.

The first of these to be used was the MTI, or Moving Target Indication, as it was called. To understand this concept, we first consider the nature of the radar signal.

A commonly used waveform is a burst or a multi-pulse waveform, consisting of several transmitted pulses, each followed by a listening interval. Strictly represented, these intervals would be strung end-to-end, but we have placed them side-by-side, since this will place in evidence the sinusoidal variation in a single range cell over the set of pulses of a coherent signal, such as a target, or perhaps some moving chaff (See Figure 8).

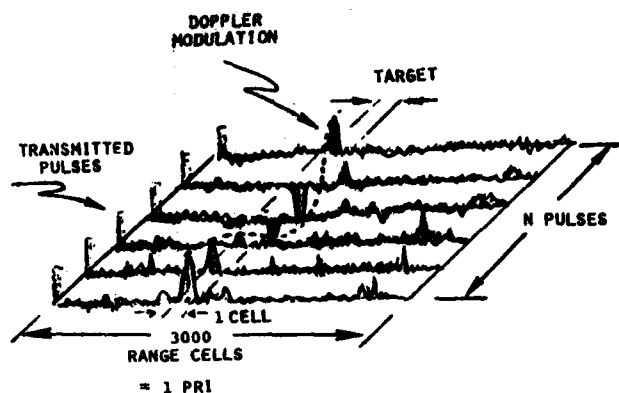


Figure 8. Radar Signal

The configuration known as an MTI (Figure 9) was originated from pure intuitive considerations, based on the obvious observation that two successive echoes that lacked any pulse to pulse modulation due to motion could be added in opposition and made to cancel, provided they were suitably aligned. This alignment was at first provided by quartz delay lines of one PRI (Pulse Repetition Interval) delay.

It was but a matter of time that this concept was recognized as a filter, and a digital implementation using memory storage to provide the PRI delay assumed the form already shown, i.e., the catenation of two-zero networks. Poles are seldom added, due to an undesirable transient response of a recursive structure (Figure 10).

The next form of radar filter to appear was the Doppler filter bank, consisting of a sequence of contiguous time gates (diode bridges) with each gate connected to a bank of narrow band crystal lattice

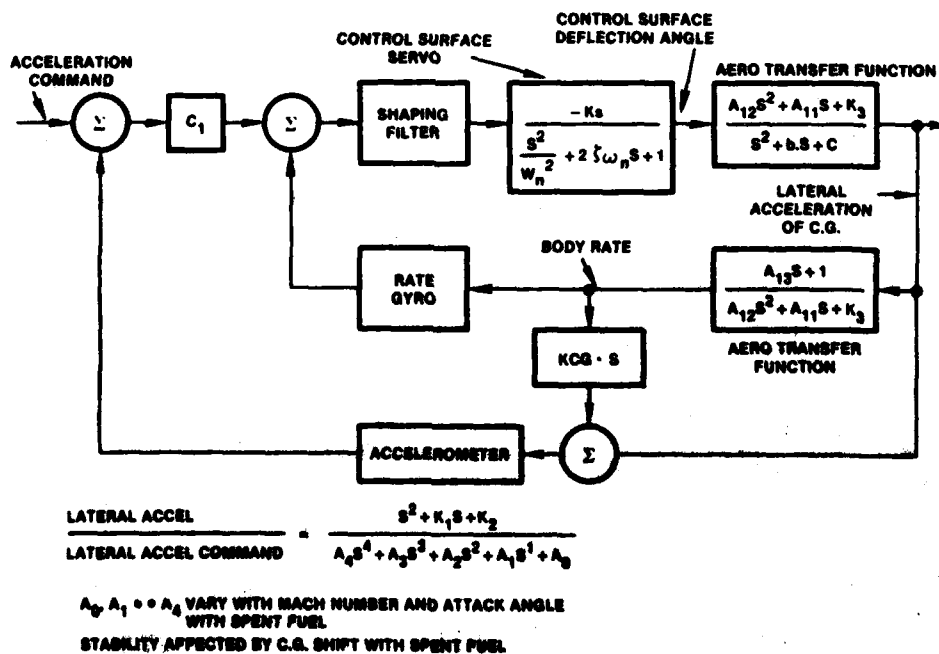


Figure 7. Digital Missile Control

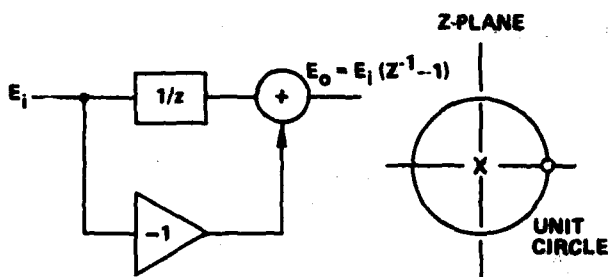


Figure 9. MTI Cancellation

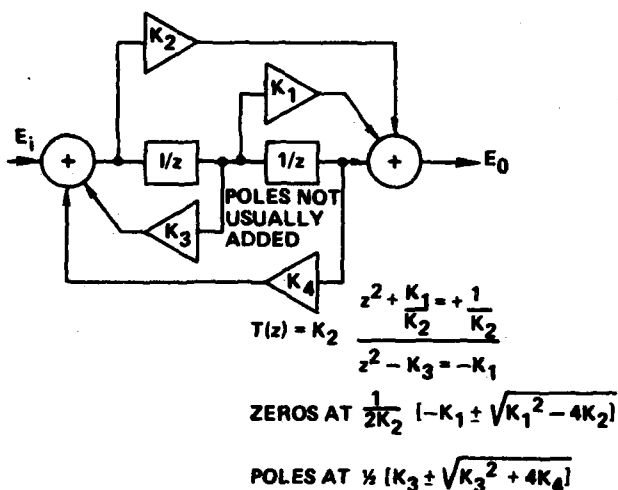


Figure 10. Radar Filter

filters, as illustrated in Figure 11. Clearly, the signal will occupy a single filter and the noise will occupy all filters, with a resulting S/N improvement determined by this ratio.

The FFT however was quickly recognized as a digital counterpart of the range gated bank of Doppler filters. The manner in which the FFT processes the signal can be seen from the illustration (Figure 12) in which the large clutter signal slowly shifts to its position in the spectrum, with the higher frequency target signal finally emerging. Once one can readily see the effect of non-observance of the quantization theorem ... the statistics of the output signal are clearly not those of the input.

The pulse compression filter may be considered as a filtering operation applied in the range direction of a single PRI, rather than operation on modulations occurring from one PRI to the next, as with the MTI or Doppler filter.

The form of the pulse compression filter results from the equivalence illustrated in Figure 13 ... i.e., a conventional band pass filter implemented as a correlator. If the time of arrival of the echo from the target is exactly known, a simple correlation is performed, with the FFT used to provide the integration of the output of the cross correlator. If the time of the echo arrival of the pulse echo is not known, two remedies may be applied. The first consists of applying the FFT to both inputs to the correlator, resulting in signals identically aligned in time at the FFT outputs. (The time shift becomes a phase shift of the transformed signal.)

If the time misalignment ... i.e., uncertainty in time of arrival of the pulse echo ... is slight, a second method is used. The two signals (transmitted pulse and echo) are simply correlated, with

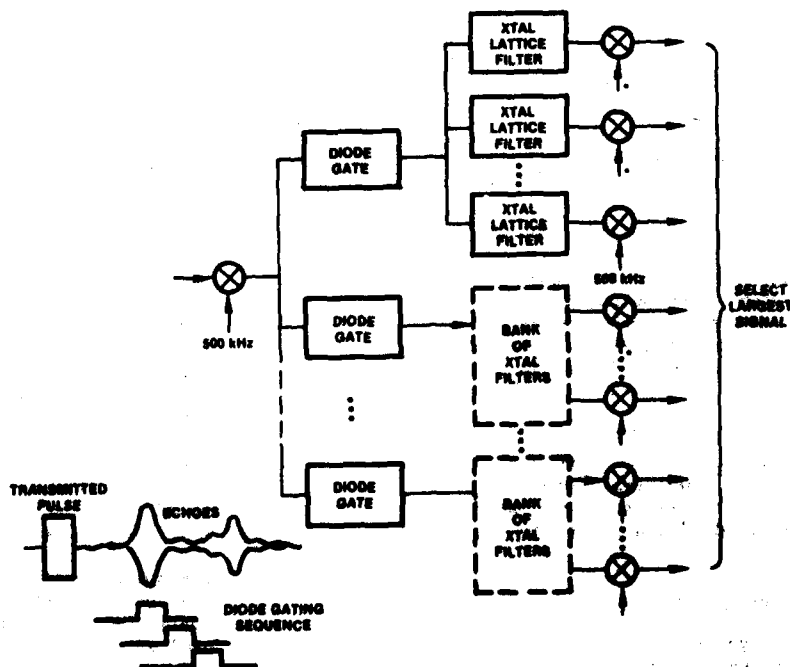


Figure 11. Doppler Filter Bank

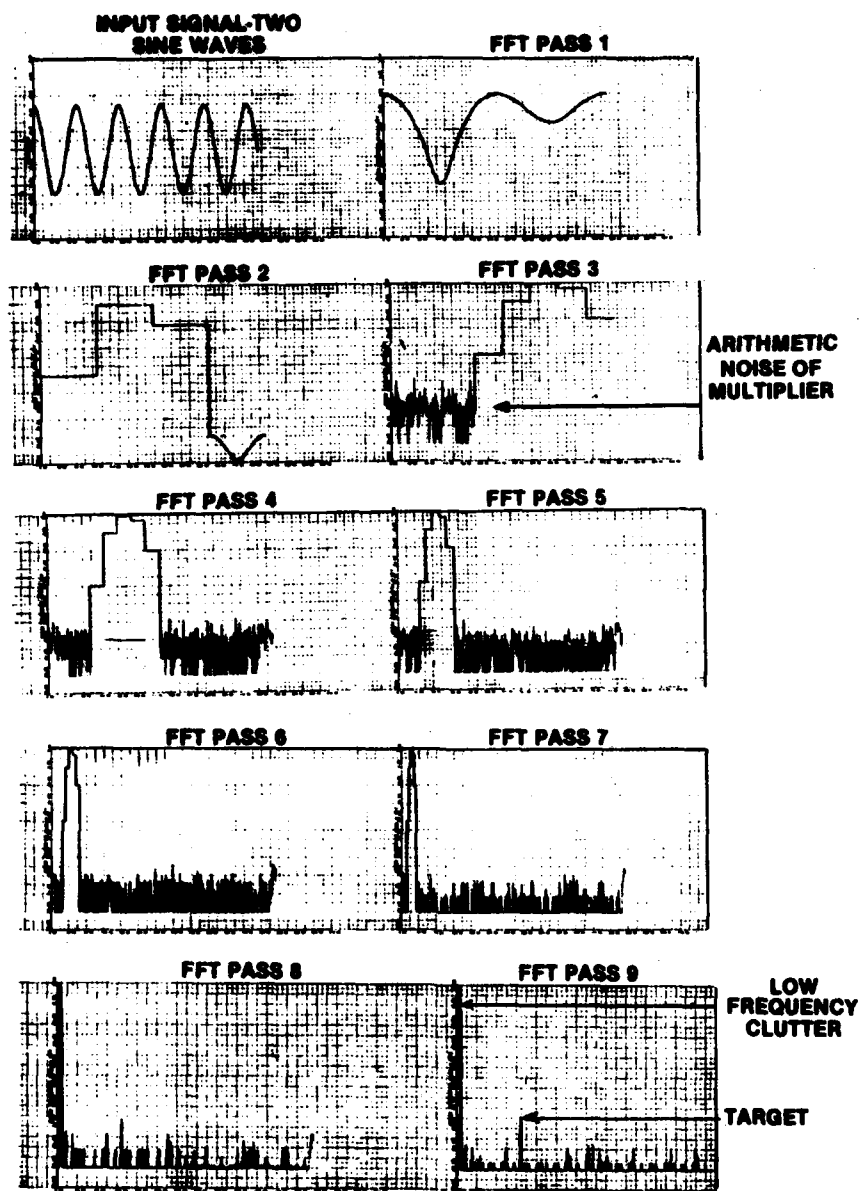


Figure 12. Doppler Filter (FFT)

the transmitted signal (Local Oscillator) extending over the entire receive window. The output from the correlation is first band limited and then coherently integrated with the FFT. This process is sometimes referred to as "stretch processing", and is illustrated in Figure 14.

Pulse compression filters are useful in Early Warning Radars, where detection of small objects at ranges of thousands of miles is required. Long pulses with lots of energy are transmitted, but the resolution in range is preserved with a pulse compression filter.

Maximum Entropy Method

A number of filter designs have been originated that are known as Maximum Entropy filters. The development was motivated by a number of factors:

- a) Conventional spectral analysis filters using transform methods have

poor selectivity ... i.e., poor stop-band rejection when closely spaced signals are to be separated. The reason for this is the forced truncation of the signal into batches either to fit into a finite computer memory or because only limited data are available. The question naturally arises ... could the signal not be better filtered if it could be extended in some rational way (See Figure 15).

- b) The speech coding problem arose from an attempt to do a filtering operation; i.e., if 10 constants, used as filter coefficients, can "whiten" (or remove a spectral bump) from a signal, cannot the same 10 constants reconstruct the signal? If so, one can send a few constants over a channel more

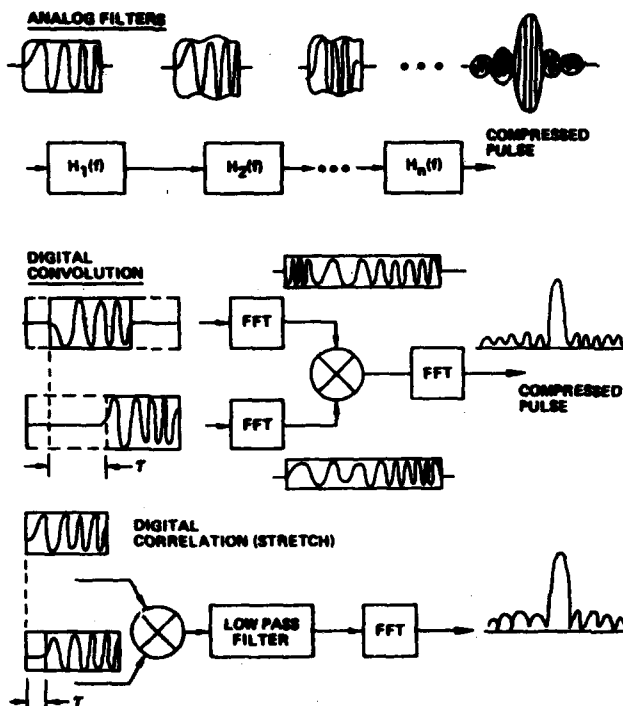


Figure 13. Pulse Compression

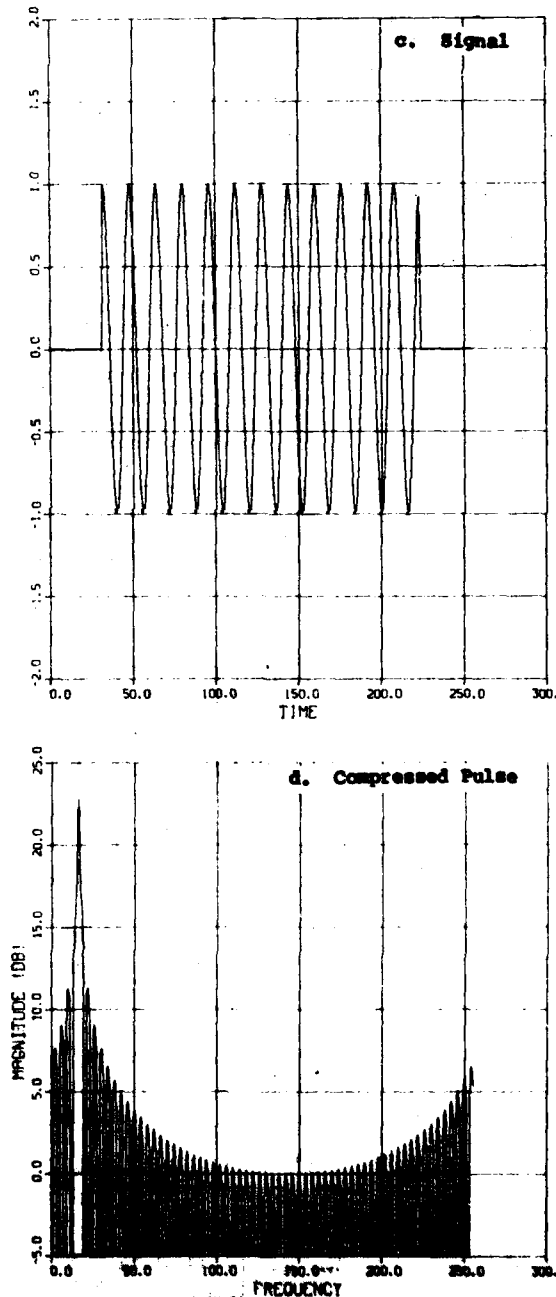
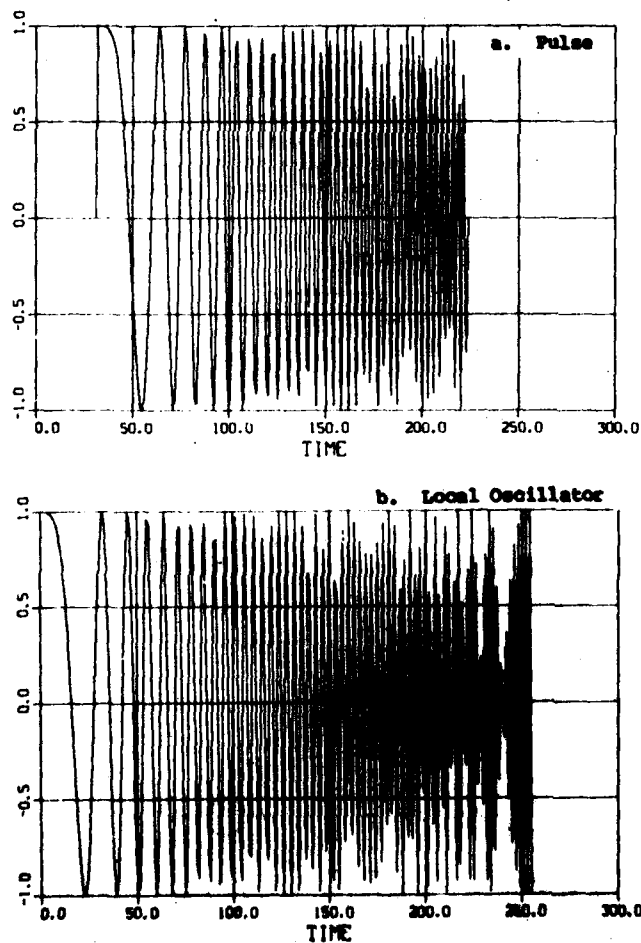


Figure 14. Stretch Pulse Compression

efficiently than a much longer sample of the actual speech.

Both the high resolution spectral analysis and the speech synthesis developments led to a filter that appears also useful as an adaptive whitening filter for removing clutter reverberations in radar and sonar signals, and much investigation along these lines is currently underway.

The MEM (Maximum Entropy Method) filter is best understood by comparing it to transmission of waves through a layered medium, as is done naturally in the atmosphere (thus coloring the sky blue), or is done in microwave filters. The filtering can be

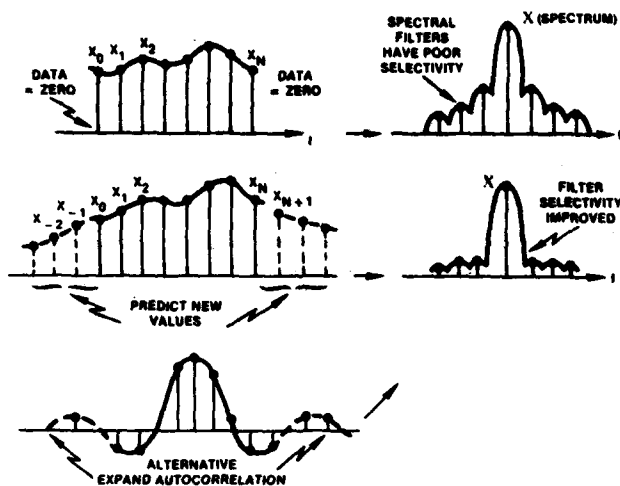
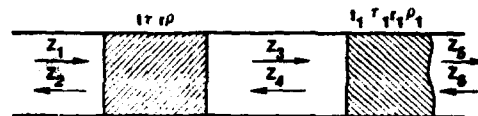


Figure 15. Spectrum Analysis

accomplished by a series of "obstacles" that both reflect and transmit incident radiation, as illustrated in Figure 16. This analogy is introduced for two reasons:

- Two "obstacles", properly spaced in wavelengths, can form a resonant structure by virtue of multiple reflections giving rise to repeated transmissions, thus allowing a "ringing" or extension of a transient incident signal.
- The reflection coefficients for the "obstacles" obey a Riccati equation, which arises naturally from physical considerations.



$$(1 + r_1 \rho + r_1^2 \rho^2 + \dots) \begin{pmatrix} Z_5 \\ Z_4 \end{pmatrix} = \begin{pmatrix} l_1 \rho_1 \\ r_1 \tau_1 \end{pmatrix} \begin{pmatrix} Z_3 \\ Z_6 \end{pmatrix}$$

$$= \frac{1}{1 - r_1 \rho}$$

$$\begin{pmatrix} Z_3 \\ Z_2 \end{pmatrix} = \begin{pmatrix} l_1 \rho \\ r_1 \tau \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_6 \end{pmatrix}$$

$$\begin{pmatrix} Z_5 \\ Z_2 \end{pmatrix} = \begin{pmatrix} l_1 (1 - \rho \tau_1)^{-1} l_1 \rho_1 + l_1 \rho_1 (1 - r_1 \rho)^{-1} r_1 \\ r_1 \tau_1 \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_6 \end{pmatrix}$$

IN GENERAL

$$S(x, y) = \begin{pmatrix} u(x, y) \rho(x, y) \\ v(x, y) \tau(x, y) \end{pmatrix}$$

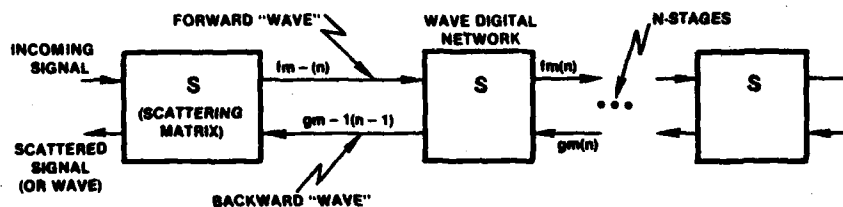
$$S(x, z) = S(x, y) \cdot S(y, z)$$

Figure 16. Scattering

This structure may be transplanted directly into the world of filters, with the "obstacles" replaced by sections of a lattice (Figure 17). The lattice sections are not the usual second order structure as in a crystal filter, for example, but are actually first-order sections.

However, what one should observe here is that the separation of a second-order structure into two first order sections will give rise to a Riccati equation, which in essence takes the place of the second order system (Figure 18).

Thus, we see that the structure described has two properties ... it will extend a transient input by a pseudo-resonance (Figure 19) in such a way that the emerging signal has the proper correlation ...



$$\begin{bmatrix} f_m(n) \\ g_m(n) \end{bmatrix} = \begin{bmatrix} S \end{bmatrix} \cdot \begin{bmatrix} f_{m-1}(n) \\ g_{m-1}(n-1) \end{bmatrix}$$

$$f_m(n) = f_{m-1}(n) + K_m g_{m-1}(n-1)$$

$$g_m(n) = g_{m-1}(n-1) - K_m^* f_{m-1}(n)$$

$$K_m = \frac{f_{m-1}(n) \cdot g_{m-1}(n-1)}{(f_{m-1}(n))^2 + (g_{m-1}(n-1))^2}$$

S = SCATTERING MATRIX

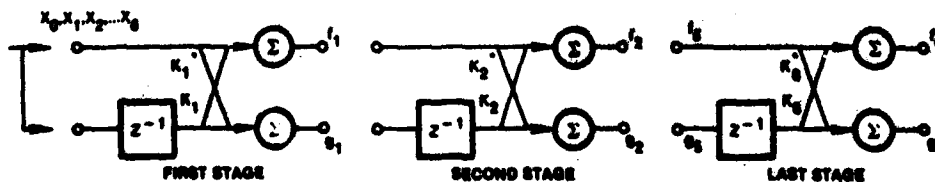


Figure 17. Lattice Filter

CONSIDER LINEAR OPERATOR L

$$L = P_0(x) + P_1(x) \frac{d}{dx} + \frac{d^2}{dx^2}$$

WE WISH TO FACTOR INTO FORM

$$Ly = \left[\frac{d}{dx} + a_1(x) \right] \left[\frac{d}{dx} + a_2(x) \right] y$$

MULTIPLY:

$$Ly = y'' + (a_1 + a_2)y' + (a_2' + a_1a_2)y$$

$$P_1 = a_1 + a_2 \quad a_2' + a_1a_2 = P_0$$

ELIMINATE a_1

$$a_2' = a_2^2 - P_1a_2 + P_0 \quad (\text{RICCATI EQUATION})$$

NATURAL RESULT OF CLOSURE IF WE DEMAND

$$Ly = [L_1 \cdot L_2]y$$

Figure 18. Riccati Equation

i.e., if it is initially first-order Markov, it will remain so. Secondly, the filter structure corresponds to an iterated lattice, with an associated Riccati equation tagging along. Although this first order nonlinear equation has no closed form analytical solution, in general its solution by numerical digital means is rather easy (easier, in fact than its second order counterpart).

Actually, the MEM filter arose not from scattering theory, but from an effort to specify a filter that would maximize the entropy (randomness) of a given signal. Such a filter would have zeros properly located to attenuate the periodic resonances in the correlated signal. The solution to this problem is the Yule-Walker equations, which simply state that the autocorrelation of the filter is the same as the correlation function of the signal that it whitens. The solution leads directly to the lattice equations. The reflection coefficients are the coded speech, or the coefficients of a tapped-delay line filter, as illustrated in Figure 20.

The ability of the MEM filter to whiten a signal ... i.e., remove a "bump" from the signal spectrum, is useful in adaptive digital beam-forming in antenna technology. If we consider a plane monochromatic wave to be sampled by discrete elements, each equipped with an A/D converter, it can be readily seen that to properly phase shift each element signal so that all signals add in phase is equivalent to steering a beam in the direction of the arrival of the wave (see Figure 21). Likewise, phase shifting each wave sample to add anti-phase will steer an antenna pattern null in the direction of arrival of the wave. This latter technique has been applied to both radar and sonar (as well as communication) systems in the form of a sidelobe canceler, which is a technique for nulling out interfering signals in the sidelobes of an antenna while allowing the main lobe to receive with full gain.

The nulling process, accomplished by removing the modulation across the sampled wave, is exactly the same "whitening" operation that the adaptive filter accomplishes, and can be carried out in exactly

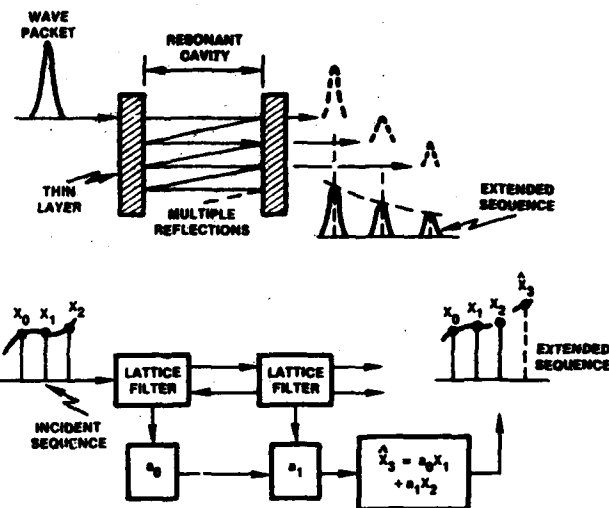
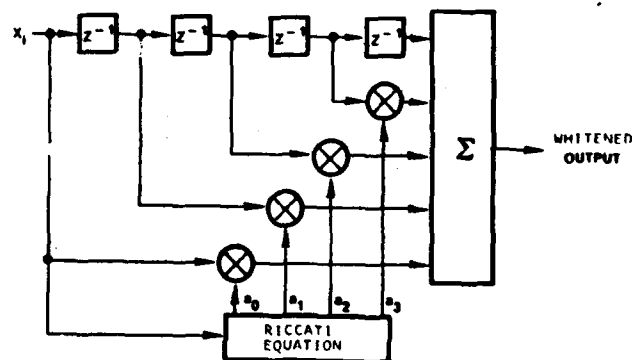


Figure 19. Scattering



$$a_m, k = a_{m-1, k} - a_{m, m} \cdot a_{m-1, m-1}$$

Figure 20. Tapped Delay Line

the same way. Thus, all of the developments of adaptive filters based on the MEM principle can be directly applied to adaptive beamforming.

Advanced Topics

Clearly substantial progress has been made since the beginning of digital filtering, which roughly coincides with the introduction of semiconductor memory. Further developments are being pursued. Dynamic Time Warping, as a technique to aid speech recognition, has usefulness in filtering that is yet to be developed. The technique consists of a transformation procedure applied to the filter coefficients (or encoded speech) that performs an arbitrary mapping of the coefficients to realize a time shift or local compression or expansion that will effect a better fit between the data (encoded) and a model, as illustrated in Figure 22. The mapping is dynamically and recursively altered to achieve the minimum error between the reference and the data. The central function in this procedure is a look-up table which can be regarded as both a time warping and weighting that provides emphasis or diminution of stationary or non-stationary segments.

Further developments in the realm of filtering, which may be regarded as the same problem as spectral analysis, are directed at achieving more optimality in the processing. An example is the method of Pisarenko for finding (or eliminating)

AD-A136 673

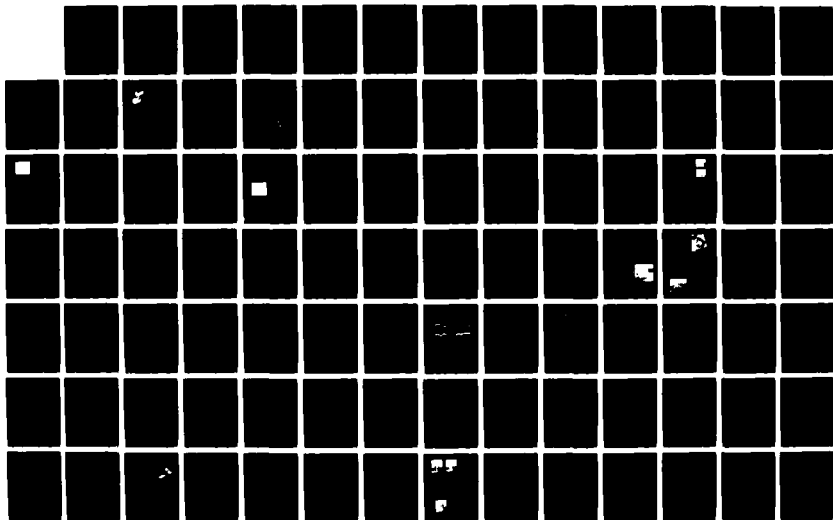
PROCEEDINGS OF THE ANNUAL SYMPOSIUM ON FREQUENCY
CONTROL (37TH) 1-3 JUNE 1983 (U) ARMY ELECTRONICS
RESEARCH AND DEVELOPMENT COMMAND FORT MONMOUTH, NJ 08050-1983

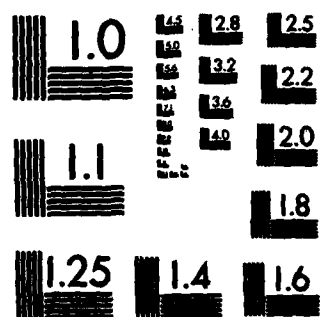
UNCLASSIFIED

F/G 9/1

NL

516





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

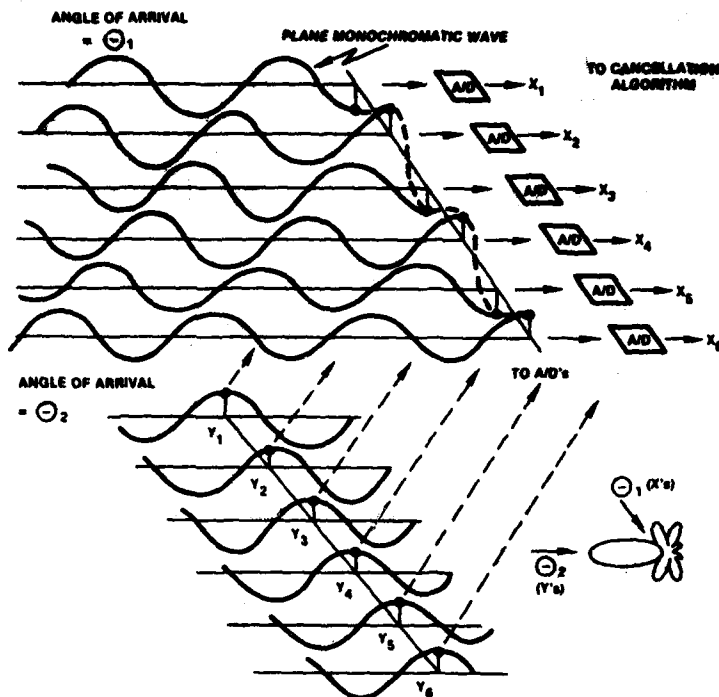


Figure 21. Jammer Cancellation

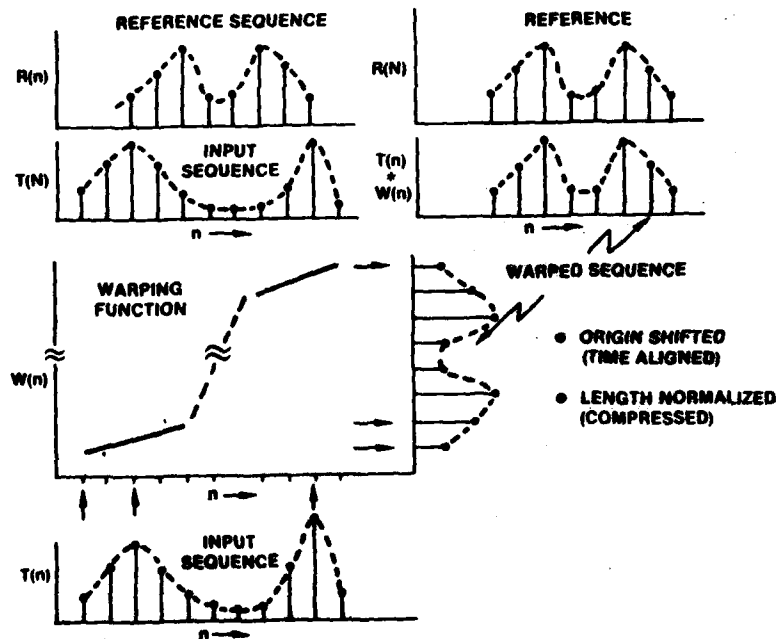


Figure 22. Dynamic Time Warping

sinusoidal signals in a noise background. For signal records of substantial length that correspond to sharply band limited signals, the entire process can be modelled as comprising harmonically related sines and cosines representable by a Fourier expansion. For short data records, however, the spectra are better represented by sines and cosines that are not harmonically related. These eigenfunctions, as they are called, can be determined analytically for a given spectrum model, such as the Butterworth spec-

trum. The procedure presumes a knowledge of the correlation function, from which the minimum eigenvalue can be found. This in turn establishes an algebraic equation, which may then be solved to yield the sinusoid frequencies comprising the correlated spectrum and the power in each spectral line. This information then provides the basis for eliminating the frequencies thus determined by using the tapped delay line structure already described.

MICROELECTRONIC ANALOG ACTIVE FILTERS

Rolf Schaumann

University of Minnesota
Department of Electrical Engineering
123 Church Street S.E.
Minneapolis, Minnesota 55455

Abstract

The paper outlines the state-of-the-art design methods for analog active RC filters suitable for microelectronic hybrid or monolithic implementation. Only practical designs are presented which will lead to filters that have been proven to result in reliable performance. Space restrictions do not permit the inclusion of detailed design procedures and equations, but limits on operation are identified and numerous references are given that enable the reader to arrive at a working filter.

Introduction

Traditionally, signal filtering has been accomplished by making use of resonance effects in LC series and parallel tank circuits. With the increasing emphasis on miniaturization, however, inductors had to be eliminated because no method exists for implementing high quality inductive elements in integrated form. Instead, it was soon found that gain and feedback could be used to achieve resonance effects so that filtering functions can be built in microelectronic form. Consequently, so-called active RC filters, circuits consisting of resistors, capacitors and active devices, usually operational amplifiers (op-amps), became widely used for analog signal processing needs over a wide frequency range. Active RC filters are usually implemented in hybrid thin or thick film technology and more recently also in fully integrated monolithic form. Such microelectronic active filters are useful for operation from sub-audio frequencies until the radio frequency range, and recent work is concerned with extending active filter operation into the range of several hundred megahertz.

The limits on operating frequencies are imposed mostly by the active devices used in filter design. At the low end, element values, R and C, for setting the required time constants tend to become unreasonably large and cause fabrication difficulties. More importantly, the amplifier drifts and offset voltages give rise to signals that are difficult to distinguish from the useful signals being processed by the filter. The high-frequency limit is imposed by parasitics and especially by the frequency dependence of the amplifiers' gain (bandwidth) and by slew-rate.

Based on early work, active RC filters are incorrectly understood as being very sensitive to fabrication tolerances and component variations. Intensive research in the last few years, however, has resulted in design techniques that yield filters whose sensitivities are very low (at their theoretical minimum) so that reliable high-quality, low cost circuits can be produced that realize the required transmission characteristics. It should be noted that in contrast to LC filters there are virtually no restrictions that the input-output transfer functions have to obey in order to be realizable by active filters. Thus, in addition to the usual lowpass, highpass, bandpass, and band-rejection filters, one can realize delay networks, gain and phase equalization or correction circuits, all-pass networks (with or without gain), and pulse shaping circuits, to name a few.

It is worth mentioning also that active filters are very easily trimmed or tuned to account for fabrication tolerances. A further significant advantage is that tuning or adjusting can be effected electronically, even during operation, so that deviations caused by, e.g., extreme environmental, say temperature, aging, or power supply, variations can be eliminated automatically. Electronic tuning can also be used to build adaptive and tracking filters, or to develop systems where the same circuit (hardware) by means of analog or digital control signals can be reconfigured (programmed) to realize different filter transfer characteristics, such as, e.g., a bandpass and a bandstop filter, or a bandpass with different selectivities and passband frequencies. In addition, apart from technological advances it is mainly the electronic tuning property that makes possible the implementation of analog filters in monolithic form where large absolute value component tolerances and drifts must be accounted for, and direct element trimming, say by a laser, is impossible or expensive and should preferably be avoided.

To appreciate the popularity of active RC filters, one only needs to look at the technical literature. Apart from a variety of excellent textbooks¹⁻³ there are literally thousands of papers introducing new or modified filters, or discussing their performance. A fairly complete list of references is available in two volumes^{4,5} that at the same time give an indication of the state-of-the-art in active filter design. Among the hundreds of different circuits and design methods proposed, only very few have proven to be useful in practice. The criteria that usually decide between acceptance or rejection of a new filter structure are sensitivity measures^{1,2} that indicate how much a performance parameter P, say gain or selectivity, changes when a circuit component k varies, e.g.,

$$S_k^P = \frac{d \ln P}{d \ln k} = \frac{dP/P}{dk/k}$$

Clearly, sensitivities should be as small as possible if a filter is to behave well in practice, when components vary due to fabrication tolerances or environmental effects such as temperature changes, aging, radiation, humidity, etc.³

For virtually all applications, the most critical component is the active device, usually an operational amplifier. Unless the op-amp gain with its frequency-dependence is modeled accurately as

$$A(s) = \frac{w_t}{s + \sigma} = \frac{w_t}{s} \quad (1)$$

where w_t is the gain-bandwidth product, when developing a filter circuit the designer has no hope of achieving a satisfactory performance even at operating frequencies in the low audio range.

The circuits and design procedures discussed in the following have been proven in practice to result in reliable, working filters. Within the space of this

paper, these circuits will be identified and some performance characteristics will be highlighted. For detailed design procedures and equations the reader is referred to the literature.

First and Second Order Filters

Voltage transfer functions $V_2/V_1 = T(s)$ of the form

$$T_1(s) = \frac{as + b}{s + c} \quad (2)$$

can be realized in practice by passive RC networks and so need not concern us here. Second-order functions of the form

$$T_2(s) = \frac{a_2 s^2 + a_1 s + a_0}{s^2 + s\omega_0/Q + \omega_0^2} \quad (3)$$

with pole frequency and quality factor ω_0 and Q , respectively, are realized most conveniently^{2,6,7} by a single amplifier biquad (SAB), shown in Fig. 1 for the bandpass case (i.e., $a_2 = a_0 = 0$).

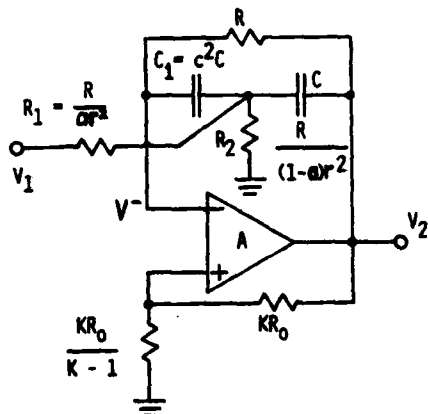


Fig. 1. SAB bandpass

Detailed design guidelines and equations are available in the literature^{2,6,7,8}. The circuit uses only a single amplifier (no unnecessary power consumption!) and, if designed correctly, behaves as good as or better than all other single-amplifier and most multi-amplifier filters. Implemented in hybrid technology^{5,8}, with an op-amp chip and chip capacitors bonded to a resistive thin film network on a ceramic substrate, the filter is realized in large numbers in microelectronic form and has proven very reliable in its performance in the field.

To obtain lower sensitivities, especially to the active device parameters that usually are very inaccurate and vary widely under practical operating conditions, numerous multi-amplifier filters have been proposed in the literature^{1,2,4,5}, notably the general impedance converter (GIC) circuit^{2,9,10}. The simplest and at the same time best one appears to be the recently suggested composite single amplifier biquad (CAB)¹², in which the op-amp A in the SAB of Fig. 1 is simply replaced by the composite amplifier of Fig. 2. A comparison¹² of the experimental performance of a bandpass design based on an SAB, GIC and CAB circuit is shown in Table 1 for an $\pm 10\%$ measured change in gain-bandwidth product ω_t (see eq. (1)) of the op-amps.

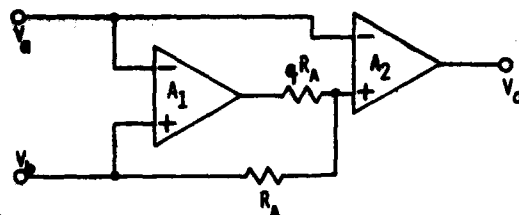


Fig. 2. Composite amplifier

	f_c/kHz	f_o/Hz	Q	f_o/Hz	Q
Nominal Design Values	(ideal)	8414	31.8	18020	20
SAB	1050	8005	47.4	16392	26.5
	560	7717	38.0	15364	22.1
GIC	1050	8278	35.2	17291	23.6
	560	8180	32.7	16812	20.6
CAB	1050	8416	32.1	18012	19.8
	560	8416	38.3	18010	29.7

Table 1

Note especially the uncomfortably large f_o -errors in both the SAB and the GIC filters, whereas the CAB proves to be very insensitive indeed.

If parameter errors in active filters become too large, they can be corrected by use of predistortion; it is noted though that predistortion can account only for errors caused by known nominal element deviations. If elements vary further, e.g. due to environmental effects, predistortion is of no help. As an example, the SAB bandpass in Table 1 was to have a center frequency of 8414 Hz; if op amps with $f_t = 1\text{MHz}$ are available, measurements show that a frequency error $\Delta f = 400\text{ Hz}$ must be expected, i.e. a -4.75% error. Predistortion means that the filter should be designed with $f_o = 8800\text{ Hz}$ so that the implementation with $f_t = 1\text{MHz}$ op amps gives a bandpass with $f_o = 8400\text{ Hz}$ as designed. However, further deviations in f_t (e.g. to 560 kHz in Table 1) will not be accounted for.

High-Order Filters⁵

If the requirements of spectrum shaping are more stringent, such as flatter passband gain, large stop-band attenuation, or steeper cut-off, higher order filters must be used. They are realized by one of three methods^{1,2}: (a) cascading low (first or second)-order sections; (b) embedding low-order sections into a resistive feedback network; and, (c) simulating, if it exists, a passive LC filter.

Cascade Structures^{1,2}

The classical approach to realizing a high-order transfer function $H(s)$ consists of factoring $H(s)$ into

$$H(s) = \prod T_i(s) \quad (4)$$

where $T_i(s)$ are low-order functions as in (2) and (3), realizing each $T_i(s)$ as described in the previous section, and finally connecting all the low-order filters in cascade (Fig. 3). The resulting structures have

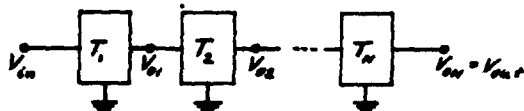


Fig. 3. Cascade topology

the advantage of being modular, easy to design and easy to tune because each transmission zero and natural frequency of the total circuit is realized in one section T_i without any interaction from the other sections. Cascade designs are used in practice in large numbers^{5,13,14} with adequate performance for filters of order up to about 10, and with sensitivities as good as those of the constituent low-order sections.

Multiple Feedback Topologies^{5,15}

Lower sensitivities can be obtained by enforcing some interaction (coupling) between the low-order sections in the cascade. One way to achieve this interaction is to embed the sections into resistive feedback structures, among which one of the more successful¹⁶ ones is the 'follow-the-leader feedback' (FLF) topology¹⁵ shown in Fig. 4. The circuit realizes the transfer function of the form

$$H(s) = \frac{-K \prod_{i=1}^N T_i(s)}{1 + \sum_{i=2}^N F_{1i} \prod_{j=1}^i T_j(s)} \quad (5)$$

where $K = R/R_{1N}$ and the feedback factors $F_{1i} = R/R_{fi}$. The interaction caused by feedback is apparent from (5); clearly, for $F_{1i} = 0$ eq. (5) reduces to the cascade expression (4). Multiple feedback topologies

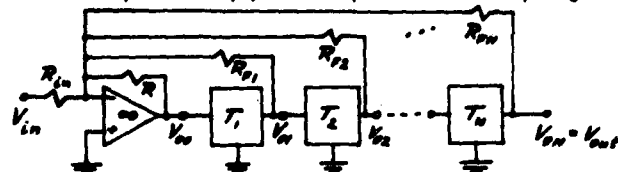


Fig. 4. The FLF topology.

have been shown¹⁵ to exhibit in the passband much improved sensitivities when compared with cascade designs; in the transition bands and stopbands, the sensitivities of all filters are about equal in magnitude; see Fig. 5 for some comparative data.

The price paid for the sensitivity advantage is a more difficult design than for cascade filters, much more complicated tuning, need for optimization, and a worse high-frequency behavior than cascade filters due to unpredictable phase shifts, caused by parasites, in the feedback loops¹⁶.

Ladder Simulations^{2,5}

Observe from Fig. 5 that passive LC ladders have the lowest deviations due to component tolerances. Thus, further improvements in sensitivity can be obtained by actively simulating passive LC ladder filters. This method yields the to-date best performing high-order filters. Again, various approaches do exist: one can implement the flow graph of the LC prototype, using active integrators and summers^{5,17}, or one can simulate inductors^{5,18} by use of gyrators,

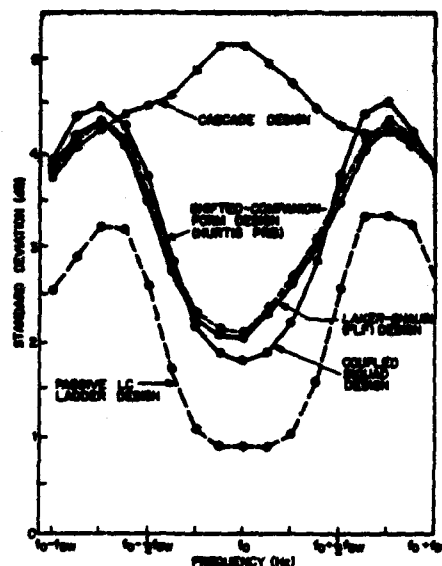
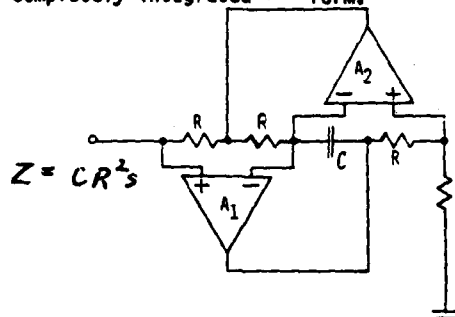


Fig. 5. Sensitivity comparison for a sixth-order Butterworth bandpass filter in different realizations.

e.g. Fig. 5; both methods result in high-quality filters that can be implemented in hybrid^{5,19} or even in completely integrated^{5,20} form.



to two orders of magnitude higher than normally designed active RC filters²². The drawback is that the filters' frequency parameters become dependent on amplifier characteristics that are not well determined and vary with power supply and temperature²². Consequently, the filters need to be combined with some kind of self-tuning feature^{22,24}, such as phase-locked²⁴ or vector-locked^{25,26} loops (Fig. 8) so that parameter values are maintained under varying operating

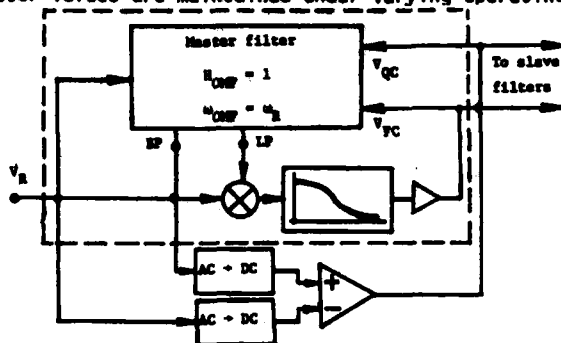


Fig. 8. Two parameter vector locked loop

conditions. Using such design methods, active filters can be obtained that operate at frequencies up to the gain-bandwidth product of the active devices used and maintain their design values automatically to within a small fraction of one percent²⁴.

Integrated Filters⁵

In an attempt to implement analog filtering functions together with all remaining signal processing circuitry on an integrated circuit chip, several approaches have been followed that lead to monolithic analog filters.

The most successful method is that of switched capacitor (SC) filters^{1,5,27}. SC circuits are sampled data filters where resistors are simulated via charge flow on switched capacitors so that all filter parameters are determined by ratios of capacitors and a clock frequency (Fig. 9). Switches (MOS transistors),

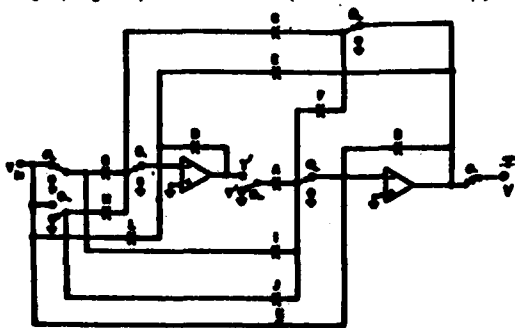


Fig. 9. General SC filter topology.

capacitors, amplifiers and clock generators can be integrated, and both clocks and component ratios can be realized accurately in IC form so that reliable monolithic analog filters are obtained that require no post fabrication tuning. Design details for these very useful and practical filters are available in the literature^{1,5,27}. With current technology, the upper frequency limits are approximately 200 kHz; the limits are given by amplifier bandwidth, settling time and finite ON-resistance of the MOS switches that prohibit instantaneous charge transfer.

In addition to the analog sampled data SC approach, several authors have proposed realizations of "regular"

active RC filters in monolithic form^{5,20,25,28-32}. The approaches follow the entire spectrum of active RC filter design, i.e., second-order sections^{25,31}, cascade filters²⁵, and ladder simulations^{20,28-31}. Most successful results to date have been obtained in the low audio range where the highest volume applications exist, but some attempts are being made to achieve analog filters at video^{20,32} and higher frequencies (>10 MHz)²⁵. The main problem confronting all monolithic active RC filter designers is that of inaccurate element values obtained in fabrication and that of component drift during operation. Since tuning of an IC circuit is impossible or at least expensive, some automatic, electronic tuning is employed by all monolithic filter designs. The methods usually involve embedding the filter into a control loop (phase-locked or vector-locked loop (Fig. 8)) which compares filter performance to an external standard, a reference frequency or clock. Some of these circuits are "master-slave" systems where the filter follows the reference signal, thereby implementing tracking filters. All the monolithic filters reported to date are special purpose designs, although a few of the approaches may emerge as general design techniques^{25,30,31}.

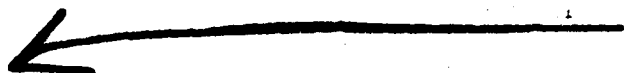
Conclusions

In this paper, a summary of the state-of-the-art and current research trends in active filter design has been presented. The possibilities and major limitations of the filters have been pointed out, as well as the most successful methods of implementing analog active filters in practice. References have been given to allow the interested reader to familiarize himself in greater depth with specific design approaches suitable for his applications.

References

1. M. S. Ghauai and K. R. Laker: *Modern Filter Design: Active RC and Switched Capacitor*, Prentice Hall, 1981.
2. A. S. Sedra and P. O. Brackett, *Filter Theory and Design: Active and Passive*, Matrix Publishers, 1978.
3. C. S. Lindquist, *Active Network Design with Signal Filtering Applications*, Steward & Sons, 1977.
4. S. K. Mitra, *Active Inductorless Filters*, IEEE Press Selected Reprint Series, 1971.
5. R. Schaumann, M. S. Soderstrand, and K. R. Laker, Eds., *Modern Active Filter Design*, IEEE Press Selected Reprint Series, 1981.
6. A. S. Sedra and L. Brown, "A refined classification of single amplifier filters", *Int. J. of Circ. Theory and Appl.*, Vol. 7, pp. 127-137, Jan. 1979.
7. A. S. Sedra, M. A. Zharab, and K. Martin, "Optimum Configuration of Single-Amplifier Biquadratic Filters", *IEEE Trans. Circ. and Syst.*, Vol. CAS-27, pp. 1155-1163, 1980.
8. J. J. Friend, C. A. Harris, and D. Hilberman, "STAR: An active biquadratic filter section", *IEEE Trans. Circ. Systems*, Vol. CAS-22, pp. 115-121, Feb. 1975.
9. B. B. Bhattacharyya, W. B. Mikhael, and A. Antoniou, "Design of RC active networks using generalized immittance converters", *J. Franklin Inst.*, Vol. 297, pp. 45-58, Jan. 1974.

10. C.-F. Chiou and R. Schaumann, "Performance of GIC-derived active RC biquads with variable gain", Proc. IEE, Vol. 128, part 6, Electronic Circuits and Systems, No. 1, pp. 46-52, 1981.
12. R. Schaumann, "Two amplifier active RC biquads with minimized dependence on op-amp parameters", IEEE Trans. Circ. Syst., Vol. CAS-30, Nov. 1983.
13. H. W. Renz, et al., "RC active filters in single-layer tantalum RC-film technology", Proc. IEEE, Vol. 67, pp. 37-42, Jan. 1979.
14. R. A. Friedenson, et al., "RC active filters for the D3 channel bank", BSTJ, Vol. 54, pp. 507-529, March 1975.
15. K. R. Laker, R. Schaumann, and M. S. Ghausi, "Multiple-loop feedback topologies for the design of low-sensitivity active filters", IEEE Trans. Circ. Syst., Vol. CAS-26, pp. 1-21, Jan. 1979.
16. R. Schaumann and M. S. Ghausi, "High-order multiple feedback filters", Proc. ECCTD, pp. 289-298, Sept. 1980.
17. K. Martin and A. S. Sedra, "Design of SFG active filters", IEEE Trans. on Circ. Syst., Vol. CAS-25, pp. 185-195, April 1978.
18. L. T. Bruton, "Multiple-amplifier active-RC filters with emphasis on GIC realizations", IEEE Trans. Circ. Syst., Vol. CAS-25, pp. 830-845, Oct. 1978.
19. W. Saraga, D. Haigh, and R. G. Barker, "Microelectronic active RC channel bandpass filters in the frequency range 60-108kHz for FDM SSB telephone systems", IEEE Trans. Circ. Syst. Vol. CAS-25, pp.1022-1031, Dec. 1978.
20. K. W. Moulding, et al., "Gyrator video filter IC with automatic tuning", IEEE J. Solid-State Circ., Vol. SC-15, 963-968, Dec. 1980.
21. K. Radhakrishna Rao and S. Srinivasan, "A bandpass filter using the operational amplifier pole", IEEE J. Solid-State Circ., Vol. SC-8, 245-246, 1973.
22. J. R. Brand and R. Schaumann, "Active R filters: review of theory and practice", IEE J. Electron Circ. Syst., Vol. 2, pp. 89-101, July 1978.
23. M. A. Soderstrand, "Active R ladders: High-order, high-frequency, low sensitivity active R filters without external capacitors", IEEE Trans. Circ. Syst., Vol. CAS-25, p. 1032-1038, Dec. 1978.
24. J. R. Brand, R. Schaumann, and E. M. Skel, "Temperature-stabilized active R bandpass filters", Proc. 20th Midwest Symp. Circ. Syst., pp. 295-300, Aug. 1977.
25. R. Schaumann and C.-F. Chiou, "Design of integrated analog filters", Proc. ECCTD, pp. 407-411, Aug. 1981.
26. D. Senderowicz, D. A. Hodges, and P. R. Gray, "An NMOS integrated vector-locked loop", Proc. ISCAS, pp. 1164-1167, May 1982.
27. P. R. Gray, D. A. Hodges, and R. W. Broderson, Analog MOS Integrated Circuits, IEEE Press Selected Reprint Series, 1980.
28. K.-S. Tan and P. R. Gray, "Fully integrated analog filter using bipolar JFET technology", IEEE J. Solid-State Circ., Vol. SC-12, pp. 814-821, Dec. 1978.
29. J. O. Voorman, W. H. A. Bruls, and P. J. Barth, "Integration of analog filters in a bipolar process", IEEE J. Solid-State Circ., Vol. SC-15, pp. 963-968, Aug. 1982.
30. M. Banu and Y. Tsiividis, "Fully integrated active RC filters", Proc. ISCAS, pp. 602-605, May 1983.
31. R. L. Geiger and J. Bass, "Performance characteristics of switched resistor filters", Proc. ECCTD, Sept. 1983.
32. K. W. Moulding and P. J. Rankin, "Experience with high-frequency gyrator filters including a new video delay line IC", Proc. ECCTD, Sept. 1983.



T. Counji, T. Kawatsu, Y. Kasai and T. Takeuchi

FUJITSU LIMITED

1015 Kamikodanaka, Nakahara-ku, Kawasaki, 211, Japan

Y. Tomikawa and M. Konno

YAMAGATA UNIVERSITY

3-4-16 Jonan, Yonezawa-shi, Yamagata, 922, Japan

Abstract

This paper describes the design methods and the experimental results of the timing tank mechanical filter. The composite longitudinal mode transducer was made into a sandwich structure by affixing a constant modulus alloy plate to each electrode surface of the piezoelectric ceramic plate. The equivalent electrical circuit constants and the suppression of spurious responses due to the third order longitudinal mode of the transducer were analyzed.

The dimensions of the experimental filter were 20L x 12W x 5H mm, the same size as a 14-pin dip IC. The center frequency of the filter is 200 kHz and it has low insertion loss of 0.5 to 0.7 dB and a bandwidth of 2.1 to 2.4 kHz 3 dB. The I/O phase difference at the center frequency was about 40 degrees. The inclination of the phase was as small as 0.07 to 0.08 deg/Hz and the phase linearity was excellent. The attenuation band has 30 dB or more attenuation over the wide range from 0 to 1.5 MHz to improve the signal to noise ratio of the equipment.

The filter input/output impedance was between 8.5 k Ω and 8.9 k Ω and the impedance matching with CMOS circuits was satisfactory.

The temperature characteristics showed 0.6 dB insertion loss variation and 11 degree phase variation between 5 and 60°C.

Introduction

Along with the development of communications LSI, the digital network is rapidly expanding. To meet this expansion, we have been investigating application of digital transmission technology to subscriber loops because they are relatively unused.

In digital subscriber transmission, non-telephone services such as data and facsimile are carried out simultaneously with telephone services. Consequently, a subscriber house requires equipment having functions such as speed conversion and code conversion.

The timing tank mechanical filter is used in the equipment and is an important element in regenerating the waveform of burst-mode signals that are transmitted and received. For this need, LC filters were first investigated. Because of LC filter problems such as not matching LSI components in shape and size, stability, and frequency adjustment, electro mechanical filters which can handle the above problems were next investigated. As a result a small-size high-performance filter using a composite longitudinal mode transducer and fully satisfying severe requirements was completed.

In the first part of this paper, the construction of the new composite longitudinal mode transducer is presented. The second part contains an analysis and experimental results of the transducer. The third part reports on design methods and experimental results of the mechanical filter.

Requirements

To satisfy the equipment specifications, the performance required for the mechanical filter is as follows:

- (a) Center frequency: 200 kHz
- (b) Relative bandwidth: 1% or more
- (c) Insertion loss: 3 dB or less
- (d) Attenuation: 30 dB or more
(0 to 180 kHz, 220 kHz to 1.5 MHz)
- (e) Input/output impedance: 5 k Ω or more
- (f) Phase stability: ± 20 deg or less
- (g) Volume: 1.2 cm³ or less

To satisfy all the above severe requirements as to characteristics and small-size, we investigated the high-performance transducer and the structure of the filter.

Composite longitudinal mode transducer

Structure

To satisfy requirements (b) and (e), a reciprocal performance of low capacitance ratio and high inductance is needed. Moreover, a transducer structure which has few spurious mode responses is expected to satisfy Requirement (d). To further satisfy (g), the transducer must be small. Therefore, as a result of investigating various transducers, we devised the composite longitudinal mode transducer shown in Fig. 1. This transducer has a structure by affixing a constant modulus alloy plate to each electrode surface of a piezoelectric ceramic plate which has a rectangular cross section, making a sandwich structure, and the first mode resonance is used. It has following features.

- (a) When the mechanical filter is constructed, it is easy to weld the roupler or the supporting wire on the metal plate of the transducer which is mechanically strong and is appropriate for mass production.
- (b) It is possible to compensate the temperature characteristics of the piezoelectric ceramics by the constant modulus alloy.

Results of analysis

When designing mechanical filters, it is useful to analyze the equivalent electrical circuit constants of composite transducers. For analysis, it is essential to determine third mode response suppression conditions, third mode, and fifth mode nodal points. The structure of the transducer causes only its odd-order resonant modes to be excited. Therefore, the half section of the transducer in Fig. 1 (b) must be considered in the analysis. For theoretical development using mathematics, refer to the Appendix.

- (a) **Resonance frequency**
 Figures 2 to 4 show the results of calculating the resonance frequencies, which correspond to the frequency values at the first, third, and fifth resonant modes, respectively. The values, however, are ones that are normalized by the first resonant mode value of a single substance of piezoelectric ceramics.

Step ratios l_2/l_1 and t_m/t_p have a considerably large effect on these values; a greater effect is seen at the high-order resonant modes. The above results are useful for the design of transducers and the estimation of spurious frequencies.

- (b) **Equivalent inductance and capacitance ratio**
 Figure 5 is the result of analysis with respect to equivalent inductance. The resulting values do not differ greatly from those for a single piezoelectric ceramic plate near a practical dimension ratio, l_2/l_1 , of 0.4.

Figure 6 is the result of analysis in respect to capacitance ratio r_1 . Step ratio t_m/t_p has a comparatively large effect on the capacitance ratio, r_1 , of the first resonant mode, and it is assured that r_1 increases in proportion to the step ratio.

Figure 7 is the relationship between step ratios l_2/l_1 and t_m/t_p when a capacitance ratio of $r_3 = \infty$ is assumed for the third resonant mode, that is, when the third resonant mode is suppressed. This relationship is useful as one of the transducer design conditions.

Figure 8 shows the nodal points for the third and fifth resonant modes determined under the third resonant mode suppression condition. The locations of the nodal points also vary considerably with step ratio t_m/t_p . This result is useful, for example, for determining the location of a coupler to be fixed to the fifth resonant mode nodal points.

Figures 9 and 10 show the equivalent mass and force factor A calculated for a transducer.

Experimental results

The results of measurements of the experimental transducer is described below. Table 1 shows the dimensions of the experimental transducer. The transducer is so thin and narrow that the width-extensional mode and thickness mode resonant frequencies can both be 1.5 MHz or more.

Figure 11 shows the results of measurements of an experimental transducer. The third resonant mode response is not completely suppressed but is low enough in comparison with the main resonance. The width-extensional mode and thickness mode resonance frequencies are both sufficiently high as shown in the same figure (Fig. 11).

Figure 12 shows an equivalent circuit of the experimental transducer at the main resonant mode. Table 2 shows the values of the equivalent circuit constants measured for the experimental transducer. Table 3 compares the results of analysis and measurement. Both results correspond well.

Mechanical filter design

To meet the requirement for size reduction, no electrical parts such as coils and capacitors can be used in the mechanical filter. The mechanical filter was therefore composed only of two transducers, one coupler, and supporting wire. Figure 13 shows the structure of the mechanical filter. This mechanical filter has two composite longitudinal mode transducers coupled in parallel by a flexure mode coupler. An equivalent circuit of the mechanical filter is shown in Fig. 11 (a) and (b), where S_0 is the stiffness of the coupler determined in terms of slide-slide-end thin bar stiffness values since the coupler is welded to the longitudinal transducers.⁴ Transformation ratio n_1 is determined as the ratio of the normalized function value at the coupling point to that at the tip of a transducer and can be obtained from equations (6) and (7) in the Appendix. Equivalent mass m_1 (first resonant mode) and force factor A can be calculated from equations (15) and (16), respectively. Thus, the filter design can be accomplished according to (b) of Fig. 11.

Table 4 shows the design procedure. An effective parameter theory of narrow band approximation with a rudder circuit was used in designing the filter.⁵ Butterworth filter was employed as the normalized lowpass filter.⁶ Note, however, that the design of a double-transducer filter with the image parameter theory is almost the same in phase characteristics as that with the effective parameter theory described above.⁷ (a) shows a normalized lowpass filter with the Butterworth characteristics, where $R_0 = 1.0$ and $S_1 = S_2 = 1.414213$. This filter is transformed to a band pass filter as shown in (b). To realize a circuit of (b) in the form of a piezoelectric transducer, gyrator transformation is employed as shown in (c).

Next, to transform it to a mechanical filter, narrow band approximation is done in the vicinity of the center frequency for circuit transformation as shown in (d) and (e). Since C_0 and r are determined by the transducer used, R' and the regulated frequency f_T' of the transducer can be obtained. In this circuit transformation, however, two X values were obtained as shown in Table 4, so a smaller value was employed to realize a high-impedance design with high terminal resistance R' to meet the input impedance requirements. The stiffness of the coupler is determined so that the S_0 value in Fig. 14 can agree with that in Table 4. If the two transducers are shaped differently, Norton transformation is also used in the filter design after the circuit transformation of (e).

Experimental results

In the first experimental filter design, a coupler was fixed at the fifth resonant mode nodal points ($X_0/(l_1 + l_2) = 0.450$) shown in Fig. 8 with the method shown in Fig. 13. With respect to transducers, the one as shown in Table 1 was used for the transducer on the side of filter output, but the width $b = 0.8$ mm was adopted for the transducer on the side of the filter input to prevent overlapping of spurious responses of the transducers in the width-extensional resonant mode.

Figure 15 shows the external appearance of the experimental mechanical filter. Its dimensions are as small as 20L x 12W x 5H mm (1.2 cm³). It is similar in shape to a 14-pin Dip IC.

Figures 16 and 17 show the attenuation and phase characteristics in the vicinity of the pass band. The center frequency of the filter is 200 kHz, the

3 dB bandwidth is 2.4 kHz, the phase difference at the center frequency is 40 degrees, and the phase inclination is 0.07 deg/Hz. The input and output impedances are 8.5 k Ω and 8.9 k Ω , respectively. Figure 18 shows the attenuation characteristics in a frequency range from 0 to 1.5 MHz. An attenuation of 30 dB or more is obtained in this wide frequency range. Although the third resonant response is relatively high, the fifth resonant response can be fully suppressed because of affixing a coupler at the fifth resonant mode nodal points. Thus, the values of the above characteristics are assured to fully satisfy the requirements.

In the next filter design, a coupler was fixed in the middle ($X_0/(l_1 + l_2) = 0.583$) of the third and fifth resonant modes. This design was intended to compensate for the attenuation at the third resonant response, which causes a high spurious response at the sacrifice of suppressing the fifth resonant response. This mechanical filter was designed by using the image parameter theory. The two transducers employed on the input and output sides of the filter have the same dimensions. Figures 19 to 21 show the characteristics of the experimental filter. In this design, the phase inclination is 0.082 deg/Hz according to Fig. 20.

Figure 22 shows the temperature characteristics of the experimental filter given in Fig. 15. In the temperature range from 5 to 60°C (reference temperature: 25°C), the characteristics are good, showing an insertion loss variation of 0.6 dB and a phase variation of 11 degrees.

Conclusion

Suggestions were made for realization of a 200 kHz timing tank filter by use of composite transducers. The results of analysis for the filter were described. Based on the design and results of analysis, several filters were constructed and examined. Experiments showed that composite longitudinal mode transducers allowed a coupler and supporting wire to be easily welded to its metal surface and improved the temperature characteristics. Moreover, the mechanical filter construction, consisting of two transducers, fully satisfied the strict requirements, including the requirement for volume.

The features of the mechanical filter and their effects on the performance of the equipment are as follows:

- Since it is small and needs no electric components (inductor, capacitor), the mechanical filter can match the other components (LSIs) in size and shape.
- Since it has wide bandwidth and high impedance characteristics, the mechanical filter permits a short rise and fall time for the output waveform and allows impedance matching with high input/output impedance CMOS circuits.
- Since it permits high attenuation (30 dB or more) over a wide frequency range (0 to 1.5 MHz), the mechanical filter improves the signal to noise ratio of the equipment.
- Because of its good temperature characteristics, the mechanical filter permits stable output waveform with a few phase variations.

Acknowledgement

The authors wish to thank Mr. Sumio Sugawara, Yamagata University, for his assistance in the

calculation of filter characteristics with valuable discussions. They also wish to thank Mr. Takamoto Kojima, Director of the Transmission Group and Mr. Masayuki Yokouchi, Director of the Technology Development Dept., FUJITSU LIMITED for their useful advice and continuous encouragement.

Appendix

- Resonance frequency and resonant mode
The resonance frequency conditions are first determined. Let the u_1 and u_2 displacements in the x direction at parts I and II shown in (b) of Fig. 1 be represented by the following equations:

$$\begin{aligned} u_1 &= \bar{A} \sin(\omega x/v_1) \\ u_2 &= \bar{B} \cos(\omega x/v_2) + \bar{C} \sin(\omega x/v_2) \end{aligned} \quad (1)$$

where v_1 and v_2 are the velocities of sound at parts I and II respectively. Assuming that the densities of piezoelectric ceramics and constant modules alloy are ρ_p and ρ_m and their elasticity constants are c_{11}^E and c_m , let the v_1 and v_2 values be expressed by the following equations:

$$v_1 = \sqrt{c_{11}^E/\rho_p}, \quad v_2 = \sqrt{c_m/\rho_p} \quad (2)$$

$$k_1 = \sqrt{(c_m t_m / c_{11}^E t_p + 1) / (\rho_m t_m / \rho_p t_p + 1)}$$

The connection conditions for parts I and II are as follows:

$$\begin{aligned} u_1 &= u_2 \\ s_1 c_1 \partial u_1 / \partial x &= s_2 c_2 \partial u_2 / \partial x \\ s_1 &= b(2t_m + 2t_p), \quad s_2 = b(2t_p) \\ c_1 &= c_{11}^E (c_m t_m / c_{11}^E t_p + 1) / (t_m / t_p + 1) \\ c_2 &= c_{11}^E \end{aligned} \quad (3)$$

Taking the above conditions into consideration, the resonance frequency conditions are determined as follows:

$$\begin{aligned} \tan(\alpha_1) \cdot \tan(\alpha_2) &= z_{01}/z_{02} \\ z_{01} &= \rho_1 s_1 v_1, \quad z_{02} = \rho_2 s_2 v_2 \\ \rho_1 &= \rho_p (\rho_m t_m / \rho_p t_p + 1) / (t_m / t_p + 1) \\ \rho_2 &= \rho_p \end{aligned} \quad (4)$$

$$\alpha_1 = \omega l_1 / v_1, \quad \alpha_2 = \omega l_2 / v_2$$

Determine the α_1 value as a function of t_m/t_p and l_2/l_1 using equations (4). The resonance frequency f_n can then be calculated by the following equation:

$$f_n = \alpha_1 / 2\pi l_1 (c_{11}^E / \rho_p)^{1/2} \cdot k_1 \quad (5)$$

n: Resonant mode order

The resonant mode is therefore determined, using the undetermined constant \bar{A} of equation (1), by the following equations:

$$\begin{aligned} u_1 &= \bar{A} \sin(\omega x/v_1), \quad 0 \leq x \leq l_1 \\ u_2 &= \bar{A} \sin(\alpha_1) / \cos(\alpha_2) \times \cos(\alpha_2(1 + l_1/l_2) \\ &\quad (x/l_1 + l_2 - l_1)), \quad l_1 \leq x \leq (l_1 + l_2) \end{aligned} \quad (6)$$

In these normalized functions of displacements, the undetermined constant \bar{A} of equations (6) can be obtained by the following equations:

$$\begin{aligned} A &= \sqrt{2 k_2} \\ k_2 &= \sqrt{\frac{1 + (l_2/l_1)/(1 + t_m \rho_m / t_p \rho_p)}{1 - \frac{\sin 2\alpha_2}{2\alpha_1} + \frac{\sin(2\alpha_2/2\alpha_1) l_2}{1 + t_m \rho_m / t_p \rho_p} \frac{\sin^2 \alpha_1}{l_1 \cos^2 \alpha_2}}} \end{aligned}$$

- (b) Equivalent inductance and capacitance ratio
The equivalent inductance L and capacitance ratio r are calculated in the following way:
The basic piezoelectric equation is as follows:

$$D_3 = \epsilon_{33} \frac{\partial \phi}{\partial x} + \frac{e}{2\pi} \frac{\partial u}{\partial x} \quad (8)$$

in which ϕ is the electric potential.

According to the above equation, transducer charge Q_m is determined, taking the total length of the transducer into account, by the following equation:

$$Q_m = 2\epsilon_{33}b \int_0^{l_1} \frac{\partial \phi}{\partial x} dx \quad (9)$$

The kinetic energy T_m of the transducer is obtained by:

$$T_m = \frac{1}{2} \rho_p (2t_p + 2t_m) \int_0^{l_1} u^2 dx \quad (10)$$

so that it can also be expressed, using the electromagnetic energy obtained by equation (9), by the following equation:

$$T_m = \frac{1}{2} L (-\omega^2) Q_m^2 \quad (11)$$

Then, equivalent inductance L is determined by the following equations:

$$\frac{L}{L_0} = \frac{k_3}{2} \left(1 + \frac{t_m \rho_m}{t_p \rho_p}\right), L_0 = 2t_p l_1 \rho_p b^2 / (2\epsilon_{33} b)^2 \quad (12)$$

$$k_3 = \left(1 - \frac{\sin 2\alpha_1}{2\alpha_1} + \frac{l_2/l_1}{(1 + \frac{\rho_m t_m}{\rho_p t_p})} \cdot \frac{\sin^2(\alpha_1)}{\cos^2(\alpha_2)}\right) \times (1 + \sin 2\alpha_2 / 2\alpha_2) / \sin^2(\alpha_1)$$

where L_0 is the equivalent inductance when there is no metal portion ($2t_m = 0$) at the third resonant mode suppression electrode ($l_2/l_1 = 0.5$).

Now that the equivalent inductance L has been calculated, capacitance ratio, r , can be determined easily. According to equations (5) and (12), the following conditions are assumed:

$$r = \frac{C_0}{C}, C = \frac{1}{L\omega^2}, C_0 = \frac{e^2 T (2t_p)^2}{2\epsilon_{33} b^2} \quad (13)$$

Therefore, capacitance ratio r is determined by the following equations:

$$r/r_0 = \frac{1}{k_3} \left(1 + \frac{t_m \rho_m}{t_p \rho_p}\right) k_3 / 2\alpha_1^2 \quad (14)$$

$$r_0 = \frac{1}{2} / k_3^2$$

k_3 : Coupling coefficient

where r_0 is the capacitance ratio when there is no metal portion ($2t_m = 0$) at the third resonant mode suppression electrode ($l_2/l_1 = 0.5$). Equation (14) can also be used for determination of the step ratio that can

suppress the third resonant mode when the response is required to be flat at the third resonance (third resonant mode).

Equivalent mass m and force factor A for the transducer under the third resonant mode suppression condition can also be determined easily as values for the current total length of the transducer, in the following manner:

Equivalent mass: $m = \bar{m} k_4$

$$\bar{m} = (\rho_p 2t_p) b \cdot 2l_1 (1 + l_2/l_1) / 2 = (\rho_p 2t_p) b l_1 / 2$$

$$k_4 = \frac{2}{3} \left(1 + \frac{t_m \rho_m}{t_p \rho_p}\right) \frac{\cos^2(\alpha_2)}{\sin^2(\alpha_1)} \times \left(1 - \frac{\sin 2\alpha_1}{2\alpha_1}\right) + \quad (15)$$

$$\frac{l_2/l_1}{(1 + \frac{t_m \rho_m}{t_p \rho_p})} \cdot \frac{\sin^2 \alpha_1}{\sin^2 \alpha_2} \left(1 + \frac{\sin 2\alpha_2}{2\alpha_2}\right)$$

$$(l_2/l_1 = 0.5)$$

$$\text{Force factor: } A = A_0 \cos(\alpha_2) / \sqrt{3} \quad (16)$$

$$A_0 = 2 \epsilon_{33} b \sqrt{3/2} \quad (l_1/l_2 = 0.5)$$

where \bar{m} and A_0 are the equivalent mass and force factor for a single substance of piezoelectric ceramics ($2t_m = 0$).

Reference

- 1) R. Komiya: Trans. Inst. Electron. Commun. Eng. Japan, vol. 64, No. 11, PP1163 - 1168 (1981)
- 2) H. Inose: Inst. Trans. Electron. Commun. Eng. Japan, vol. 64, No. 11, PP1124 - 1129 (1981)
- 3) Y. Tomikawa, et al: Paper of Technical Group on Ultrasonics, IECE of Japan, 83 - 3 (1983)
- 4) K. Nagai and M. Konno: Electromechanical Vibrators and Their Applications (in Japanese), Corona Company, 1974
- 5) T. Kawana, et al: IECE Ultrasonics Symposium, D - 6 (1979)
- 6) Zverer: Handbook of Filter Synthesis, John Wiley and Sons INC (1967)
- 7) Y. Katsube, et al: Denchi Kagaku, Japan (in Japanese), vol. 4, PP21 - 27 (1980)

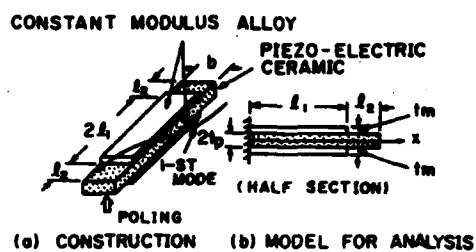


Fig. 1 Construction of Transducer

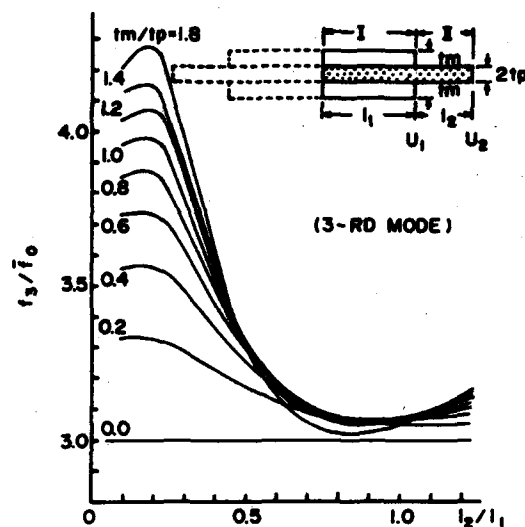


Fig. 3 Normalized resonant frequency of the 3-rd mode.

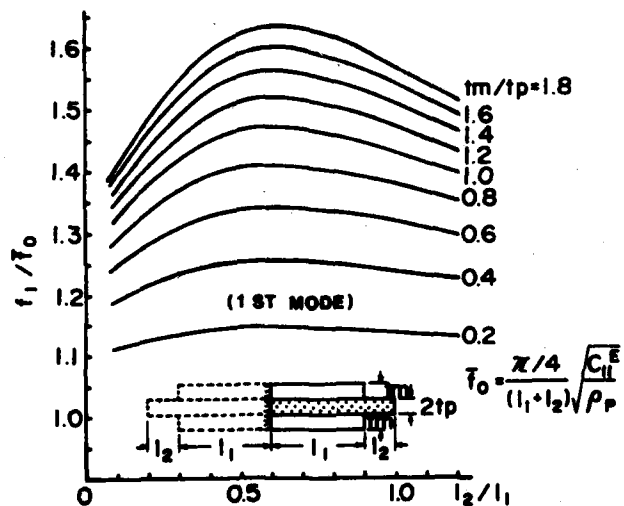


Fig. 2 Normalized resonant frequency

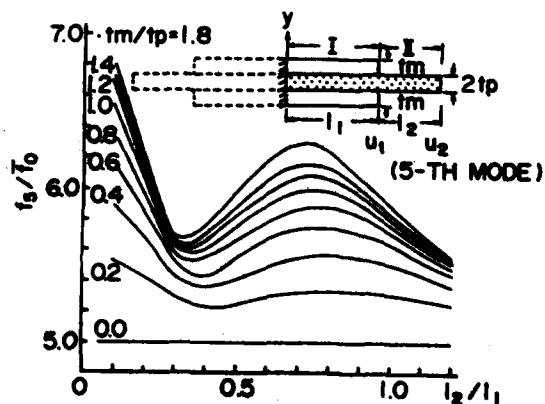


Fig. 4 Normalized resonant frequency of the 5-th mode.

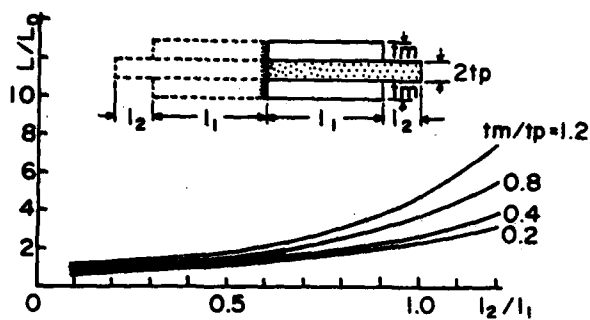


Fig. 5 Normalized equivalent inductance versus l_2/l_1 .

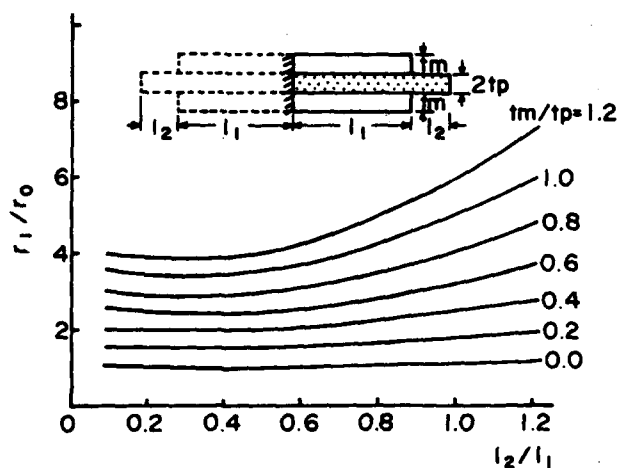


Fig. 6 Normalized capacitance ratio versus l_2/l_1 .

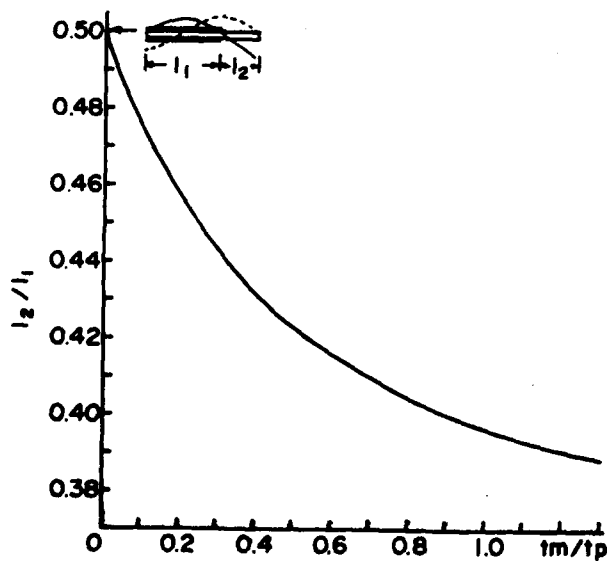


Fig. 7 l_2/l_1 versus tm/tp at the 3-rd mode suppression.

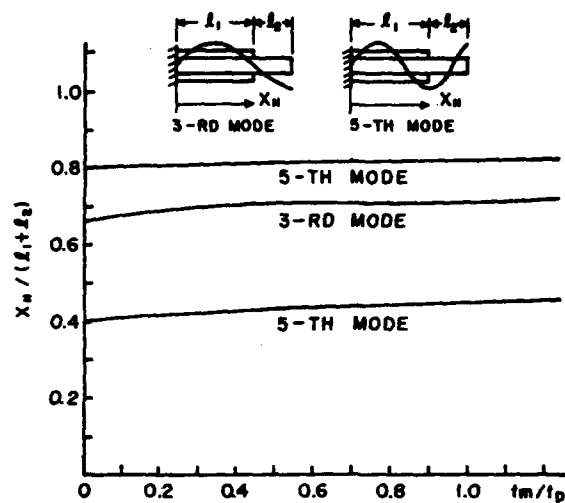


Fig. 8 Nodal Point at the 3-rd and 5-th Mode

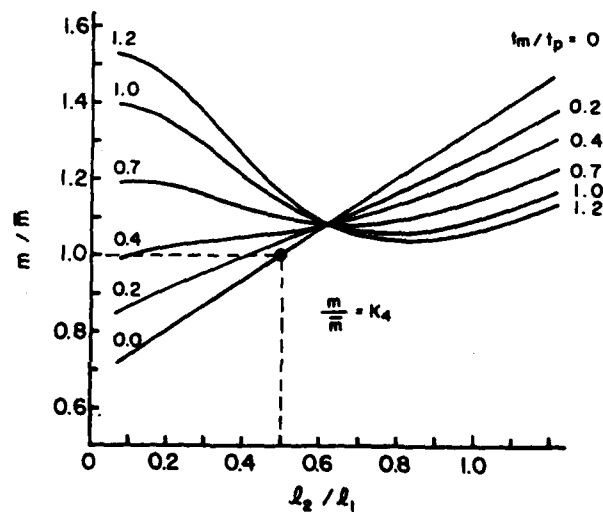


Fig. 9 Normalized Equivalent Mass Versus l_2/l_1

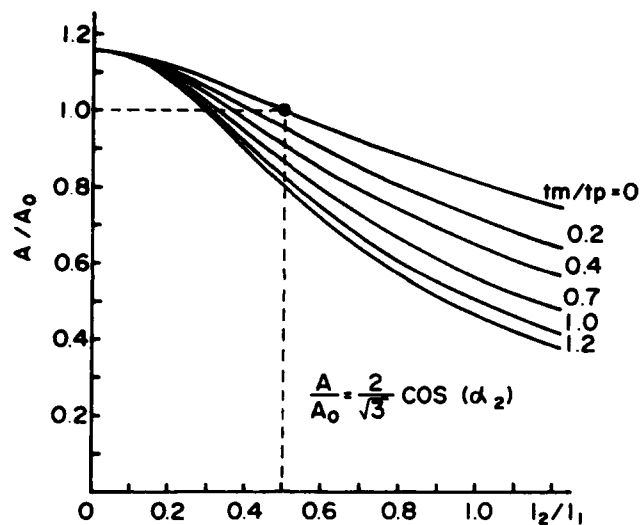
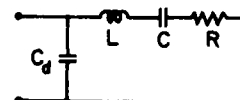


Fig.10 Normalized force factor versus l_2/l_1 .

Table1 Dimensions of vibrators

$2l_1/l_2$	$2l_1$	$2l_2$	l_2/l_1	tm	$2tp$	tm/tp	b
12.5	9.0	3.5	0.389	0.25	0.4	1.25	1.0

UNIT OF DIMENSIONS : mm



L : EQUIVALENT INDUCTANCE
 C : EQUIVALENT CAPACITANCE
 R : EQUIVALENT RESISTANCE
 $f_1 = 1/2\pi\sqrt{LC}$: RESONANT FREQUENCY
 C_d : DAMPED CAPACITANCE
 $\gamma = C_d/C$: CAPACITANCE RATIO
 $Q = 2\pi f_1 L/R$: QUALITY FACTOR

Fig.12. Equivalent Circuit of the Transducer

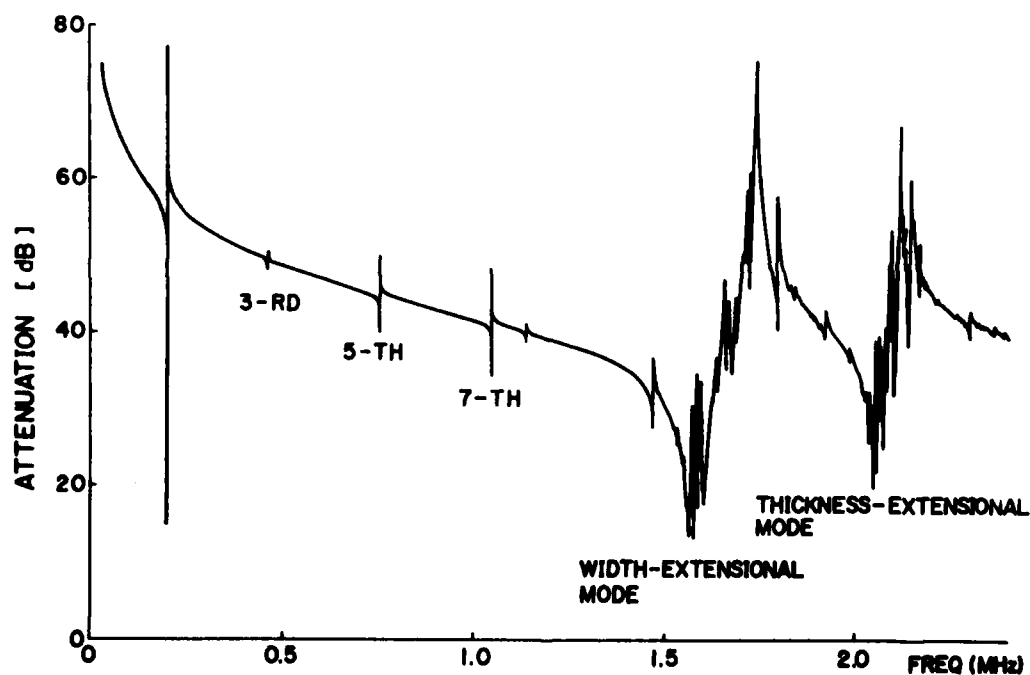


Fig.11. Frequency Characteristic of One Transducer

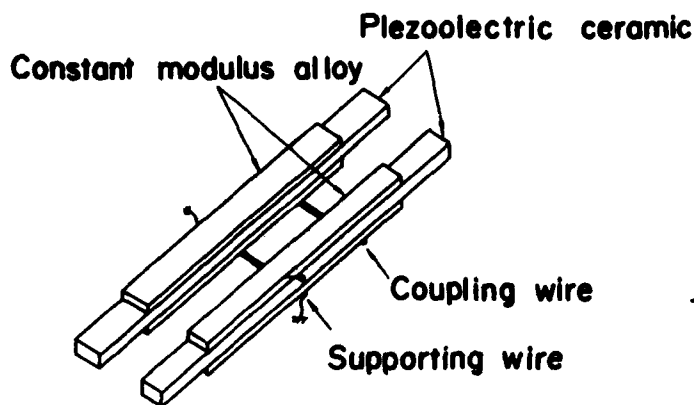


Fig.13 Construction of mechanical filter

Table3 Comparison between calculated and measured values.

	CALCULATED	MEASURED
f_3/f_1	2.39	2.35
f_3/f_1	3.06	3.00
T/r_0	3.42	3.56~3.70
L/L_0	1.44	0.94~1.01

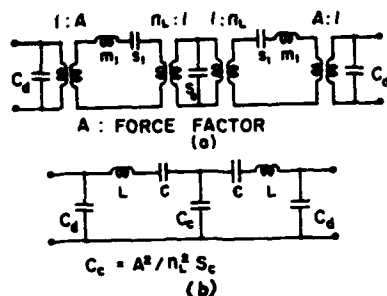


Fig.14 Equivalent Circuit of the Filter in Fig.13

Table 2 Examples of measured equivalent circuits

	SAMPLE NO.1	SAMPLE NO.2
f_1 (kHz)	199.152	199.784
L (mH)	173.0	186.5
C (pF)	3.69	3.40
R (Ω)	71.9	84.3
Q	3010	2780
C_d (pF)	184	176
T	49.8	51.7

PIEZOELECTRIC CERAMICS :
Nepoc-4 $K_{31}=0.28$

Table 4 Design Procedure Using Butterworth Filter

(a)		$L_1 = \frac{R}{2\pi\omega} S_1$ $C_1 = \frac{\Delta f}{2\pi f_0^2 R} \frac{1}{S_1}$
(b)		$L_2 = \frac{\Delta f R}{2\pi f_0^2} \frac{1}{S_2}$ $C_2 = \frac{1}{2\pi\omega R} S_2$ $G = \sqrt{\frac{C_2}{L_1}}$ $C_c = \frac{G}{2\pi f_0}$
(c)		$L = \frac{L_1}{C_c C_1}$ $C = \frac{C_c C_1}{C_c - C_1}$ $f_T = 1 / 2\pi\sqrt{LC}$
(d)		$f_T' = f_T \sqrt{1 - \frac{1}{X}}$ $X = \frac{C_d/C}{\frac{1 \pm \sqrt{1 - 4Q^2(1+Q^2)}}{270^2}}$ $\alpha = RC \cdot 2\pi f_T$ $r = \frac{C_d}{C}$ $R' = 1 / 2\pi f_T' C_d \alpha x$ $C' = 1 / (\frac{1}{C} - \frac{1}{C_d})$
(e)		

f_0 : Center frequency Δf : Band width

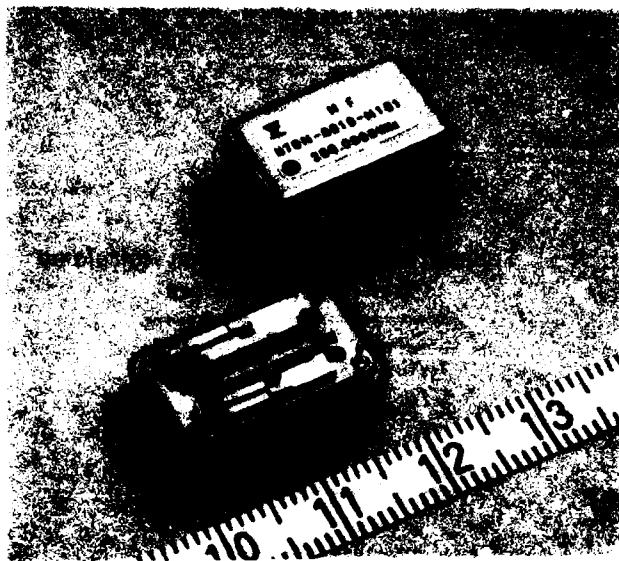


Fig. 15 Appearance

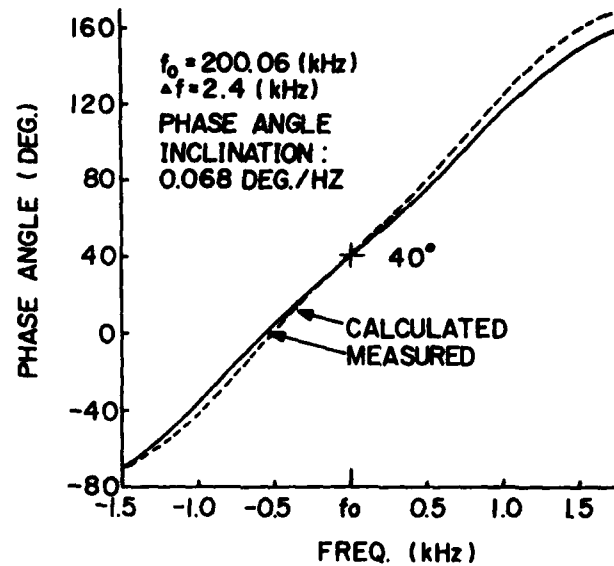


Fig. 17 Phase Characteristic.

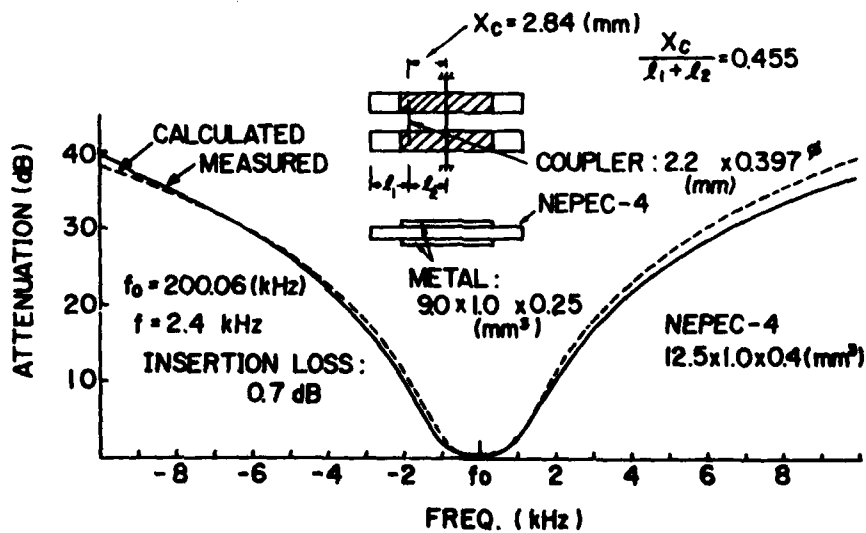


Fig. 16 Frequency Characteristic.

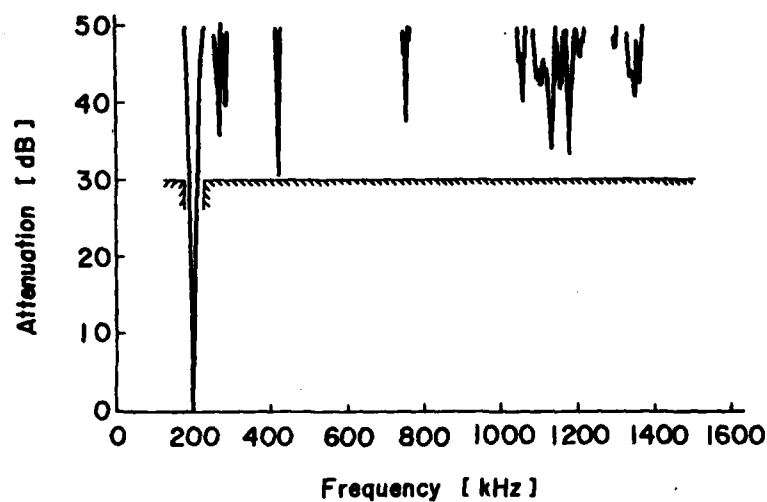


Fig. 18 Measured Frequency Characteristic.

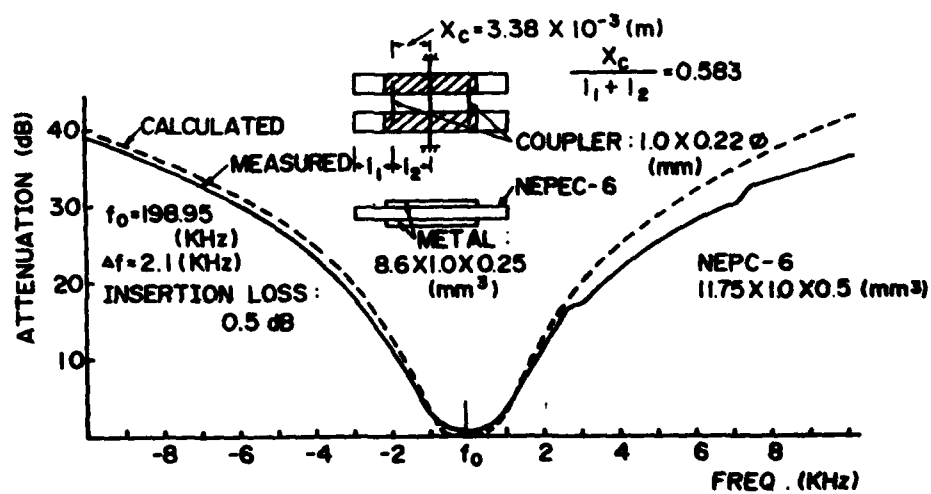


Fig. 19 Frequency characteristic.

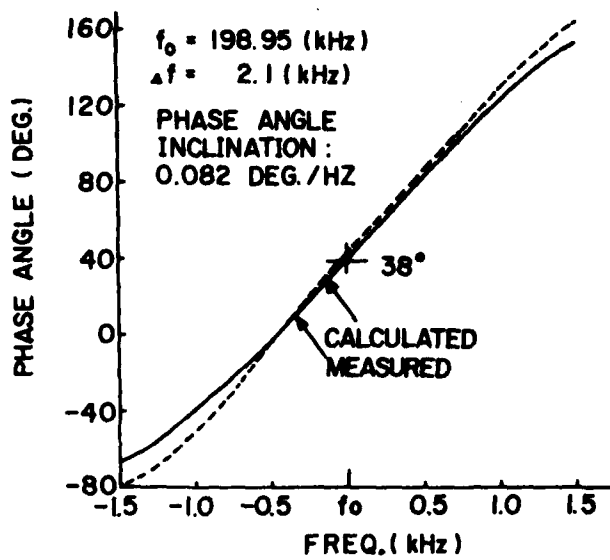


Fig. 20 Phase Characteristic.

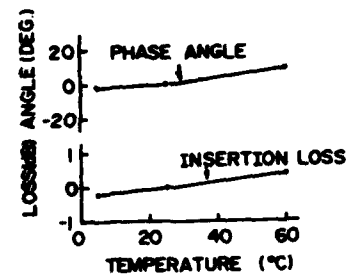


Fig. 22 Temperature characteristic.

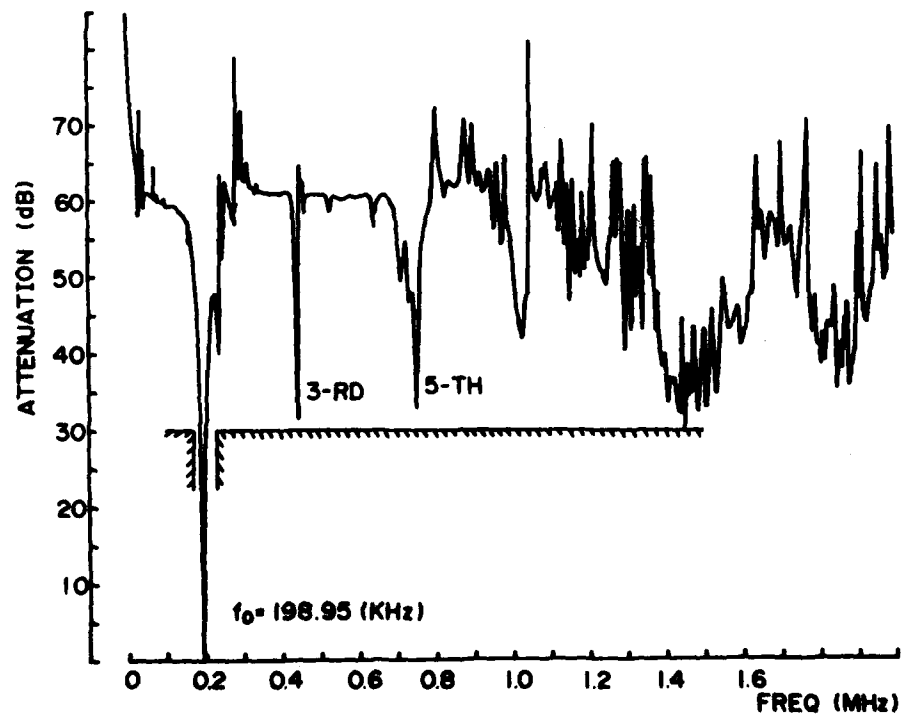


Fig. 21 Measured frequency characteristic.

AD P002504

STRIPLINE FILTERS - AN OVERVIEW

Lawrence N. Dworsky

Motorola, Inc.
Schaumburg, Ill.

ABSTRACT

Stripline filters and the closely related microstrip filters are coupled transmission line networks which are fabricated using one of a number of "printed circuit" fabrication techniques. In order to understand the operation and parameters of stripline filters it is necessary to understand the fundamentals of single and coupled transmission line resonators.

The fundamental properties of transmission line resonators are reviewed, and specific examples are given for typical stripline parameters. Basic relationships are given for the size, characteristic impedance, and Q of stripline resonators. In general it is found that stripline resonator Q will be limited by the ohmic losses in the conductors. The characteristics of the skin effect and the proximity effect which determine these losses are reviewed.

Several examples of coupled stripline filters are shown.

INTRODUCTION

Stripline circuit design and manufacture, begun in the early 1950's, has evolved into a cost effective, compact, and repeatable transmission line technology. The subject of stripline filters can be broadly broken into three subcategories. These are:

1. Properties of TEM resonator systems.
2. Unique electrical and mechanical characteristics of the stripline system.
3. Synthesis of desired filter response in transmission line networks.

Each of these subcategories deserves a treatment much, much more complete than is reasonable for an overview paper such as this. Therefore, some choices had to be made. The first choice was to not address the synthesis problem. The reason for this is that of the three subcategories above, this one is the least specific to stripline technology. There are many excellent references, but practical considerations dictate that the list of references be kept modest. Therefore, we list only two references in this area.^{1,2}

Of the remaining two subcategories, we will draw generalizations from the subject of TEM resonators

and apply them to the stripline system as required.

In general, a transmission line with a uniform, homogeneous, isotropic dielectric will support a mode characterized by transverse electric and magnetic fields (TEM mode) at all frequencies. When the dimensions of the structure are such that waveguide modes begin to propagate at some frequency, then we may consider that frequency as the upper practical limit of the transmission line as a TEM system. The basic properties of the TEM transmission line are described by the inductance and capacitance per unit length, L and C. These parameters are frequency independent, being determined only by the geometry of the system and the dielectric constant of the medium (we are not considering materials with a permeability other than unity).

L and C are related to the velocity of propagation of the line by the relationship

$$LC = 1/v^2 = \epsilon_r/v_0^2 \quad (1)$$

and to the characteristic impedance of the line by

$$\sqrt{L/C} = Z_0 \quad (2)$$

where v_0 = the speed of light
 ϵ_r = the (relative) dielectric constant
 v = the propagation velocity of the line
 Z_0 = the characteristic impedance of the line.

Figure 1 shows, in cross section, the coaxial cable and the strip transmission line (stripline). The coaxial cable is probably the most common TEM transmission line. These two lines are very similar: Both consist of an enclosing outer, or "ground" conductor, in inner or "center" conductor, and a dielectric medium separating the two conductors. Insofar as their basic characteristics, they are the same - only the particular parameters and the methods of manufacture are different.

As is shown in figure 1, the cross sectional dimensions of the stripline are w - the width of the center conductor, t - the thickness of the center conductor, and b - the height, or thickness, of the box. Typically the width of the box is so much greater than b that its actual dimension

does not enter into any calculations.

Since filter structures are typically built using resonant lengths of line, the length of such a resonant length is a measure of the size that a filter structure will be. Figure 2 shows the lengths of quarter wave resonators vs. frequency, for several common dielectric constants. An air dielectric, $\epsilon_r = 1$, is rarely used for stripline structures, but is included for perspective. $\epsilon_r = 2.3$ is typical of the various teflon, glass-filled teflon, and polyolefin dielectric materials. $\epsilon_r = 10$ is an approximate value for alumina, which is probably the most common ceramic dielectric. $\epsilon_r = 36$ is an approximate value for barium titanate, a ceramic which is beginning to see wide acceptance as a dielectric material.

STRIPLINE PARAMETERS

The most widely used formulas for stripline parameters were first presented by Cohn.³ These formulas are approximations to exact formulas which are expressed in terms of elliptic integrals. Cohn showed that, for $w/(b-t) < .35$,

$$Z_0 \sqrt{\epsilon_r} = 60 \ln(4b/\pi d) \quad \text{ohms} \quad (3)$$

$$\text{where} \\ d = \frac{w}{2} \left[1 + \frac{t}{\pi w} \left(1 + \ln \frac{4\pi w}{t} + .51\pi \left(\frac{t}{w} \right)^2 \right) \right] \quad (4)$$

and that for $w/b > .35$,

$$Z_0 \sqrt{\epsilon_r} = \frac{94.15}{1 - t/b + \frac{C'}{.0885 \epsilon_r}} \quad \text{ohms} \quad (5)$$

$$\text{where} \\ C' = \frac{.0885}{\pi} \left[\frac{2}{1 - t/b} \ln \left(\frac{1}{1 - t/b} + 1 \right) - \left(\frac{1}{1 - t/b} - 1 \right) \ln \left(\frac{1}{(1 - t/b)^2} - 1 \right) \right] \quad (6)$$

Note that equations (3) through (6) all depend on ratios of dimensions. Therefore actual units do not matter, so long as all dimensions are expressed in the same units.

Figure 3 shows $Z_0 \sqrt{\epsilon_r}$ as calculated from the above formulas, for two cases of t/b . The value $t/b = .04$ is probably as large a value as would be seen in practice. For $b = 100$ mils, for example, t would be 4 mils. This is quite thick for a laminated or deposited strip of some sort. The range $0 < t/b < .04$ therefore encompasses most practical cases. As may be seen, the line thickness is not ignorable, but on the other hand is not very consequential insofar as Z_0 is concerned.

Ohmic losses in stripline conductors are dependent upon three contributing factors. The first factor is the conductivity of the conductors themselves. Copper has the best conductivity (other than very expensive silver) available, and is almost universally chosen as the stripline conductor material.

The second factor contributing to the ohmic loss is the Skin Effect. That is, at RF frequencies, currents in conductors flow close to the surface of the conductors. In the case of stripline they flow on the outside of the center conductor, and on the inside of the outer conductor. The current falls off exponentially with depth into the conductor, the exponential decay having a characteristic length called the Skin Depth.⁴ The skin depth, δ , is shown to be

$$\delta = 1/\sqrt{\pi f \mu \sigma} \quad (7)$$

where f is the frequency
 μ is the permeability of the conductor
 σ is the conductivity of the conductor

For an infinite half-plane conductor, the total effective resistance is given by

$$R_s = 1/\sigma \delta \quad (8)$$

R_s is the lowest resistance that a conductor can exhibit at a given frequency - and then only if it is infinitely thick. For all practical purposes, R_s is achieved once the conductor is several skin depths thick. Figure 4 shows δ and R_s vs. frequency for a copper conductor.

It should be pointed out at this point that in the case of the stripline center conductor, the thickness t should not be confused with the conductor thickness described here. The center conductor should be pictured as a rectangular box, of dimensions $W \times t$, and the conductor thickness refers to the thickness of the walls of this box.

The final factor contributing to the ohmic loss is sometimes called the Proximity Effect. That is, the current distribution in a stripline is not uniform in the conductors. The current tends to "bunch" at the edges of the center conductor, and directly above (and below) the center conductor in the outer conductor.

Wheeler⁵ has shown that the resistance per unit length of a transmission line may be calculated from the inductance per unit length and R_s by the formula

$$R = \frac{R_s}{\mu} \sum \frac{\partial L}{\partial x_i} \quad (9)$$

where the x_i are the dimensions of the line as measured from some convenient reference point. In the case of stripline, it is most convenient to take the reference point as the geometric center of the line cross section, and then $x_1 = t/2$, etc. Also, by using (1) and (2), L is known if Z_0 is known.

Dielectric losses in stripline are directly related to the dielectric loss tangent of the dielectric material, $\tan(\delta)$. Sometimes this loss tangent is expressed in terms of the Q of the dielectric material, and then

$$Q_{\text{diel}} = 1/\tan(\delta) \quad (10)$$

The above quantity is usually frequency dependent, the exact relationship depending on the particular material.

The shunt loss per unit length of a transmission line is given by

$$G = \omega C / Q_{\text{diel}} \quad (11)$$

and then the Q of a quarter-wave resonator is⁶

$$Q = \frac{\omega LC}{R/Z_0 + GZ_0} \quad (12)$$

Figure 5 shows stripline resonator Q (copper conductor) vs. the box height, b , for the particular case specified in the figure. Z_0 is kept constant, therefore w was varied along with b . As may be seen, increasing b increases Q . Also, increasing b at constant Z_0 increases w . This means that the cross-sectional area of the stripline and therefore the volume of a stripline resonator structure at a given Z_0 will increase superlinearly with Q .

Figure 6 shows stripline resonator Q (copper conductor) vs. frequency, for several different values of ϵ_r . In this case b and Z_0 are kept constant. This means that w was corrected for the different values of ϵ_r . Here we see that Q increases with frequency, but that a higher dielectric constant, which shrinks the dimensions of the structure, lowers the Q .

For the cases shown in both figures 5 and 6, it was assumed that ohmic (conductor) losses were the only losses to consider. This is an accurate assumption when considering almost all of the high quality dielectric materials available today. In the case of low quality printed circuit board material, or perhaps some ceramic compositions, this assumption would not be reasonable. Equation 12, however, would still be correct.

FILTER STRUCTURES

Figure 7 shows two types of stripline low-pass filters, and their approximate equivalent lumped circuit models. The first type is based upon the approximations that a short thin length of line acts as a series inductance, while a short wide length of line acts as a shunt capacitance. The second type uses these same approximations, but in a more complicated structure, and is actually a realization of an elliptic function filter.

Figure 8 shows two realizations of a stripline high-pass filter, and their (common) circuit model. As in the case of the low pass filters, the thin sections of line act as series inductors. The grounding of one end of a line is typically accomplished using a "plated through" hole to both upper and lower outer conductors. The difference between these two realizations is the manner in which the series capacitance is constructed. The zig-zag break in the lines shown in the first case

is a very easy capacitance to realize - it is defined using the same technique, and at the same time, as the rest of the line definitions. The zig-zag structure is used to maximize the gap capacitance. Unfortunately, even with zig-zag structures, often sufficient series capacitance cannot be obtained using gap capacitance. In this case the second approach shown in figure 8 is used. Here, a "patch" of conductor material is laid over an insulating thin film of some sort, and then the insulating film is laid across a gap in the line. Various insulating sheet materials, metalized on one side and adhesive backed on the other are available for this purpose. In most cases they are thin enough so that the final assembly of the stripline filter is not at all affected by their presence.

Figure 9 shows a direct-coupled bandpass filter, generally most useful for wideband (> 25%) applications. This filter is comprised of a string of shorted quarter wave resonators, each of which acts as one pole of the filter. The resonators are separated by quarter wave line sections, which act as impedance inverters.

Figure 10 shows two different layouts of a coupled-resonator bandpass filter. In this case each line is one half wavelength long. This type of filter is most useful for narrow to medium bandwidth applications. The parameters of coupled transmission lines are described in terms of "even" and "odd" mode characteristic impedances in the case of identical line sections. These impedances are the normal modes of the coupled system. Formulas for the side, or edge, coupled lines were derived by Cohn⁷ who presents the formulas both analytically and in the form of nomographs.

Figure 11 shows a coupled-resonator bandpass filter made up of quarter wave lines. It is essentially equivalent to the structure of figure 10. This structure, known as an "interdigital" filter has the advantage of being smaller than the structure of figure 10. On the other hand, shorting the lines as required is not an easy job, and therefore the structure of figure 10 is both less expensive to build and more repeatable.

CONSTRUCTION TECHNIQUES

Stripline construction can be broadly divided into two categories, depending upon the mechanical properties of the dielectric material. The first category includes the various teflon, filled teflon, polyolefin, etc. materials which may be generally categorized as "soft" materials. These materials are supplied in sheets, with copper metalization laminated to both sides of the sheets. Metal patterns are defined and etched using standard printed circuit board technology. The stripline is then assembled by drilling as many holes as necessary, and bolting the sheets together. In most cases stiff metal backing plates are added to either side for rigidity. When a stripline filter is a subassembly of a larger system, often a metal casting is provided with "nests" for the various subassemblies. In this case the sheet of dielectric containing the stripline filter is cut to the

correct size and dropped into its nest. When all the subassemblies are in place and connected, a cover is attached.

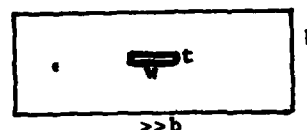
The second category of stripline construction includes the various ceramic materials, such as alumina. These materials are quite hard and brittle, and cannot be counted on to deform about a center conductor when a stripline sandwich is formed. Typically, ceramic substrates are supplied unmetallized. Conducting metalization must be applied to the ceramic, using some form of thin film deposition, thick film deposition, or chemical bath. Patterning of conductors is achieved using either shadow masking, silk screening, or photolithographically defined etching, as appropriate. When forming the stripline sandwich, it is necessary in this case to bond mirror image center conductor patterns directly to each other. This may be accomplished by tinning and soldering the layers, or with a conductive adhesive. Note that in this type of structure a thin air layer is left trapped between the dielectric layers, in between the center conductors. This air layer will lower the "effective" dielectric constant of the system, and must be compensated for in the design.

CROSS-SECTIONS OF (2) UNIFORM TRANSMISSION LINES

COAXIAL CABLE

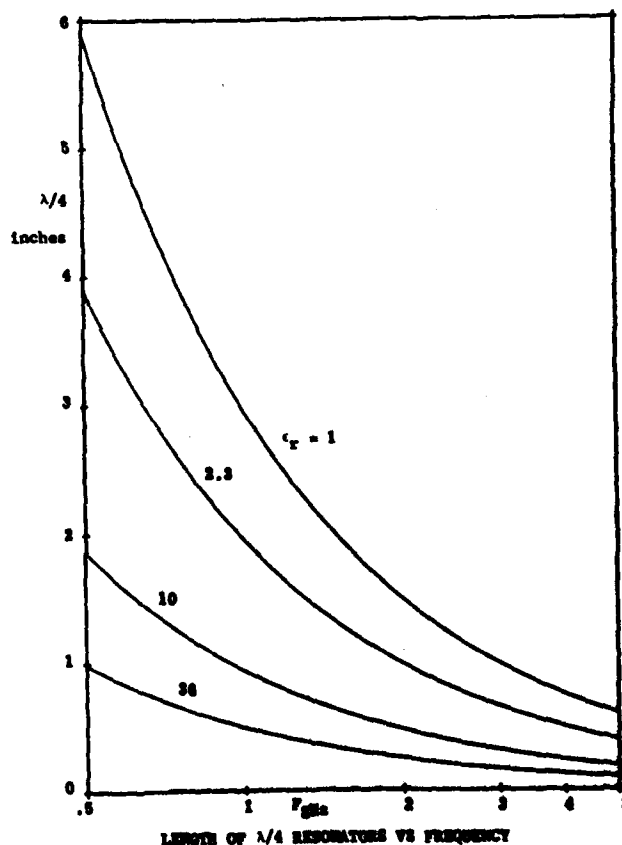


STRIPLINE

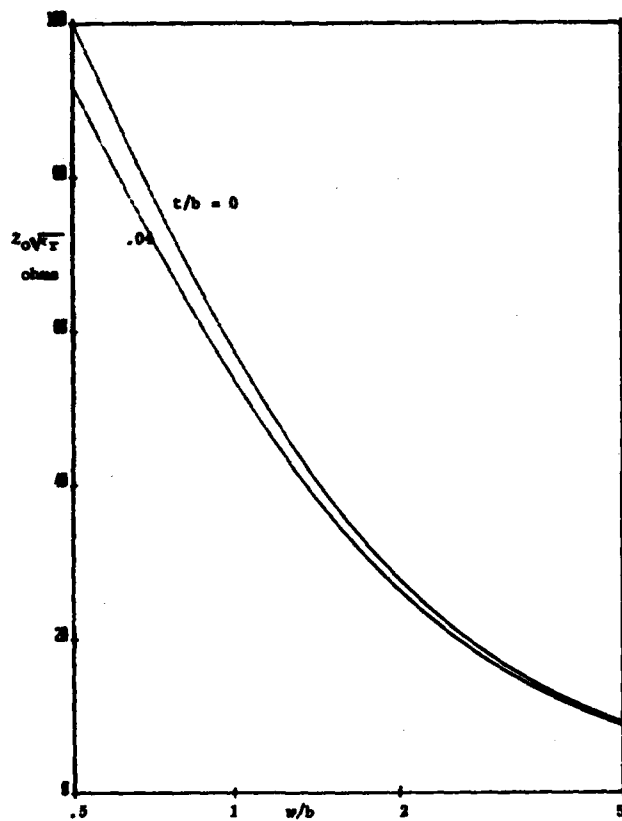


REFERENCES

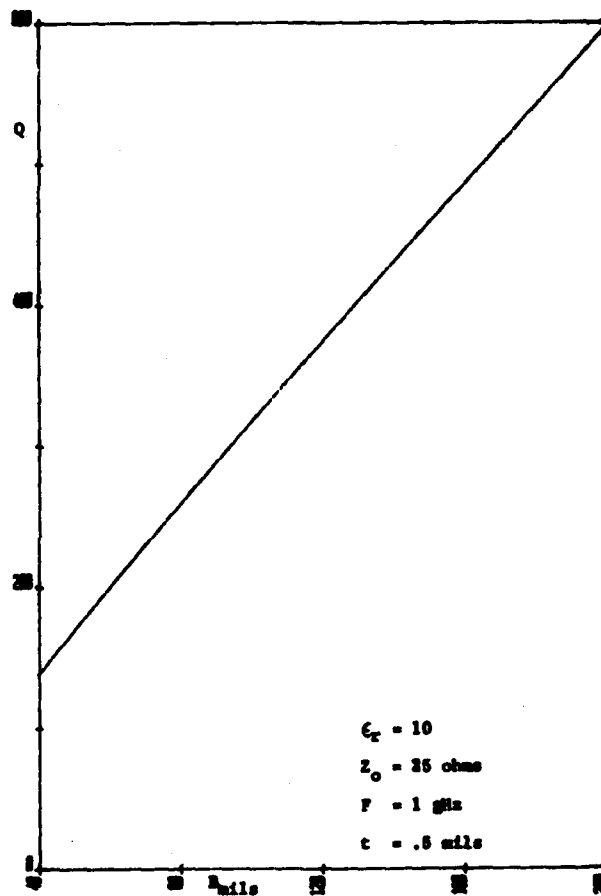
1. Temes & Mitra, "Modern Filter Theory and Design." John Wiley & Sons, New York, 1973.
2. Matthaei, Young & Jones, "Microwave Filters, Impedance Matching Networks and Coupling Structures." McGraw-Hill, New York, 1964.
3. S. Cohn, "Problems in Strip Transmission Lines." IEEE Transactions on Microwave Theory and Techniques, Vol. 3, #2, March, 1955.
4. L Dworsky, "Modern Transmission Line Theory and Applications." John Wiley & Sons, New York, 1979.
5. H. Wheeler, "Formulas for the Skin Effect." Proceedings of the IRE, Vol. 30, #9, Sept., 1942.
6. Ramo, Whinnery & Van Duzer, "Fields and Waves in Communications Electronics." John Wiley & Sons, New York, 1965.
7. S. Cohn, "Shielded Coupled Strip Transmission Line." IEEE Transactions on Microwave Theory and Techniques, Vol. 5, #10, Oct., 1957.



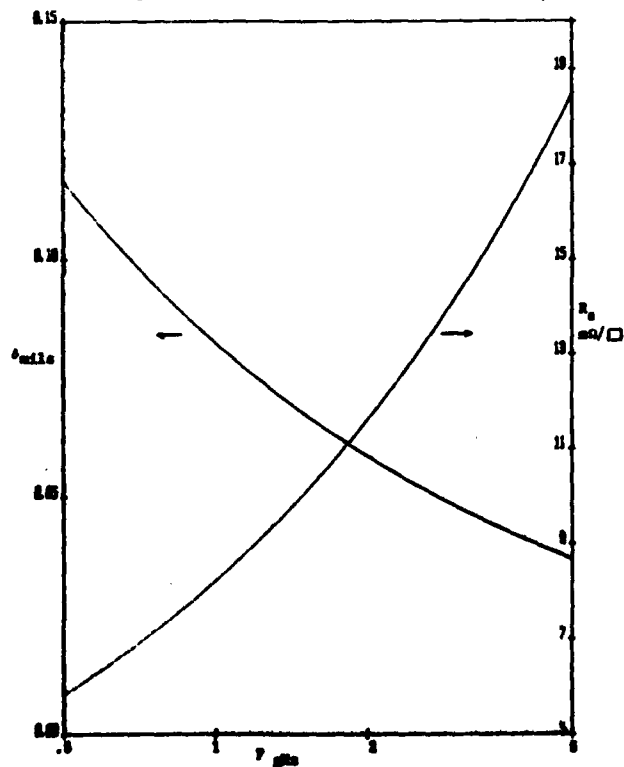
CHARACTERISTIC IMPEDANCE OF STRIPLINE



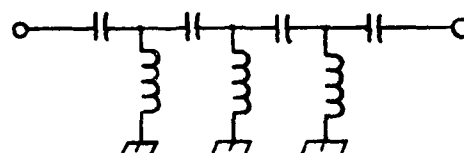
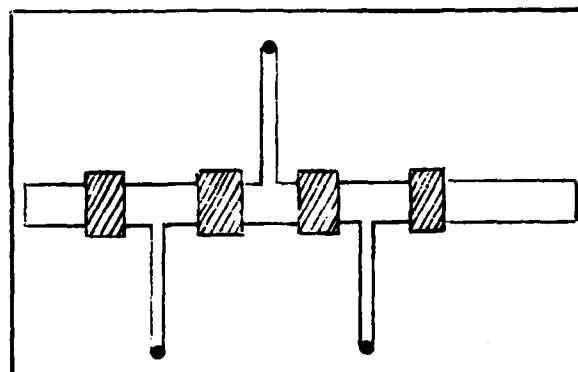
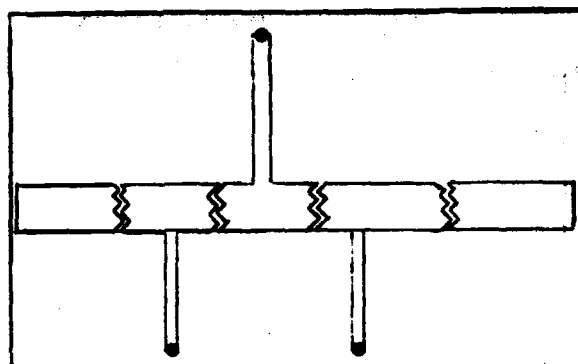
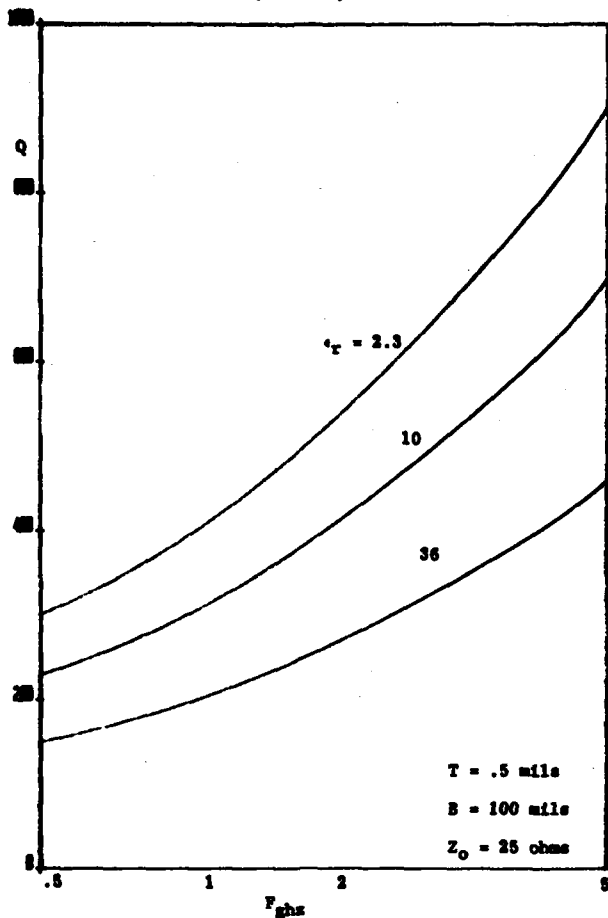
RESONATOR Q VS BOX HEIGHT



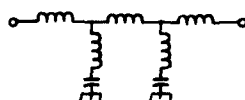
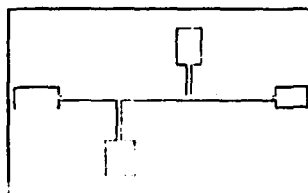
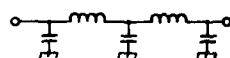
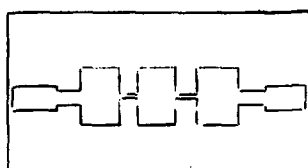
SKIN DEPTH AND SHEET RESISTANCE OF COPPER VS FREQUENCY



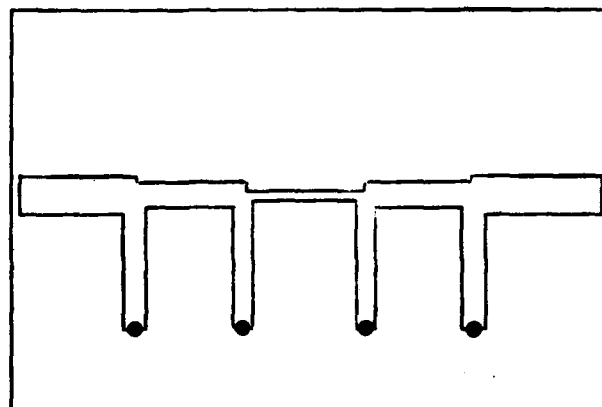
RESONATOR Q vs FREQUENCY



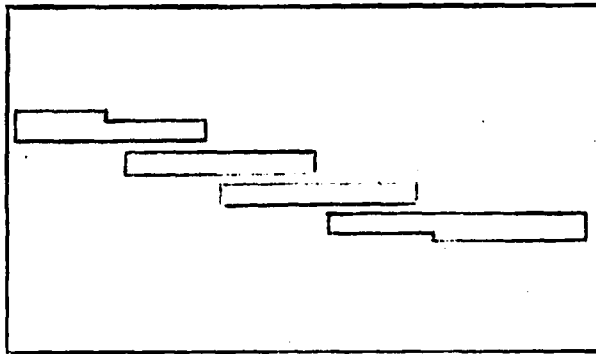
TWO HIGH PASS FILTERS AND THEIR APPROXIMATE LUMPED ELEMENT EQUIVALENT



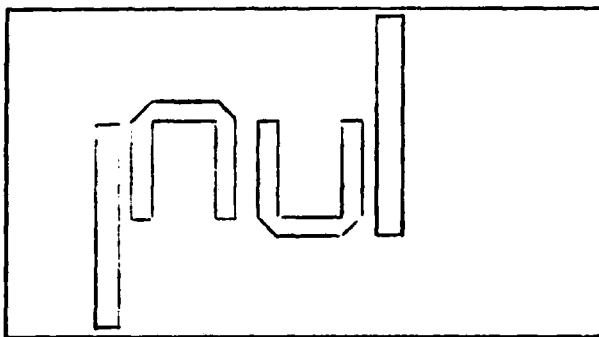
TWO LOW-PASS FILTERS AND THEIR APPROXIMATE LUMPED ELEMENT EQUIVALENT



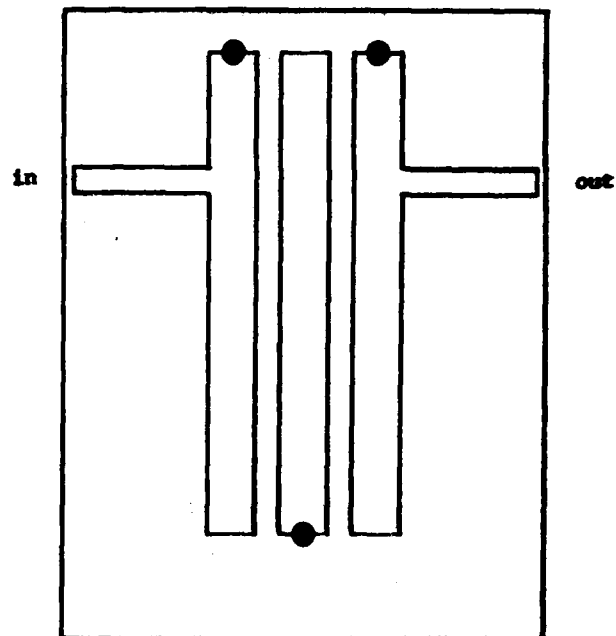
A DIRECT-COUPLED BANDPASS FILTER USING QUARTER WAVE RESONATORS AND RESONATOR SEPARATIONS



TWO LAYOUTS OF A COUPLED-LINE BANDPASS FILTER



TWO LAYOUTS OF A COUPLED LINE BANDPASS FILTER USING HALF WAVE RESONATORS. NOTE THAT NO GROUNDING HOLES ARE REQUIRED



AN "INTERDIGITAL" BANDPASS FILTER

RESONATORS ARE $\lambda/4$ LONG

SOLID CIRCLES REPRESENT METALIZED HOLES



PRECISION L-BAND SAW OSCILLATOR FOR SATELLITE APPLICATION

THOMAS O'SHEA, VERNON SULLIVAN AND ROBERT KINDELL

SAWTEK INC.

POST OFFICE BOX 18000

ORLANDO, FLORIDA 32860

AD P002505

Abstract

A 740 MHz SAW oscillator is developed for application as a fixed local oscillator in the transmitter of a satellite system. The SAW oscillator, in a 2.4 cubic inch package contains a heater, second stage amplifier/buffer, power leveling and harmonic suppression filtering. The advantages of this system are the reduced size and complexity due to the elimination of the multiplier chains required in bulkwave oscillators, a short term stability of 2×10^{-10} for 100 second averaging times, low single side band noise floor and 1.1:1 VSWR.

Theoretical and practical consideration in the optimization of SAW resonator design parameters for short term stability and phase noise are discussed. A fabrication process resulting in improved long term stability and turning point control is presented. Fabrication tolerance variations in resonance frequency are dealt with by a plasma process trim and by design optimization of static to motional capacity ratio or pulling range of the resonator. The introduction of semiconductor doping techniques reduces metal migration and attendant instabilities due to drive level.

Factors affecting the design of precision high frequency SAW oscillators are discussed with particular emphasis on the choice of transistor type and oscillator configuration. FET and bipolar transistors are compared to obtain optimum power leveling, desirable input and output impedance characteristics and lowest possible noise factor for improved short term stability. Pierce, Colpitts and Clapp oscillator configurations are examined in this study. Practical considerations of stray capacitance and biasing requirements which can shunt the various elements of these configurations are reviewed in making a final design choice. Due to size constraints a component type ovenizing system is chosen for stabilizing this oscillator.

Experimental results of prototype units are reported for phase noise, Allen Variance of fractional frequency deviation, power consumption and temperature stability. Noise floor, close-in phase noise and Allen Variance for 1 sec. to 100 sec. averaging times are presented. Testing considerations, including reference oscillator requirements and EMI/RFI effects, are also presented.

INTRODUCTION

This paper will describe the development of a prototype SAW resonator controlled 740 MHz precision local oscillator for satellite applications. The oscillator is housed in a 2.4 cubic inch package. It uses discreet components and a miniature component oven for temperature compensation and enhancement of short term stability.

Considerations in the design and fabrication of resonators for this application will be discussed as well as those of the oscillator.

Table #2, discussed in a later section of this paper, lists the goals of this program. Since this is for a satellite application, these goals are required to be met in a thermal vacuum environment. Those considered most challenging were the setability of ± 3.4 ppm, a long term stability or "aging" of ± 5 ppm for the first year, and a short term fractional frequency deviation of 1×10^{-10} which was especially critical at the longer averaging times.

In order to achieve these goals it was necessary to optimize both the resonator design and oscillation configuration. The design of the resonator will be considered first.

SAW RESONATOR DESIGN

Noise Considerations

In a feedback oscillator AM noise is limited because the active device of such an oscillator is operated in a saturated condition. The dominant noise form is then FM or phase noise. The commonly accepted way of analyzing phase noise is based upon an expression by Leeson^{1,2,9} which is shown in Figure 1. Though this is not an exact expression it is a useful vehicle for optimizing the noise performance of such a feedback oscillator.

Using this expression we see that there are several steps that can be taken to minimize the noise in a feedback oscillator. In order to establish the best noise floor the terms in the numerator of the coefficient should be minimized. The first factor to be minimized is the multiplication factor N. To achieve a minimum value of unity requires that the resonant device be designed to operate at at

desired output frequency instead of at a subharmonic.

Secondly, the insertion loss of the resonator device should also be minimized since the gain G required to meet oscillation conditions is very nearly proportional to this loss. Finally the noise figure F of the amplifier should also be minimized. Little can be done to minimize thermal noise KT . The frequency dependent terms which govern the region closer to the carrier are characterized by the phase slope of the resonant device which should, therefore, be maximized for optimum noise performance.

Since the technology presently exists to build SAW resonators to over 1 GHz using conventional photolithographic processing, fundamental frequency operation was adopted at the conception of this program. The requirements for highest Q and lowest insertion loss dictated the use of a SAW resonator rather than delay line, and specifically, a 740 MHz SAW resonator with the lowest possible loss and highest possible phase slope.

Equivalent Circuit Parameters

Figure 2 is a summary of equivalent circuit relationships which are used for design of a single port SAW resonator. These simplified equations are valid at frequencies near resonance and have been derived from earlier work by this author and others.³

$$\left(\frac{P_{ss}}{P_c}\right)_{\frac{48\text{dB}}{\text{Hz}}} = 10 \text{ LOG} \left[\frac{N^2 G F K T}{P_o} \left(\frac{\omega_o}{(A_m)^2} + \frac{1}{(A_m)^2} + 1 \right) \right]$$

P_{ss} = SINGLE SIDE BAND NOISE POWER

P_c = OSCILLATOR POWER

N = MULTIPLICATION FACTOR

G = AMPLIFIER GAIN \approx I.L. OF RESONATOR

F = NOISE FIGURE OF AMPLIFIER

KT = THERMAL NOISE

ω_o = FLICKER FREQUENCY = $\frac{F_o}{2\pi}$

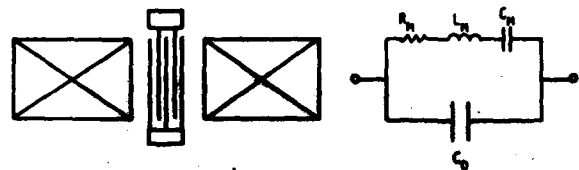
$\Delta\omega$ = OFFSET FREQUENCY

τ = RESONATOR PHASE SLOPE OR GROUP DELAY

Single Sideband Phase Noise Spectrum

Figure 1

By examining the equivalent circuit in this Figure one can see that the motional resistance, R_m is the main source of insertion loss. Furthermore this motional resistance is directly proportional to the radiation resistance, R_o of the interdigital transducer pumping the cavity of the resonator. Most of the factors determining the radiation resistance are material constants, however, one design variable N_e , the number of effective finger pair has a dominant effect because it is a squared term. This factor is the main reason for not choosing a two-port resonator configuration. The maximum SAW resonator cavity length is determined by interference of other longitudinal modes.



M = EFFECTIVE CAVITY LENGTH
 N_e = EFFECTIVE NO. OF FINGER PAIR
 A = APERTURE OF TRANSDUCER

Γ = TOT. REFLECTION COEFFICIENT

$$R_o = (8k^2 F_o A_c N_e^2)^{-1}$$

$$C_o = N_e A_c$$

MOTIONAL ELEMENTS

$$R_m = R_o \frac{1 - \Gamma}{1 + \Gamma}$$

$$L_m = \frac{M R_o \tau}{F_o (1 + \Gamma)^2}$$

$$C_m = \frac{1}{(2\pi F_o)^2 L_m}$$

OTHER CHARACTERISTICS

$$Q_u = \frac{2\pi M \tau}{1 - \Gamma^2}$$

$$\frac{C_o}{L_m} = \frac{\pi^2 \tau^2 M}{2 K^2 (1 + \Gamma)^2 N_e}$$

Equivalent Circuit
 Single Port Resonator

Figure 2

Therefore, for a given cavity length, the single transducer of a one port resonator would have twice the number of finger pair as a two-port; thus, yielding one fourth the motional resistance.

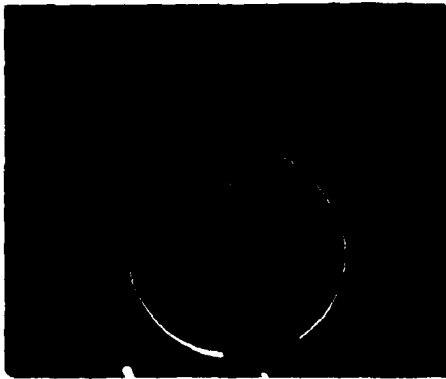
Examining these relationships also demonstrates that the maximum unloaded Q is achieved by maximizing cavity size M and the total reflection coefficient, Γ . The static to motional capacity ratio, C_o/C_m is also an important figure of merit because it establishes the pullability of a resonator in a given oscillator circuit. It is desired to minimize this ratio for the largest tuning range. This is also accomplished by using a single port resonator since N_e would be maximized in this type of design.

Experimental Results

Figure 3 shows the network analyser S_{11} trace in polar form of the single port SAW resonator designed for this application.

This 740 MHz resonator is of all metal construction, and both transducers and reflectors are $\lambda/4$ geometries. The aperture is 100λ , the cavity size is 100.75λ and the transducer is 98λ in length and has a transverse cosine weighting. The metal is 600A° thick and it has several percent copper dopant for increased power handling.¹⁰

We can see from this figure that this resonator has an inductive swing, a motional resistance of approximately 20 ohms, and that there are no competing modes as is demonstrated by the smoothness of the trace. Experimental measurements also show that the unloaded Q is 16,800.



Reflection Coefficient, S_{11}
for a Single Port Resonator

Figure 3

Short Term Stability

In order to achieve the short term stability goals of the program a miniature component oven was chosen to stabilize the temperature drift of the SAW resonator. Maximum stability could only be achieved by making the control point of this oven coincide with turnover point of the SAW resonator. Since the control points of this oven could only be adjusted $\pm 5^\circ\text{C}$, the turnover point variations due to fabrication influences must also remain within these limits. Figure 4 shows the temperature vs frequency curve for one of these prototype SAW resonators.

In this figure the solid curve represents a least squares fit to the data points shown as crosses. The goodness of the fit of this curve together with the low second order coefficient indicate that mounting stresses were well under control. The 63°C turning point temperature of this unit was nearly ideal for this application and the observed standard deviation was $\pm 3^\circ\text{C}$.

Long Term Stability

Long term stability is mainly a matter of good process definition which eliminates systematic causes and good process controls for minimization of random causes. Figure 5 illustrates the process developed to achieve good aging and tight make tolerances.

Section (a) Figure 5 illustrates the process flow and "Rework" paths of the wafer fabrication area. This area could be characterized as conventional hard contact printing process with wet chemical etch. Careful wafer cleaning prior to metalization is key to removing contamination and achieving good metal adhesion. Control of linewidth and metal thickness is required to set the center frequency and the turnover point within acceptable limits.

The distinguishing feature of the assembly process, Figure 5(b), is the inclusion of two types of resonator packaging concepts. The first is a conventional coined, Kovar TO-5 package and the second a Kovar HC-18 type bulk crystal package with modifications for clip mounting directly to the SAW resonator die. Electrical connection to the HC-18 package was provided by conductive epoxy. At the time of this writing the HC-18 packaged units have not been completely evaluated. Work is continuing to determine if improvements in aging, temperature vs frequency distortion, and hysteresis can be achieved with this package.

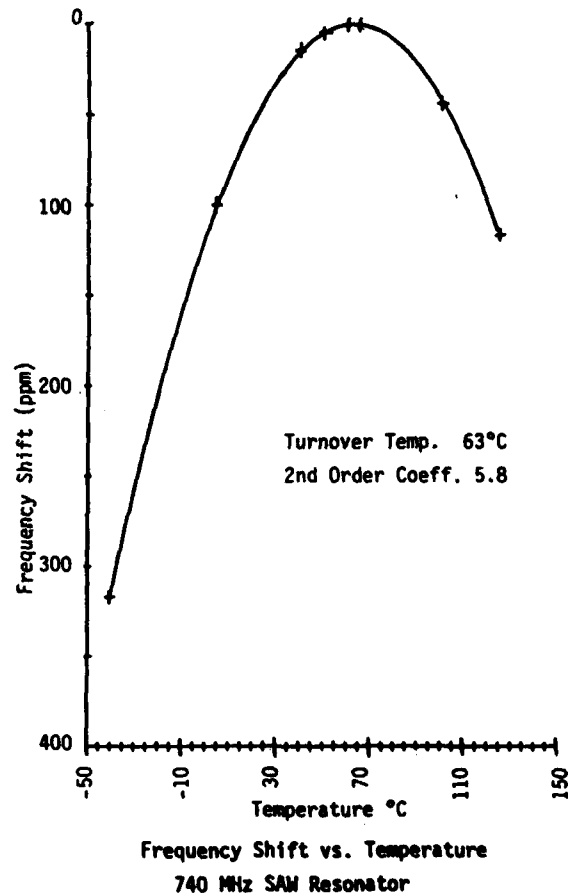
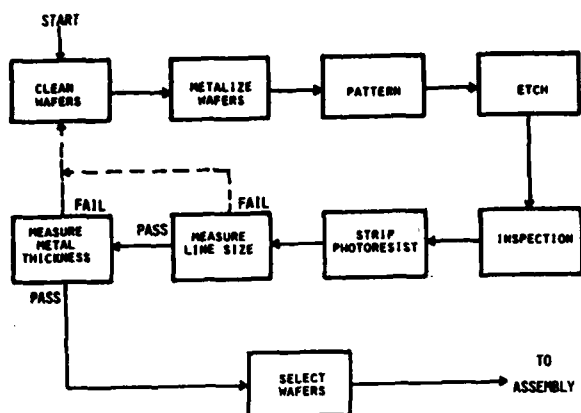
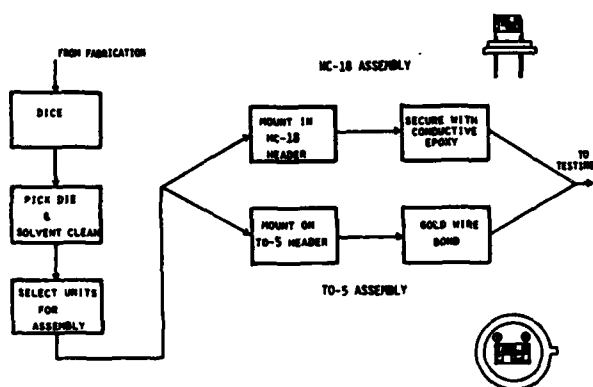


Figure 4

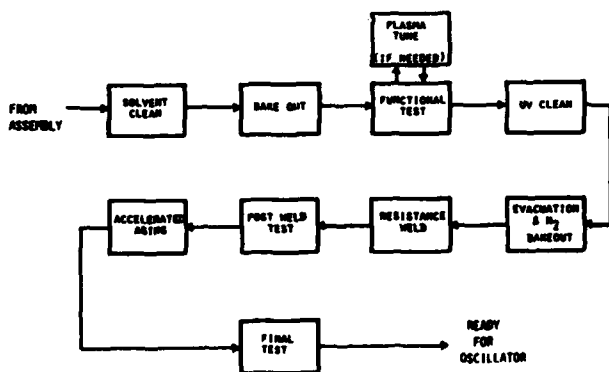
Figure 5(c) illustrates the process flow in the final assembly area. This includes a functional electrical test and plasma etching of units where necessary to adjust center frequency within a ± 50 ppm window. After functional testing units were given an ultraviolet/ozone cleaning to remove the final traces of contamination, the enclosures are evacuated and baked at 150°C in a nitrogen atmosphere. The resonators units were then immediately resistance welded in a nitrogen atmosphere and leak checked with a helium mass spectrometer. These sealed resonators were then fully electrically tested and "pre-aged" at elevated temperature for several days. The aging data was used to select well behaved units for final assembly into oscillators.



(a) FABRICATION



(b) ASSEMBLY



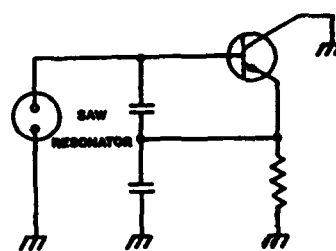
(c) SEAL AND TEST

SAW Resonator Process Flow
Figure 5

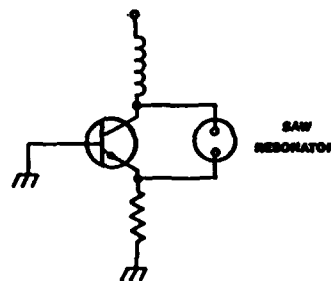
OSCILLATOR CONFIGURATION

Three types of oscillator configurations were considered: the Colpitts, the Clapp, and the Pierce. A simplified version of each is given in Figure 6.

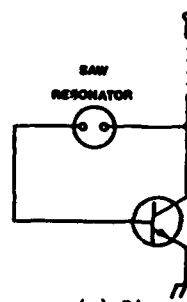
Several oscillators were built at approximately 200 MHz using the Colpitts configuration in Figure 6(a). These units



(a) Colpitts



(b) Clapp



(c) Pierce

Oscillator Configurations

Figure 6

worked well at this frequency but difficulty was encountered when this configuration was tried at 740 MHz. The capacitance values of the series connected capacitors in the feedback path are on the order of 5 to 15 pf at 200 MHz. At 740 MHz these capacitor values become much smaller and are approximately of the same order of magnitude as the stray circuit capacitance. This stray capacitance tends to swamp the discreet units, making value selection difficult and time consuming. For this reason the Colpitts configuration was set aside.

The Clapp Oscillator in Figure 6(b) was considered and several units tested. This circuit has a potentially better noise performance due to its grounded base configuration, but the units tested were all plagued with spurious oscillations to a degree far greater than any other configuration studied. Free running oscillations and squegging (low frequency oscillations associated with RC time constants of the bias and coupling circuits) were present in every case. Due to the difficulty involved in controlling these spurious oscillations this circuit too was set aside.

The Pierce Oscillator in Figure 6(c) had none of the drawbacks mentioned above and its performance was judged satisfactory. Based on the measured data gathered during the FET vs. Bipolar tests discussed later and work in this area by others^{5,6}, the Pierce Oscillator was chosen for further development.

CHOICE OF ACTIVE DEVICE - FET vs. BIPOLAR

Presently, a number of FET oscillators utilizing SAW Resonators are in use as injection sources in the cable television industry.⁷ Three circuits of this type were examined. All of these circuits were Pierce Oscillators that utilized 2-port SAW resonators in the frequency range of 650 to 750 MHz. These resonators had nominal insertion phases of 0° or 180° . In addition, a Pierce Oscillator using a two port 0° resonator and a bipolar transistor was also tested. The results for the four oscillators are presented in Table #1.

The published noise figure of the T1S-189 was approximately 1.5 dB greater than the bipolar unit but the most marked difference between device types was the bipolar circuit's ability to withstand externally induced changes such as tuning and voltage variations. For a supply voltage variation of $\pm 1\%$ the change in frequency for each oscillator was observed. In terms of parts per million per volt the bipolar circuit was clearly superior to all three of the FET circuits.

The tuning range of each oscillator was determined by compressing or expanding the turns of wire in the inductors that are part of the feedback loop and noting the frequency change at the extremes.

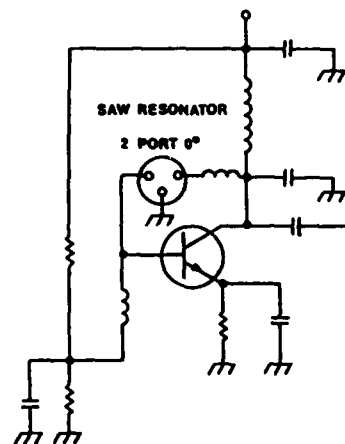
The tuning range of the FET circuits was much greater than that of the bipolar. Though a large tuning range is an advantage in cost sensitive commercial applications where large resonator frequency distributions must be tolerated, it is a disadvantage in a precision oscillator application. Large tuning range indicates a high sensitivity to the aging effects of the other feedback components. The R.F. output power of the FET circuits also varied considerably more than that of the bipolar unit when the oscillators were tuned.

The bipolar circuit also exhibited a harmonic level (at $2 f_0$) that was 7 dB lower than the best FET circuit. This was considered significant in light of a -40 dBc program goal.

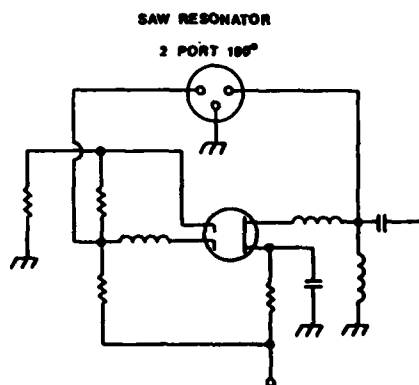
The FET circuits had the advantage of overall higher R.F. output power than the bipolar types, but none of the oscillators studied had sufficient R.F. output to eliminate the need for a second (amplifier) stage. The FET circuits also used less power than the bipolar but this was attributed to an excessively stiff bias supply on the bipolar circuit. This was later modified and the bipolar's power consumption was reduced to approximately 175 mW.

At this point FET circuit #1, was chosen as representative of the FET circuits and used for further comparison to the bipolar circuit. The schematic diagrams of the two are presented in Figure 7. Using a Hewlett-Packard 5390A Frequency Stability Analyzer and manufacturer supplied software, plots were made of fractional frequency deviation and phase noise for each of the two oscillators.

From the phase noise plot (Figure 8) it can be seen that the bipolar's performance is clearly superior to the FET, being 10 to 15 dB lower in noise for offsets of 10 Hz to 1000 Hz. The peak on the FET's curve is a "bright line" at 60 Hz that can be attributed to power-line interference.



Bipolar



FET

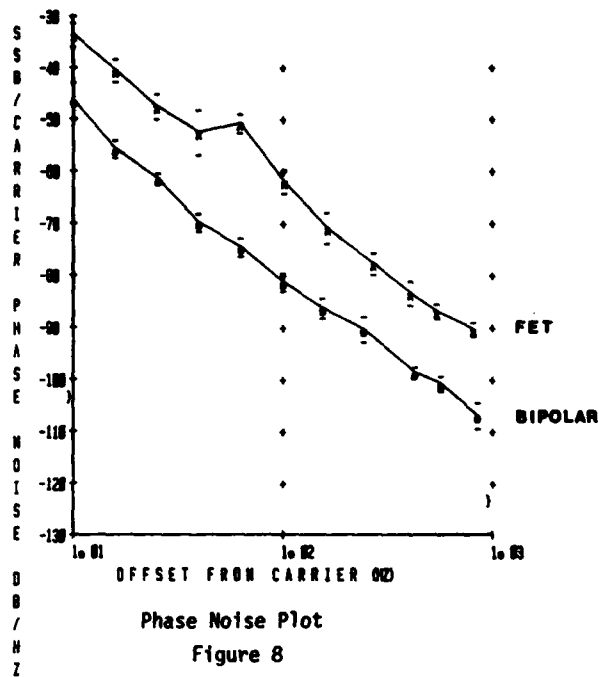
Oscillator Circuits for Bipolar and FET

Figure 7

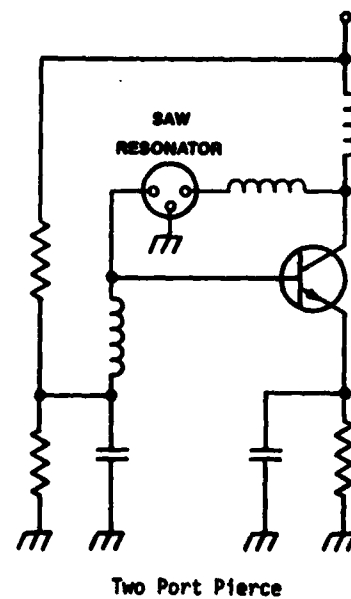
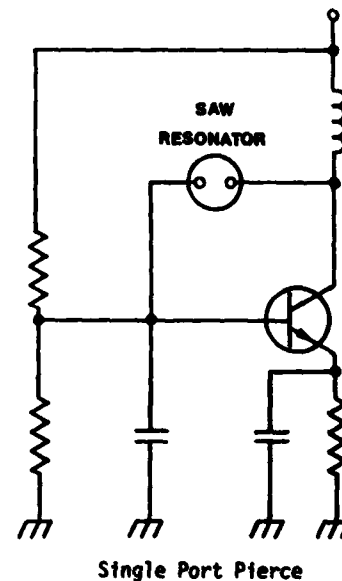
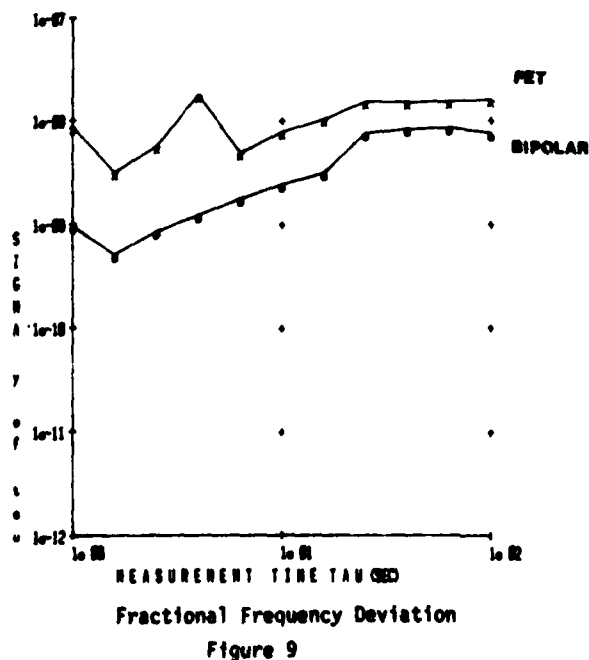
For the reasons just discussed, a bipolar device was selected as the most suitable for use in a precision oscillator.

RESONATOR TYPE: SINGLE PORT vs. TWO PORT

A single port resonator was chosen for this work primarily due to its lower insertion loss, as mentioned in the earlier discussion on SAW resonator design. There are also other advantages to the single port resonator.



A similar picture is obtained by comparing the short term stability of these oscillators. Figure 9 is a plot of sigma Y of tau (the standard deviation of fractional frequency drift or Allen Variance) versus measurement averaging time. The bipolar circuit frequency drift is approximately an order of magnitude smaller at the shorter averaging times of 1 to 10 seconds. Toward the longer averaging times from 10 to 100 seconds, the difference is smaller, though at these averaging times the temperature dependency of the resonator and long term stability may be masking these measurements. At this point, neither oscillator had been temperature stabilized with an oven.

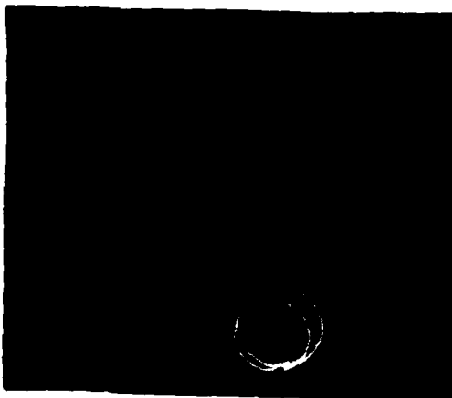


Single Port and Two Port Oscillators
Figure 10

Figure 10 shows the schematic diagrams for oscillators using each type of resonator. Note that the two port oscillator circuit requires two inductors in the feedback loop. These are necessary in order to achieve the required phase shift and reduce the insertion loss of the device. These inductors tune out the negative reactance appearing at each port of the two port resonator. (Figure 11)

Referring back to Figure 3 it can be seen that the single port resonator swings inductive. This eliminates the need for bulky, often temperature sensitive inductors. The single port circuit requires only one additional capacitor in the feedback loop to allow for tuning. Its equivalent circuit (Figure 2) also resembles that of the well known bulk mode crystal resonator. This allows for easier implementation into circuits previously developed for that type of resonator.

In summary, the single resonator has the advantages of lower insertion loss, requires fewer external components, and is easier to implement.



Reflection Coefficients S_{11} and S_{22}
for a Two Port Resonator

Figure 11

OSCILLATOR DESIGN

The resonator's effect on short term stability has been previously discussed. Here, consideration is given to the design of the oscillator for optimum short term stability.

As discussed earlier, the noise figure of the amplifying device should be as low as practicable. The transistor selected for this work has a noise figure of 1.8 dB at 1 GHz. A high ratio of circulating power to output power, i.e. a lightly loaded oscillator, minimizes the effect of load variations that tend to "pull" the oscillator frequency.⁹ The high circulating power necessary for this has been cited⁸ as causing deterioration of the resonator, but the copper doping of the

electrodes mentioned previously has been found to minimize the effect. One additional step taken here was to introduce a small amount of emitter degeneration in the oscillator circuit to reduce the effect of flicker noise generated in the base-emitter junction of the transistor.⁴ The final value of the emitter bypass capacitor is the minimum required to maintain oscillation.

TEMPERATURE COMPENSATION

The previous steps have their greatest effect on short term stability at the shorter averaging times. At the longer averaging times however, temperature begins to play a significant role.

There are two basic approaches to compensating an oscillator for medium (i.e. temperature) stability. One approach is that of a temperature compensated oscillator (TCXO) and the other involves the use of ovens. While the TCXO offers the advantage of low power consumption, it has definite disadvantages. The increased complexity of the TCXO decreases reliability and it also requires extensive development and testing.

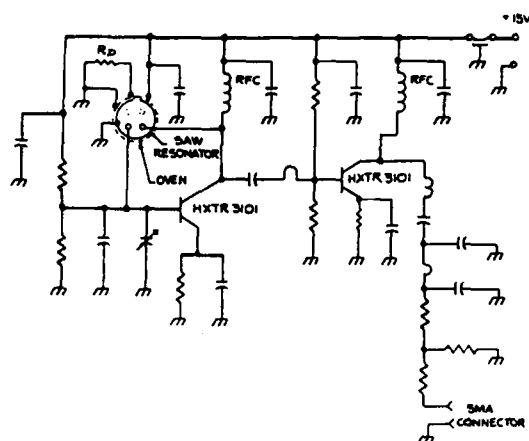
The ovenized approach avoids these disadvantages. The power required by a miniature component oven and its physical size were both within the limits of the program goals, so these were not felt to be a disadvantage for this project.

OSCILLATOR -FINAL CONFIGURATION

The oscillator is housed in a 1.2"W X 2.0"L X 1.0"H gold plated aluminum housing. The oscillator circuit is constructed on one side of a piece of RT Duroid 5870 using discrete components. The resonator and miniature component oven are mounted on the other side of the circuit board along with connections for +Vcc.

The final circuit configuration is given in Figure 12. Hewlett-Packard HXTR-3101 bipolar transistors were used as active devices in both the oscillator and amplifier stages. The resistor labeled R_p in the schematic diagram is the oven programming resistor. Its value is selected to adjust the temperature of the oven to the turnover point of the individual resonator. This enhances the short term stability of the oscillator by reducing its sensitivity to small temperature variations, since the oven maintains the resonator at a temperature where its frequency versus temperature characteristic is at or near a minimum slope.

The final frequency of the oscillator is adjusted via a stripline type variable capacitor at the base of the oscillator transistor. Coarse frequency adjustment is accomplished by resonator tuning and selection. The power output of the oscillator is approximately -3 dBm. The gain of the amplifier is approximately 15 dB. This



PRECISION L-BAND OSCILLATOR

Figure 12

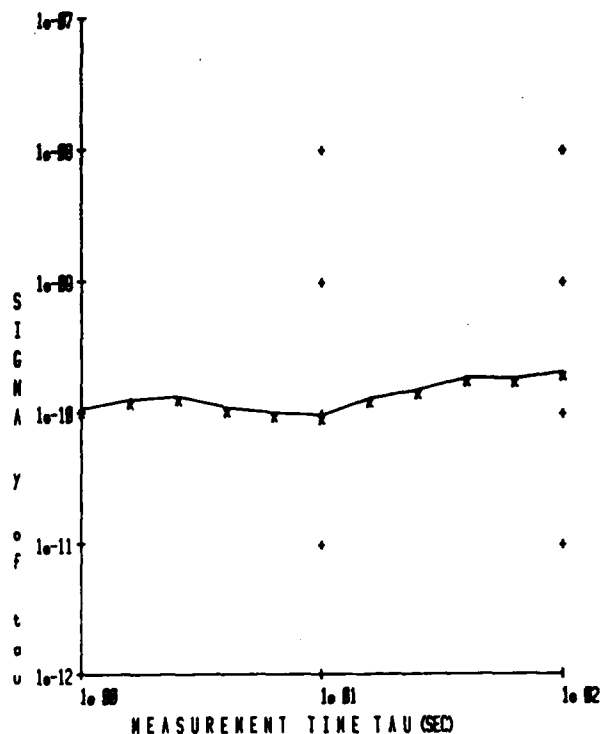
combination results in an output level that meets the +10 dBm goal with a comfortable margin. The output of the amplifier is fed through a PI network that functions as a low pass filter and through an attenuator that serves as a minimum load on the amplifier and holds the units output impedance within limits. Final output is via a removeable SMA connector which is stripline compatible.

PERFORMANCE

The final prototype configuration was then tested for short term stability, phase noise, long term stability, and performance over temperature. The short term stability, shown in Figure 13, for the specified averaging times is very close to the goals. At the 1 second averaging time the fractional frequency deviation was approximately 1.5×10^{-10} , at 10 seconds, 1.7×10^{-10} and at 100 seconds, 2.9×10^{-10} .

Figure 14 is a plot of the single sideband phase noise for 10 Hz to 1000 Hz offsets from the carrier. At the 10 Hz offset, the phase noise is approximately -52 dBC. From there it rolls off at a $1/f^3$ rate to approximately -110 dBC at the 1000 Hz offset. The advantage of a single port resonator over a two port resonator can be seen by comparing the phase noise plot of Figure 14 with that of the two port bipolar oscillator in Figure 8. At the 10 Hz offset it can be seen that the single port circuit has a 5 to 6 dB advantage over the two port circuit.

An estimate of long term stability (i.e. aging) was made by fitting a line to the frequency vs. time curve in Figure 15. Over the 24 hour period indicated in the graph, the extrapolated drift rate was +5.03 ppm/year, essentially the +5 ppm/year goal. A later measurement by the customer in a thermal vacuum environment over a three week period yielded a drift rate of +4.5 ppm/year.



Short Term Stability for
L-Band Precision Oscillator

Figure 13

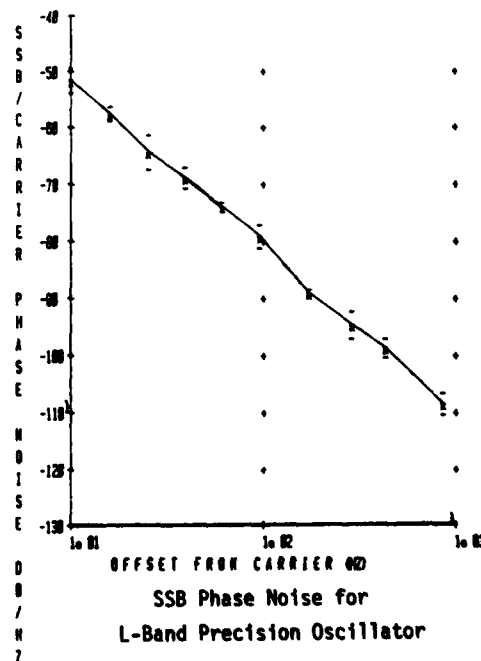


Figure 14

Over the temperature range of -15°C to $+60^{\circ}\text{C}$ the R.F. output power (Figure 16) meets the goal of $+10$ dBm minimum and exhibits less than 1 dB variation. The frequency drift (Figure 17) over the temperature range is approximately ± 1 ppm. The power consumption results (Figure 18) are very close to the original goal except at $+40^{\circ}\text{C}$ and higher. Here the oven consumes very little power and the power consumption is primarily that of the oscillator circuitry.

The final package size fell short of the goal by 0.36 cubic inches. It is believed that this goal could be reached by a careful redesign of the housing and circuit board layout.

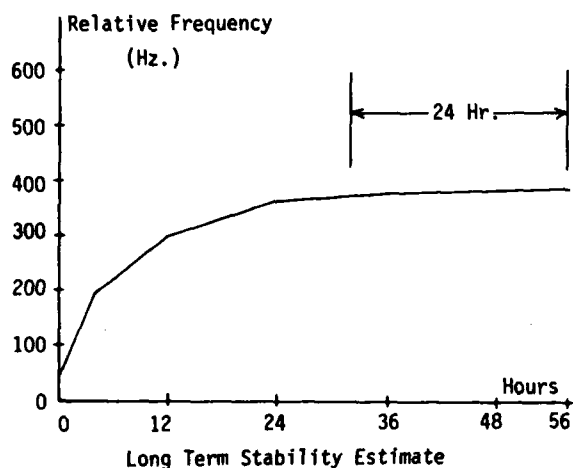


Figure 15

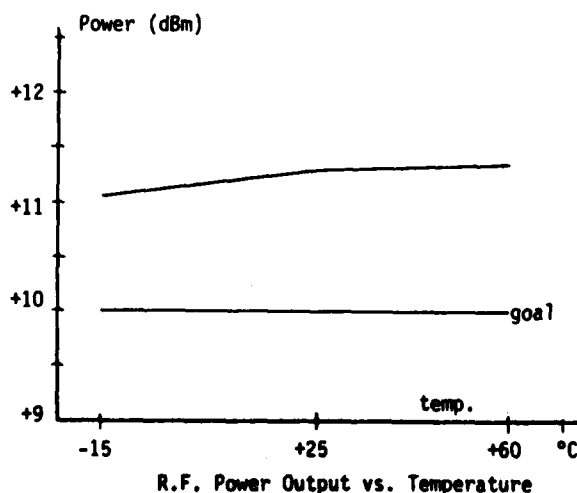


Figure 16

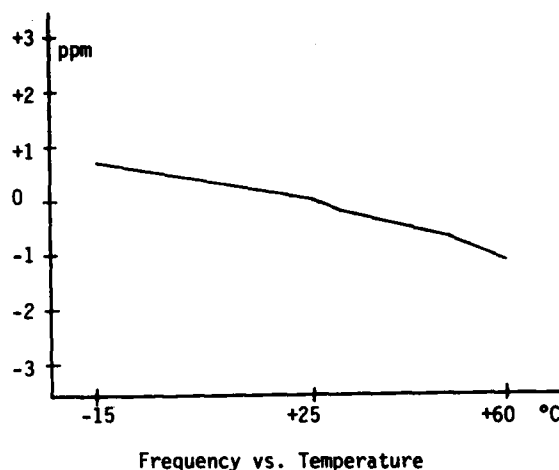


Figure 17

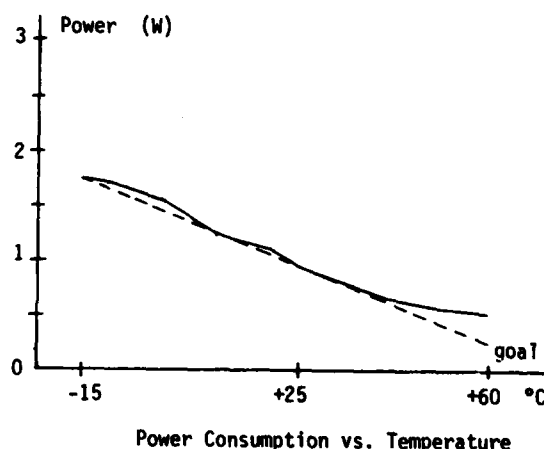


Figure 18

TESTING CONSIDERATIONS

The short term stability (Allan Variance) and phase noise ($\mathcal{L}(f)$) measurements presented here were made on the Hewlett-Packard 5390A Frequency Stability Analyzer. The system uses a 9825B Controller, manufacturer supplied software and an 8662A reference oscillator.

In order to obtain meaningful results in testing the oscillator, consideration must be given to the testing environment. Two areas that can mask test results are electromagnetic interference (EMI) and radio frequency interference (RFI). A third area of concern is the thermal environment. The steps taken to minimize these influences are outlined as follows.

The testing environment should be shielded or otherwise well removed from EMI sources because electrical noise of sufficient strength can easily modulate the oscillator output. Also, a fixed attenuator of 20 dB directly at the output of the oscillator will reduce the effects of noise entering the

PARAMETER	FET OSC. #1	FET OSC. #2	FET OSC. #3	BI-POLAR	UNITS
NOMINAL FREQUENCY	674	680	668	740	MHZ
CIRCUIT CONFIGURATION	PIERCE	PIERCE	PIERCE	PIERCE	-
RESONATOR TYPE	180° 2 PORT	0° 2 PORT	0° 2 PORT	0° 2 PORT	-
ACTIVE DEVICE	TIS-189	TIS-189	BF-960	HXTR-3101	-
DEVICE NOISE FIGURE	3 dB @ 900 MHZ	3 dB @ 900 MHZ	NOT AVAILABLE	1.8 dB @ 1 GHZ	
FREQUENCY SENSITIVITY TO ±1% SUPPLY VOLTAGE VARIATION	30	36	10	1	PPM/VOLT
TUNING RANGE	1600	1000	130	13	±PPM
POWER OUTPUT VARIATION WITH TUNING (P_0)	8	17.4	4.9	1.3	DB
POWER CONSUMPTION	132	120	286	306	MW
HARMONIC AT F_0	-17	-18	-20	-27	DBC
R.F. POWER OUTPUT	+3.3	-5	+3.4	-3.11	DBM

(UNMATCHED)

ACTIVE DEVICE COMPARISONS

TABLE 1

oscillator circuitry via the output connections. Although a shielded enclosure for reducing EMI/RFI effects was not available, most of the undesirable effects of EMI/RFI were overcome by the attenuator mentioned above and by using the shortest possible R.F. cable runs between the oscillator output and the input ports of the test equipment. In addition, the power supply leads were also shielded and decoupled at the entrance to the oscillator housing. These steps yielded a marked improvement in short term stability measurements.

The oscillator itself should also be shielded on all sides. An additional improvement was obtained when tests were conducted during the night time hours when sources of EMI were fewer and usually at lower levels than during the day. The thermal environment was also more stable at night.

During the temperature cycling more consistent results were obtained with measurements made in still air. This was also considered more consistent with a thermal vacuum environment because convection effects were minimized. Preliminary tests of the oscillator in a thermal vacuum show no significant change in performance.

CONCLUSIONS

A complete list of the goals attempted and the results achieved for this project is compiled in Table 2. The goals for frequency setting, R.F. power output, harmonic and spurious suppression were all met or exceeded. The short term stability for the specified averaging times is very close to the goals. Less than 5.0 ppm 1st year long term stability has also been achieved in a thermal vacuum environment. This is believed to be state of the art for SAW oscillators of this frequency, size and power consumption.

This work has successfully demonstrated the feasibility of L-Band Precision SAW Oscillators. The majority of the goals attempted were met and those remaining are certainly within the realm of possibility.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the following people for their help in preparing this paper: E.J. Staples of Rockwell International for his suggestions on the improvement of short term stability, R. C. Rosenfeld, for his technical advice and creative suggestions for the preparation of the manuscript, C. Register, J. Dodge, for Construction and Testing of the Oscillator Prototypes, M. Ellis for Drawings and C. Green for Typing of Manuscript.

<u>SPECIFICATIONS</u>	<u>GOAL</u>	<u>ACTUAL</u>
Frequency	740.000 MHz ± 3.4 PPM	-0.52 PPM
R.F. Power	+10 dBm Minimum	+11.60 dBm @ 25°C
Short Term Stability	Average Times (100 Samples)	
	1 Sec. 1 X 10^{-10}	1.473×10^{-10}
	10 Sec. 1 X 10^{-10}	1.682×10^{-10}
	100 Sec. 1 X 10^{-10}	2.896×10^{-10}
Long Term Stability	+5 PPM for 1st Year	1.38×10^{-8} Per Day
	± 12 PPM for 5 Years with Goal of ± 6 PPM for 5 Years	(5.03 PPM/Year) Customer verified at 4.5 PPM/Year
Harmonics	<-40 dBc	2nd <-60 dBc 3rd and higher not measured
Spurious	<-60 dBc	<-60 dBc (100 Hz to 1500 MHz)
Impedance	50 Ω Nominal	40.7 Ω
Operating Voltage	15V $\pm 1\%$	15V $\pm 1\%$
Frequency Variation with $\pm 1\%$ Change in Operating Voltage	None	<0.2 PPM
Power Consumption at Turn On	3.5W (Maximum)	3.47W
Over Temperature Range	1.75W Max. @ -15°C	1.74W
	0.25W Max. @ +60°C	0.50W
Mechanical	1.2"WX2.0"LX0.85"H (Max) 1.2"WX1.5"LX0.65"H (Goal)	1.2"WX2.0"LX1.0"H

GOALS AND ACHIEVEMENTS

TABLE 2

BIBLIOGRAPHY

1. D.B. Leeson, "Short Term Stable Microwave Sources," *Microwave Journal* (June 1970).
2. D.B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," *Proc. IEEE* **54**, 329 (February 1966).
3. J. S. Schoenwald, R.C. Rosenfeld and E.J. Staples, "Surface Wave Cavity and Resonator Characteristics - VHF to L-Band", *IEEE Ultrasonics Symposium Proceedings*, pp 253-256, November 1974.
4. D.J. Healy, III, "Flicker of Frequency and Phase and White Frequency and Phase Fluctuations in Frequency Sources", *Proceedings 26th Annual Symposium on Frequency Control*, pp 29-42, June 1972.
5. S.J. Dolochyckji, E.J. Staples, J. Wise, J.S. Schoenwald, and T.C. Lim, "Hybrid SAW Oscillator Fabrication and Packaging", *Proceedings 33rd Annual Symposium on Frequency Control*, pp 374-378, June 1979.
6. Marek E. Mierwinski and Mark E. Terrien, "280 MHz Production SAWR", *Hewlett Packard Journal*, Dec. 1981, Vol. 32, No. 12, pp 15-17.
7. Darrel L. Ash, "SAWS! Powerful Passives", *R.F. Design*, March/April 1982, Vol. 5, No. 2, pp 12-19.
8. B.Y. Lao, N.J. Schneier, D.A. Rowe, R.E. Dietterle/J.S. Schoenwald, E.J. Staples, J. Wise, "Commercial Satellite Navigation Using a SAW Oscillator", *Proceedings 35th Annual Frequency Control Symposium*, pp 345-348, May 1981.
9. T.E. Parker, "Current Developments in SAW Oscillator Stability", *Proceedings 31st Annual Frequency Control Symposium*, pp 359-364, June, 1977.
10. "Improved Metalization for Surface Acoustic Wave Devices", *Thin Solid Films*, **64** (1979) 9-15. J.I. Latham, W.R. Shreve, N.J. Tolar and P.B. Ghatge (1979).

AD P 002506

SAW RESONATOR STABILIZED
OSCILLATORCullen Bennett
RF Monolithics, Inc.
4441 Sigma Road
Dallas, Texas 75234Abstract

→ A systematic approach to the design of a feedback RF oscillator using a two-port SAW resonator as the primary frequency control element will be presented. The tradeoff between oscillator pull range and stability will be discussed. Oscillators have been developed which are optimized for minimum manufacturing cost in the UHF range. Over 100,000 oscillators have been successfully used for over one year with no known failures in the field. Data will be presented on the aging characteristics of oscillators using FET and bipolar transistors. Also, the interaction of oscillator pulling and temperature stability has been extensively investigated. ←

Introduction

The birth of Surface Acoustic Wave Resonators (SAWR) is recorded as being in 1975, with the literature recording intense amounts of continuing work being done in the area of aging and short term stability by various investigators. It is through these efforts that a utility can be made utilizing these devices to fulfill a realistic need for small, low cost VHF and UHF frequency sources. Frequency crowding has dictated tighter and tighter tolerances on frequency accuracy which is in direct opposition to low costs from a manufacturing point of view. It is with this in mind that an approach to the design of an oscillator circuit that can accommodate the wider spread in manufacturing frequency tolerances that coincide with lower costs and be able to "pull" the frequency back to the desired operating frequency is attempted. The overall frequency characteristics of the new circuit afford adjustability while still maintaining the desirable characteristics associated with SAWR oscillators.

Oscillator Design

The original credit for the feedback oscillator goes to Barkhausen with the criterion established in figure 1 as being those required for oscillations to occur. The circuit discussed here (figure 2) is an implementation of those criteria in such a manner as to utilize the SAWR as the primary frequency control element in the oscillator. The SAWR is normally considered as a device with a fixed Q, center frequency and some insertion loss. The fact that these parameters are normally given in terms of a 50 ohm system has apparently restricted the thought of using the devices in an impedance other than 50 ohms or whatever impedance the circuit will provide. The fact that the resonator exhibits a 90 degree transmission phase between the 3 dB bandwidth points for various values of loaded Q enables us to electrically establish a variable phase slope while using the SAWR in the feedback loop of an oscillator. This phase slope does have finite limits in that the device does not have infinite Q in itself and thereby sets the maximum phase slope attainable. The minimum phase slope is limited only by the characteristics of the resonator under loaded conditions and the other components in the

oscillator loop. The transmission amplitude and phase characteristics of a 674 MHz resonator are shown in figure 3 for two different impedance environments. These photos attempt to illustrate the difference in phase slope that can be realized by providing a controlled impedance environment in which the resonator can operate. It is also worth noting at this time that there is a limit to which the de-Q-ing can be accomplished without altering the desirable characteristics, the actual limit is determined from trade-offs. The 200 ohm impedance environment already shows asymmetry in the amplitude response versus frequency.

The impedance matching networks between the SAWR and the active device allow almost complete freedom to choose the operating impedance for the resonator and also allow for compensation of non ideal phase contribution by the active device. In most oscillator designs the final frequency adjustments are made by varying the values of these components. These matching networks can be either series type of matching networks such as pi or L networks or they can be a shunt type of network such as a tapped inductor or tapped capacitors. There are four possible combinations that can be utilized in an oscillator with two matching networks. The frequency vs. temperature characteristics for these four types of circuits, in oscillators, will be shown to be different for each type of implementation.

A typical example of an oscillator is one that uses a series type of matching on each side of the SAWR to establish the characteristic impedance in which the resonator will operate. This circuit also uses a dual gate MOSFET device as the active element. This type of device was chosen because of its relatively high gain, low reverse transmission characteristics as well as providing low noise amplification. In this particular example a 3N225 was used but the BF960 and the 3SK88 will work as well for providing the required gain at 674 MHz.

The goal of this design is to be able to accept a SAWR with a manufacturing tolerance of +/- 150 KHz on the center frequency and be able to pull the oscillator back to the center frequency. A SAWR at 674 MHz will typically have an unloaded Q in the range of from 10,000 to 11,000 which implies an unloaded 3 dB bandwidth of approximately 65 KHz and exhibit approximately 8 dB of insertion loss in a 50 ohm system. In order to be able to pull this device by +/- 150 KHz the loaded Q will have to be on the order of 2,200. In order to determine the impedance of the external circuit that will be required to provide this value of loaded Q the lumped element equivalent circuit should be examined. This lumped element model as shown in figure 4 is a first order approximation to the SAWR close to resonance. The external resistance that must be supplied in order to provide this desired value of loaded Q (Q_l) can be determined from the relationship

$$R_g + R_l = (Q_u / Q_l) * R_m - R_m$$

For equal generator and load resistances this value is determined to be approximately 300 ohms.

This new impedance environment will also reduce the insertion loss to approximately 3.5 dB as opposed to the 8 dB encountered in the 50 ohm environment. This new lower loss thereby reduces the requirements that are made on the amplifier section of the oscillator. If a small signal analysis is made on the circuit (no limiting due to saturation) a gain of about 6.5 dB is required of the circuit to allow for sufficient gain when the circuit is pulled to the 3 dB point on the resonator transmission characteristic. At this point we know the impedance that we would like to operate in and the next task is to design an amplifier and matching networks to provide the required minimum gain as well as provide the required transmission phase shift so that the loop will oscillate when closed upon itself. If the 3N225 FET is conjugately and simultaneously matched at the center frequency of the SAWR it is found to exhibit a stability factor of 6.6 in the 300 Ohm environment.

This stability factor insures that the device is unconditionally stable and that the device will not oscillate due to the impedance environment. A value of less than one indicates that the amplifier is only conditionally stable and that for some impedances oscillations may occur just due to the loading. S-parameters for the 3N225 device indicate that the matching networks required to simultaneously and conjugately match this structure in a 300 Ohm environment are .94 at an angle of 142 degrees ($10.6 + j104$) on the input and .97 at an angle of 124 degrees ($4.9 + j159$) on the output. These impedances will provide the required operating environment for the resonator to be "pulled" to the desired operating frequency within the tolerance window. Two pi section networks were selected for these matching networks. By selecting three element networks both the phase and the transformation ratio can be specified in the same network. A second advantage is that some of the parasitic capacitances associated with the devices can be absorbed by the matching network capacitances. An analysis of the amplifier and the matching networks indicates an S21 phase of 160 degrees when perfectly matched. This value can easily be adjusted to a value of 180 degrees while only providing a small amount of mismatch in the desired impedance. This 180 degrees and the additional 180 degrees of phase shift encountered in the SAWR at center frequency provide the required criteria of Barkhausen that the loop phase be a multiple of $2n \times 180$ degrees ($n=1,2,3,\dots$). The actual amplifier is capable of providing approximately 16 dB of gain in the 300 Ohm environment and thereby has sufficient gain to easily absorb the slight mismatches.

There have been numerous circuit implementations used to accomplish the same end result, each using a different configuration of components. In some instances, if the output impedance of the active device is low enough all of the additional impedance required can be supplied in the input matching network. This does require that the phase of the active device be of such degree as to still meet the loop requirements. Oscillators using these techniques have been constructed and operated to as high as 1 Ghs. In addition, over 100,000 oscillators using SAWR devices as the stabilizing element have been put into field tests at this time and no known failures have been reported. There have been a limited number of oscillators designed using bipolar transistors in the range of 600 MHz with good results from an inexpensive transistor. The 600 MHz is about the useful limit for the very inexpensive transistor while more costly devices will work equally as well at the higher frequencies.

Extensive work has been done in the evaluation of these circuits as functions of temperature and the amount of frequency that they are pulled from the actual device center frequency. With two possible types of tuning networks (series and parallel) and two places to utilize the circuits there are four possible configurations to consider. The four configurations will be referred to as those shown in figure 5 as types one thru four in the text. All four circuits gave similar frequency vs. temperature performance as far as the amount of frequency shift that was observed. There were some distinct differences however observed when a close analysis was performed on the shift observed in the turnover temperature of the oscillators as compared with that of the resonator used in that oscillator (when the resonator was observed in a 50 Ohm environment), versus the amount of frequency that the oscillator was pulled from the natural fabrication frequency of the resonator. Figure 6 illustrates a typical set of frequency versus temperature characteristics for a type 1 oscillator. The triangle points indicate the turnover point of the oscillator as computed from a quadratic fit to the limited number of data points. The four plots are made from separate temperature runs with the oscillators being set to four discrete frequency settings relative to the resonator that was used in that particular oscillator circuit. These settings were +70 KHz, 0 KHz, -70 KHz, and -130 KHz from the natural frequency of the resonator which was approximately 674 MHz. Figure 7 illustrates the effects of the various types of circuits on the turnover temperature of the resonator. The mechanism for this shift is not understood at this time but is being presented in raw form as documentation that there is a mechanism to be considered in the design of oscillator circuits.

Two types of oscillators, figures 8 and 9, have been fabricated and placed in temperature chambers in an attempt to acquire some insight into the aging characteristics of these circuits. One type of oscillator, figure 8, utilizing a bipolar transistor has been run at 35 deg. C and at 65 deg. C for approximately 200 days. There have been numerous momentary interruptions in the test as can be observed in the data. In figure 10 on the 65 deg. devices an interesting phenomenon was observed. The temperature in the chamber had drifted slightly (3 deg.) over a long period and was reset to 65 deg. at approximately 50 days, the resulting frequency shift was in accordance with that which would be expected (2.2 ppm/deg. C). The interesting point is that with the temperature maintained exactly at 65 deg. from that time there is a time constant of approximately 120 days with all three oscillators that appears to be returning to the same aging rate as observed before the temperature was reset. One of the primary problems with both this data and that taken on the oscillators fabricated with fet devices is that the turnover temperature of the oscillators was not the same as that of the test thereby producing an error term in the data that is very sensitive to the ambient conditions of the chamber. In the case of the bipolar circuit the oscillators turned over at approximately 0 to +10 degrees C. With the quadratic nature of the temperature characteristics and with the oscillators operating at 65 deg. C they have a frequency to temperature sensitivity of approximately 1.5 KHz per deg C at this point. The FET oscillators were under test for a year, with two units completing the period. The third unit was taken off of test and dissected to see if any visual degradation could be observed in

the resonator due to the elevated temperatures and power being dissipated in the device. None of these effects were observed. Figure 11 shows the data for the three FET oscillators of Figure 9. One oscillator has a jump in the frequency which is due to the circuit physically falling over in the test chamber and taking a fixed offset at that point. None of the data has been adjusted for perturbations incurred during the tests. There were no special precautions taken during the processing of these devices to insure ideal aging conditions in the packages.

Conclusions

SAWR oscillators can be designed to allow for frequency adjustability that can compensate for a wide frequency tolerance in center frequency. This allows for low cost, high volume production of these devices for a commercial market where ultra high stability is required. It has been observed that the type of circuit designed for these devices can have a significant bearing on the overall frequency versus temperature characteristics of the circuit. The temperature characteristics of any one circuit is normally very well behaved and can provide adequate stability for most commercial applications such as set top converters for television receivers. The aging rates observed in the various circuits appear to be different in characteristic shape, but due to the uncertainties involved in the tests with temperature a conclusive statement is not possible.

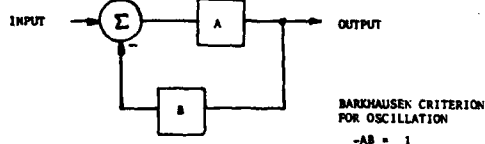


FIGURE 1

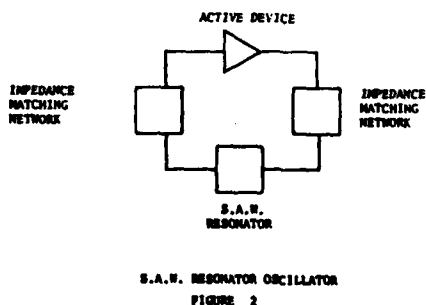
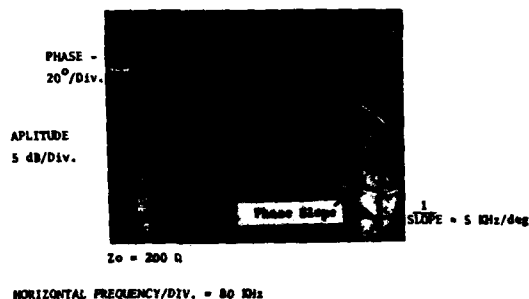
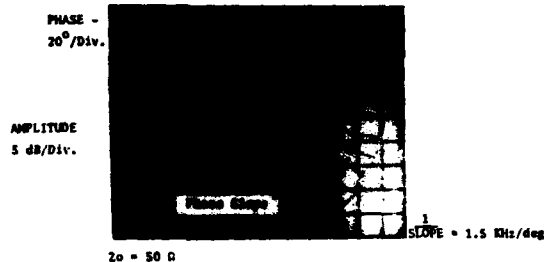


FIGURE 2



TRANSFER CHARACTERISTICS OF S.A.W. RESONATOR VS. IMPEDANCE ENVIRONMENT

FIGURE 3

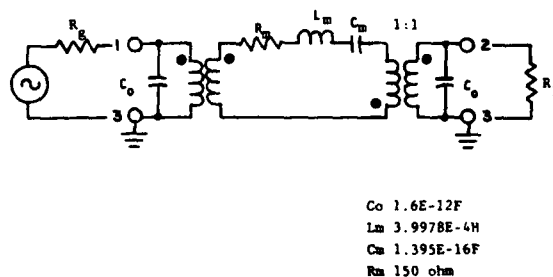


Figure 4 180 Degree SAW Resonator

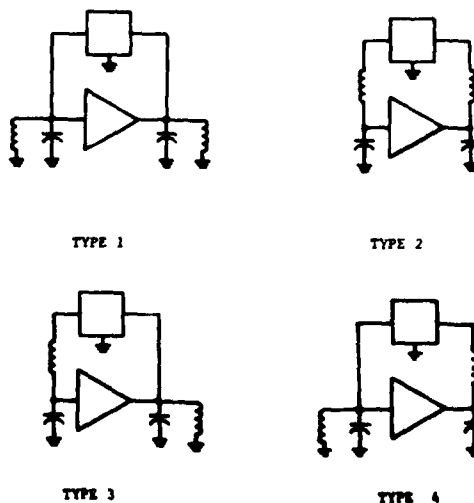


FIGURE 5

CHARACTERISTICS OF A SAW OSCILLATOR AT VARIOUS FREQUENCY SET POINTS

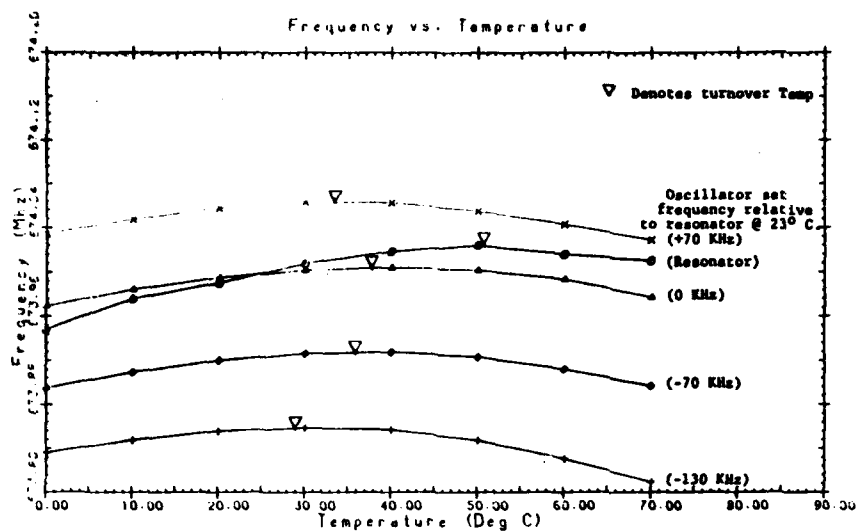


Figure 6

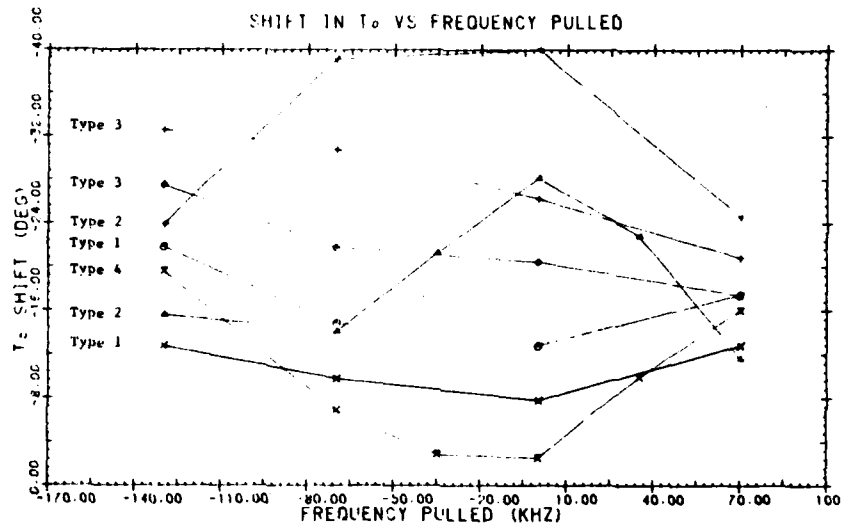
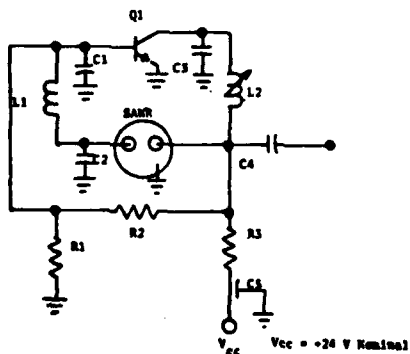


Figure 7

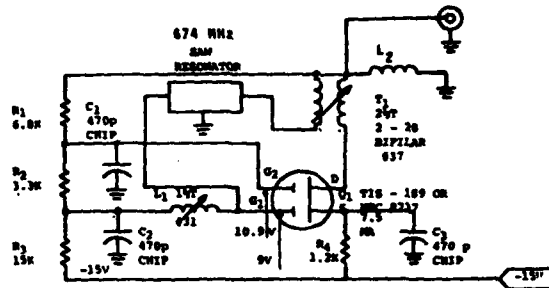


Q1 NPN-17
SAWR RP-1036
C1 5.6 pF
C2 1.0 pF
C3 1.5 pF
C4 0.02 pF
C5 500 pF Leadless

L1 - 4 turns #34 .065" ID.
modified (Brown)
L2 - 3 turns #34 .065" ID.
(Brown)
R1 2.7 K
R2 33 K
R3 1 K

SAW Resonator Oscillator

Figure 8



SAW RESONATOR OSCILLATOR

Figure 9

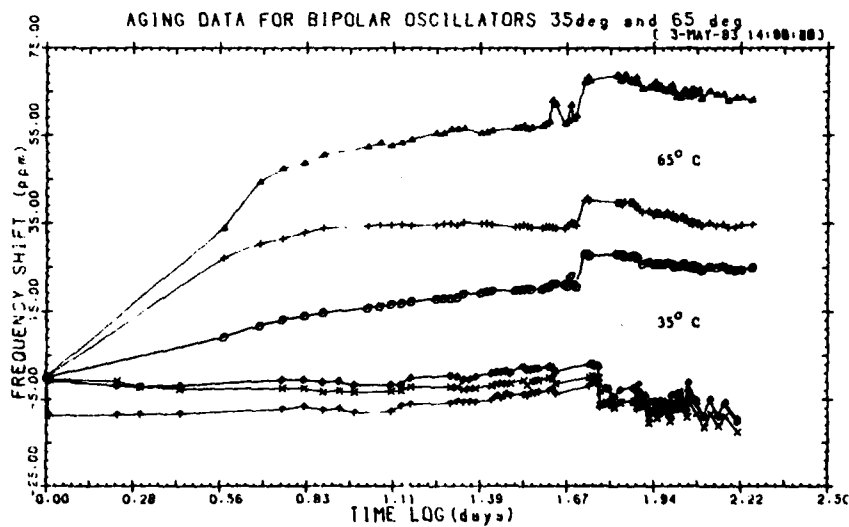


Figure 10

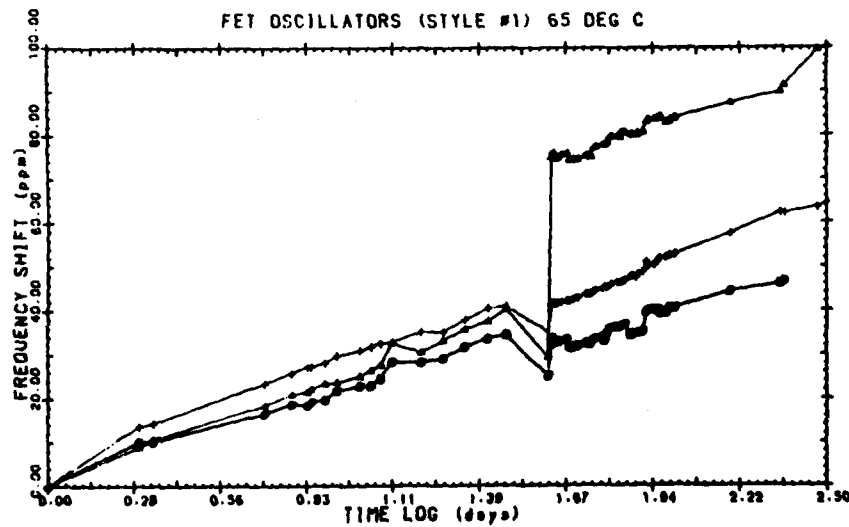


Figure 11

VERY LONG PERIOD RANDOM FREQUENCY FLUCTUATIONS IN SAW OSCILLATORS

T. E. Parker
Raytheon Research Division
131 Spring Street
Lexington, MA 02173

Summary

The long-term frequency stability of SAW oscillators is of considerable importance to many applications. Therefore, extensive work has been done to understand and reduce the observed changes in frequency of SAW oscillators which occur over periods of time ranging from days to years. With improvements in the fabrication and packaging of SAW devices, the systematic long-term frequency drift has been reduced to a level where 19 out of 23 oscillators have drifted less than 2 ppm after 1 year. However, with the reduced systematic drift, it has become apparent that there is a significant level of random frequency fluctuations occurring with periods ranging up to months and years. A typical delay line oscillator may show a systematic drift of 1 ppm in a year and an rms deviation of 0.3 ppm. The value of the systematic drift might lead one to expect a week-to-week stability of 2×10^{-8} , but the random fluctuations make week-to-week changes of 1 to 2×10^{-7} a common occurrence.

To determine the nature of the observed long period random fluctuation, a spectral analysis of aging data from 45 oscillators (mostly delay lines) has been carried out using a Fast Fourier Transform technique. The exponents of the calculated frequency power spectral densities have a distribution that peaks at -1.5 and 80% of the values fall within the range of -1.4 to -2.2. Since the power spectral density of a random walk process has an exponent equal to -2.0, it is clear that the observed frequency fluctuations are a close approximation to a random walk. From the FFT calculations and other noise measurement techniques, the FM noise power spectrum has been determined for two delay

line oscillators for noise frequencies from 1.5×10^{-8} Hz to 4×10^7 Hz. From 1.5×10^{-8} Hz to approximately 10^{-3} Hz the exponent is close to -2.0 (-4.0 for phase noise) and consequently represents a random walk in frequency. From 10^{-3} Hz to approximately 10^4 Hz the exponent is close to -1.0 (-3.0 for phase noise) which corresponds to the familiar $1/f$ or "flicker" noise. Beyond 10^4 Hz are the usual white frequency and white phase regions.

Introduction

The long-term frequency stability of SAW controlled oscillators is of considerable importance to many applications where the ultimate in frequency stability is desired. As a result, there has been a significant effort, in a number of laboratories, directed toward understanding and reducing the causes for frequency changes that occur over time periods ranging from days to years. With improvements in device fabrication and packaging, the systematic long-term frequency drift of SAW oscillators fabricated in our laboratory has been reduced to the level where 19 out of 23 oscillators have exhibited a frequency drift of less than 2 ppm after 1 year.¹ Efforts in the reduction of systematic drift are continuing and there is evidence that a total frequency drift of well under 1 ppm after the first year is not an unreasonable expectation. However, with reduced systematic drift, it has become apparent that there is a significant level of random frequency fluctuations occurring with periods ranging up to months and years. Figure 1 illustrates this point with aging data from four 400-MHz delay line SAW oscillators. All four curves show some level of systematic upward frequency drift,

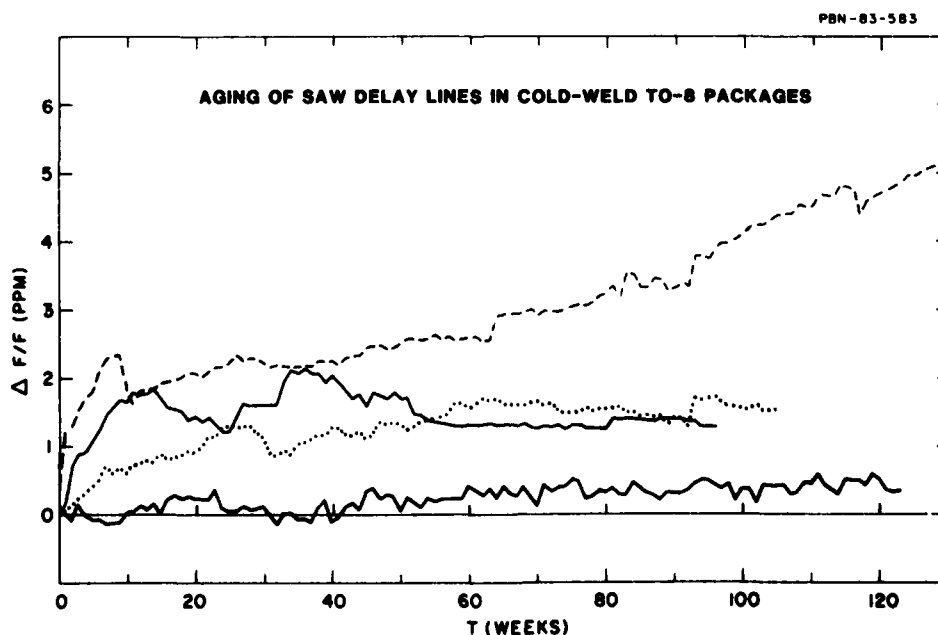


Figure 1. Typical aging data from four 400 MHz delay line oscillators sealed in cold-weld TO-8 packages.

AD P 002507

but all four also show significant levels of random frequency fluctuations. The lower solid curve clearly demonstrates the importance of these random fluctuations. After more than 120 weeks of operation, this device has drifted less than 0.5 ppm. This information alone might lead one to expect a week-to-week stability of 4×10^{-9} . However, the rms deviation from the systematic drift for this device is actually 0.1 ppm. Thus, week-to-week variations on the order of 1×10^{-7} are not unusual. At this point, it is important to make clear that there is no experimental evidence to indicate any connection between the systematic drift and the random frequency fluctuations.

Initially it was thought that the random frequency fluctuations in the aging data were caused by temperature fluctuations in the environmental chamber where the aging test was being conducted. This chamber exhibits rms temperature fluctuations of approximately 0.25°C at 30°C. However, about 2 years ago, a new type of oven was added to the aging facilities and these ovens provide a rms temperature stability better than 0.04°C at 55°C. Now that aging data is available on ten 400 MHz delay line oscillators operated in these ovens, it has become clear that temperature fluctuations are not causing the observed frequency fluctuations. The upper solid curve in Figure 1 shows the long-term frequency drift for one of the oscillators in the 55°C ovens. The magnitude of the observed random fluctuations in frequency for this curve is essentially the same as for the other three curves, which are from oscillators in the 30°C chamber. In fact, the average value for the rms frequency fluctuations is equal for devices in both types of chambers and is 0.3 ppm.

Spectral Analysis

Previous noise measurements on SAW oscillators² have shown that $1/f$ frequency fluctuations are present in the range of 0.1 Hz to 10^3 Hz and that typical rms fractional frequency deviations are on the order of 10^{-10} to 10^{-9} . ($1/f$ frequency fluctuations correspond to phase noise with a $1/f^3$ spectral density.) Calculations based on equations from Ref. 3 have shown that if this $1/f$ noise spectrum continues all the way down to frequencies of 10^{-8} Hz (~ 3 years), it could not give fluctuations in aging data any larger than about 4×10^{-9} . Therefore, there must be a different noise mechanism present at very low modulation frequencies. To determine the nature of this noise process, a spectral analysis of the aging data from 45 oscillators (43 delay lines and 2 resonators) has been carried out using a Fast Fourier Transform (FFT) technique.

The first step in performing the spectral analysis was to remove the systematic drift that is clearly present in some of the aging data. Rather than fitting the aging data to an arbitrary polynomial equation, three equations were chosen that have some physical basis for aging mechanisms.⁴ These are:

$$\Delta F/F = A \log(Bt+1) \quad (1)$$

$$\Delta F/F = A(1 - e^{-Bt}) \quad (2)$$

$$\Delta F/F = A(t)^B \quad (3)$$

where A and B are fit parameters and t is time. Out of the 45 aging curves that were analysed, 25 required the removal of systematic drift. This was accomplished by subtracting from the aging data calculated values derived from the equation that gave the best fit. Eq. (1) gave the best fit in 36% of the cases, Eq. (2) in 32% and Eq. (3) in 32%. Even after removal of the systematic drift, all of the curves had to be weighted

with a cosine taper in the first and last 10% of the data points in order to reduce leakage.⁵

The aging data for an oscillator consists of up to 130 points, each representing the average fractional deviation (in ppm) from the start frequency for 1 week. One hundred and thirty is a relatively small number of points to perform an FFT on, so a preliminary evaluation was done on computer generated data. Using a pseudo random number generator, artificial aging data corresponding to a one-dimensional random walk was generated by plotting the sum of the random numbers as a function of time. Such a random walk process has a power spectral density proportional to $1/f^2$ where f is the Fourier frequency. The computer-generated curves had 100 steps and 90 different curves were analyzed after suitable cosine weighting. Figure 2 shows the distribution of calculated exponent values for the power spectral density. As expected, the distribution peaks around 2, but the relatively small number of points in each curve has caused the distribution to broaden. The values of the calculated exponents range from 1.4 to 2.6.

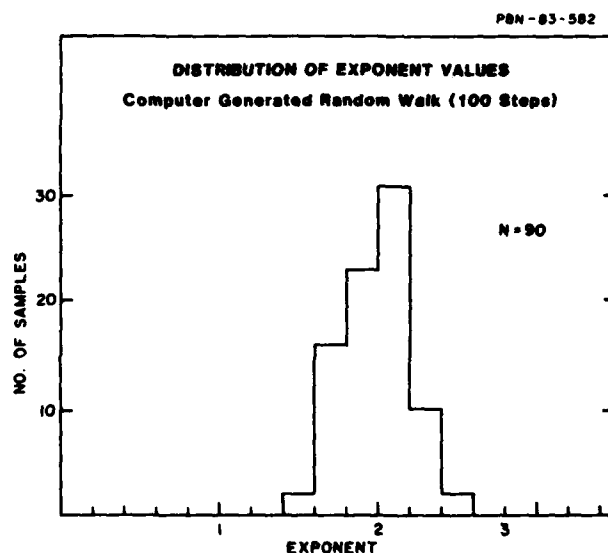


Figure 2. Distribution of calculated exponent values for 90 computer generated curves of 100 step random walks.

With confidence that 100 points were enough to give reasonably accurate values for the exponent of the power spectral density, FFT analysis of aging curves from 45 SAW oscillators was carried out. Figure 3 shows the distribution of the calculated exponent values. Though the peak of the distribution lies at 1.5, nearly 80% of the values lie between 1.4 and 2.2. Thus, it appears that the observed frequency fluctuations for most of the oscillators are a close approximation to a random walk process. (For a power spectral density of frequency fluctuations which varies as $1/f^2$, the corresponding power spectral density for phase fluctuations varies as $1/f^4$.) It is interesting to note that there also appears to be a distinct, but smaller, peak near the value 1. This indicates that a small fraction of the oscillators may have a $1/f$ power spectrum at very low frequencies, although the level must be higher than what would be obtained by extrapolating from measured $1/f$ levels in the 0.1 Hz to 10^3 Hz range.

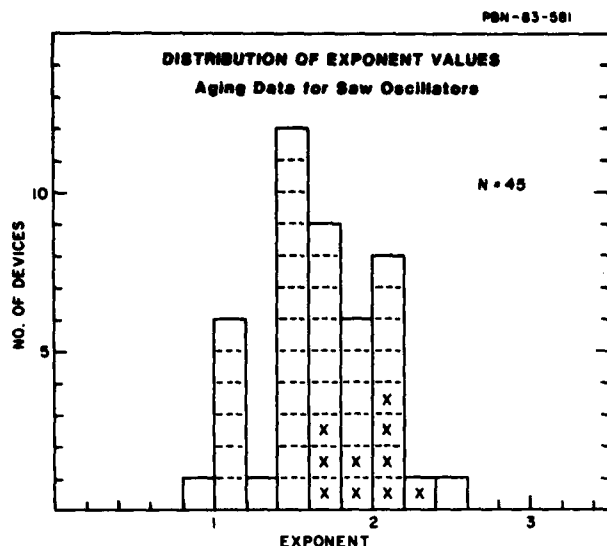


Figure 3. Distribution of calculated exponent values for the FM power spectral density of 45 SAW oscillators. The blocks marked with an X represent oscillators in the more stable ovens.

The boxes in Figure 3 which are marked with an X represent the ten oscillators which were aged in the 55°C ovens with the improved temperature stability. The calculated exponent values for these oscillators are more evenly distributed around 2 than the values for all of the other oscillators in the less stable chamber. Though we noted earlier that the rms deviations for oscillators in the two chambers were essentially the same, it does appear from the calculated values of the exponents that some whitening of the power spectral density is occurring in the less stable chamber.

In addition to the exponent values shown in Figure 3, the FFT analysis also yields the magnitude of the power spectral density. This information in conjunction with other noise measurement techniques has made it possible to evaluate the noise power spectral density for SAW oscillators from 1.5×10^{-8} Hz to 4×10^7 Hz. This complete analysis has been performed on two 400 MHz delay line oscillators and Figure 4 shows the results for one of them. (The other one is not shown since the results are nearly identical.) In Figure 4 the log of the power spectral density of frequency fluctuations for a 1 Hz bandwidth is plotted as a function of the Fourier (or modulation) frequency. In region 1, the data is from the spectral analysis of aging data. In region 2, the same FFT analysis was performed on several 1100 point data sets acquired with an HP 5345A electronic counter (with 5358A measurement storage) for gate times ranging from 5 sec to 200 sec. In region 3 the same electronic counter was used but the data was acquired and evaluated via the Hadamard variance.⁶ In region 4 an analog phase locked loop technique was used. The dashed portion of the curve from 3×10^5 Hz to 4×10^7 Hz represents data acquired from another but similar oscillator. This region of the spectrum is well understood and the extrapolation is made with considerable confidence.

The data in Figure 4 shows that from 1.5×10^{-8} Hz to 1×10^{-3} Hz the noise decreases as $1/f^2$ and therefore corresponds to a random walk in frequency. From 10^{-3} Hz to beyond 10^4 Hz the noise decreases as $1/f$ and therefore corresponds to the familiar "flicker" noise.

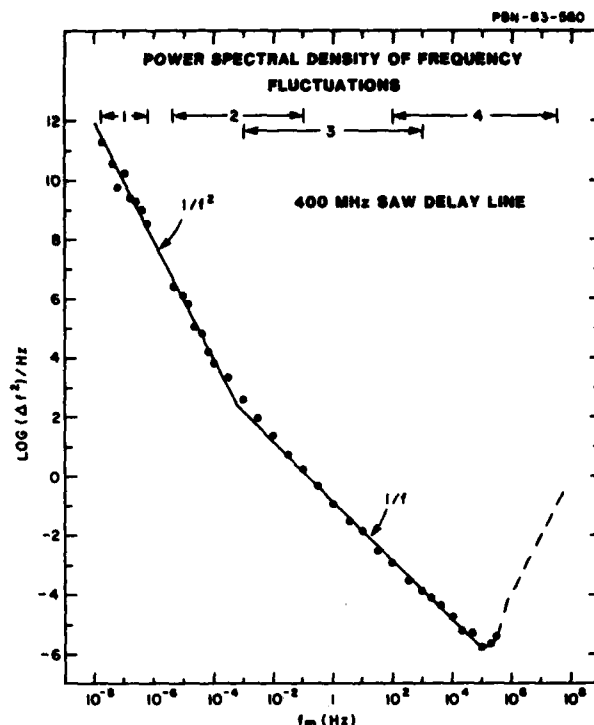


Figure 4. Power spectral density of frequency fluctuations for a 400 MHz SAW delay line oscillator. Region 1-FFT of aging data; Region 2 - FFT of 1100 point data sets; Region 3 - Hadamard variance; Region 4 - Phase locked loop.

Because the flicker noise level is fairly high on this device, there is essentially no region of white frequency noise. Beyond 10^5 Hz there is white phase noise. By using the relation

$$(\Delta\phi)^2 = \frac{(\Delta f)^2}{f_m^2} \quad (4)$$

the power spectral density of frequency fluctuations shown in Figure 4 can be translated into the power spectral density of phase fluctuations. Figure 5 shows the same data as used in Figure 4, but translated into phase fluctuations by using Eq. (4). In the regions of the curve where $10 \log (\Delta\phi^2/2)$ is negative, the data corresponds to the familiar single sideband noise to carrier ratio (in dBc). Note that the $1/f^2$ section of Figure 4 translates into a $1/f^4$ section in Figure 5 and that the $1/f$ section becomes $1/f^3$.

Another significant observation made during the course of this investigation was that there may be a correlation between the noise at very low frequencies and the $1/f$ noise at higher frequencies. This is illustrated in Figure 6. On the horizontal axis the measured phase noise level at 1 Hz (indicated as N_f) which is in the flicker noise region, is plotted for each of 41 oscillators. In the vertical direction, the corresponding phase noise level at 10^{-7} Hz is plotted (indicated as N_A). The latter will be referred to as aging noise and for most of the devices this is representative of a random walk process. Normal delay line-type oscillators are indicated by

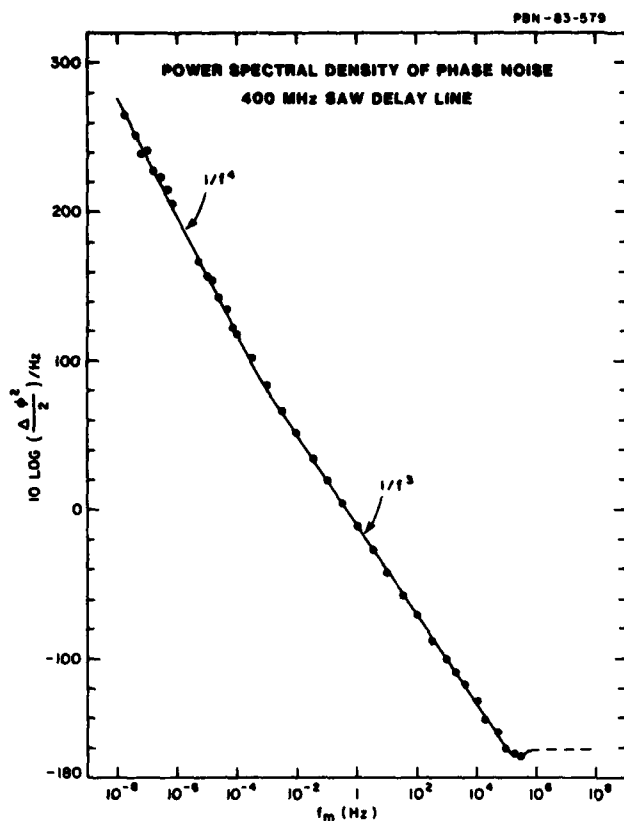


Figure 5. Power spectral density of phase noise for a 400 MHz SAW delay line oscillator. This curve was calculated for the same data used in Figure 4.

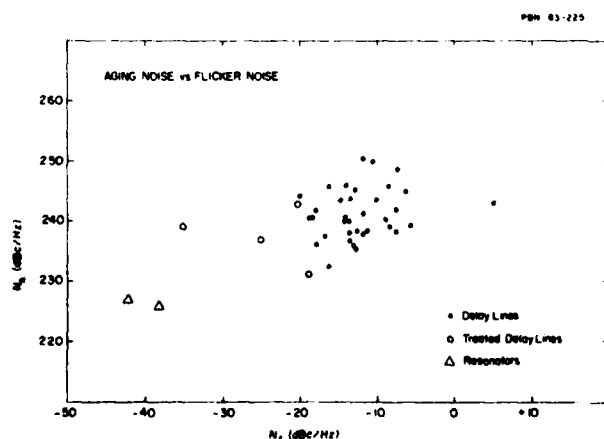


Figure 6. Plot of phase noise levels at 10^{-7} Hz (N_A) versus phase noise at 1 Hz (N_E) for 41 SAW oscillators.

solid dots, silicone-treated delay line oscillators² are represented by hollow dots and the two resonators are indicated by triangles. In general, there is considerable scatter from device to device, but several characteristics are evident. In the flicker noise region the resonators are definitely the quietest devices. Interestingly, they are also the quietest in the aging noise region. The silicone-treated delay lines are also quieter than normal delay lines in the flicker noise region but not as quiet as the resonators. In the aging noise region the treated delay lines also appear to be quieter than the normal delay lines, but the small number of treated devices and the large scatter present makes it difficult to be definite. Though the data in Figure 6 is not yet conclusive, it does appear that there may very well be a correlation between flicker noise and aging noise.

To illustrate the effect of lower aging noise, three aging curves have been plotted in Figure 7. The top curve is for a resonator, the middle curve is for a silicone-treated delay line and the bottom curve is for a normal untreated delay line. Ignoring the differences in systematic drift, the lower levels of the random fluctuations in the resonator and the treated delay line are clearly visible. It is also interesting to note that the level of fluctuations with periods on the order of several weeks tends to decrease with time. A similar decrease has also been observed in the flicker noise.

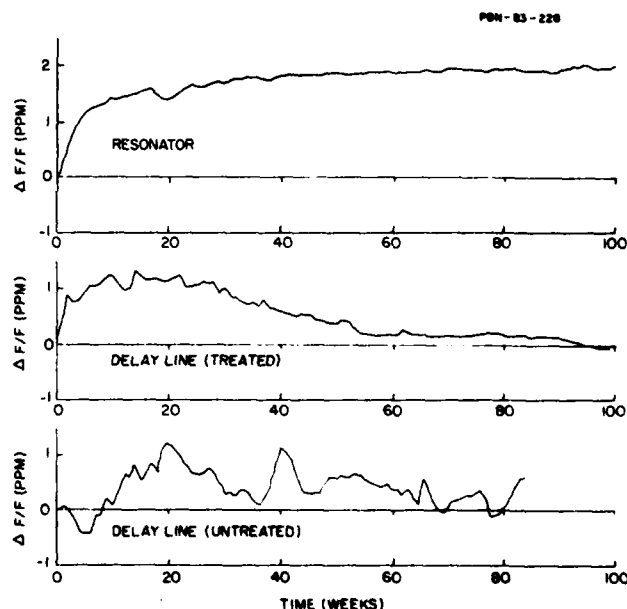


Figure 7. Aging curves for a SAW resonator, a silicone-treated SAW delay line, and an untreated SAW delay line.

At very low frequencies a $1/f^2$ (random walk) dependence has also been observed on bulkwave resonators⁷, and the source was identified as temperature fluctuations. This is not the case with the SAW oscillators in the present study because, as mentioned earlier, oscillators in two ovens with very different levels of rms temperature fluctuations exhibited the same level of frequency fluctuations (0.3 ppm). Furthermore, two resonators with the same temperature dependence as the delay lines showed a much smaller rms frequency fluctuation of 0.05 ppm even though they were in the chamber with the larger temperature fluctuations.

Conclusions

Though the investigation of very long period random frequency fluctuation in SAW oscillators has just begun, there are several conclusions that can be drawn. One is that, in most delay line oscillators, the very low frequency fluctuations are a close approximation to a random walk in frequency. Second, the aging noise is lower in resonator oscillators than delay line oscillators, and therefore, excluding differences in systematic long-term stability, the resonator is the preferred choice where ultimate long-term stability is desired. (There is insufficient data available at this time to state with certainty whether resonators exhibit a random walk at very low frequencies.) Finally, since the aging noise process extends all the way to frequencies on the order of 10^{-3} Hz, it is a relatively easy task to measure the aging noise level over a period of 1 to 2 days. This makes it possible to evaluate relative levels of random fluctuations in long-term stability without waiting weeks or months.

The fundamental question that remains to be answered, however, relates to the source of the noise. At this point, it is not clear whether the SAW device is the source or whether it is some other component in the oscillator. It is known that the SAW device is the source of flicker noise in delay line oscillators² and, therefore, any correlation between flicker noise and aging noise would be significant.

Acknowledgements

The author would particularly like to thank Don Lee, Jim Mullen, and Hans Bilger for help with the FFT analysis, and Joe Callera for encouragement and

helpful discussions. Also the valuable assistance of John Lang and Merton Bennett in fabricating and testing of the devices is acknowledged.

References

1. T. E. Parker, "Precision Surface Acoustic Wave (SAW) Oscillators", 1982 IEEE Ultrasonics Symposium, p. 268 (1982).
2. T. E. Parker, "1/f Phase Noise in Quartz Delay Lines and Resonators", 1979 IEEE Ultrasonics Symposium, p. 878 (1979).
3. J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healey, D. B. Leeson, T. E. McGunigal, J. A. Mullen, W. L. Smith, R. L. Sydnor, R. E. Vessot, and G. M. R. Winkler, "Characterization of Frequency Stability", IEEE Trans. on Instrumentation & Meas., Vol. IM-20, pp. 105-120 (1971).
4. T. E. Parker, "Analysis of Aging Data on SAW Oscillators", 34th Annual Symposium on Frequency Control, p. 292 (1980).
5. J. S. Bendat and A. G. Piersol, Random Data: Analysis and Measurement Procedures, New York: Wiley-Interscience, pp. 316, 323 (1971).
6. R. A. Baugh, "Frequency Modulation Analysis with the Hadamard Variance", 25th Annual Symposium on Frequency Control, p. 222 (1971).
7. Y. Noguchi, Y. Teramachi, and T. Masha, "1/f Frequency Fluctuations of a Quartz Crystal Oscillator and Temperature Fluctuation", 35th Annual Frequency Control Symposium, p. 484 (1981).

AD P 002508

STRESS INDUCED EFFECTS ON THE PROPAGATION OF SURFACE WAVES

B. K. Sinha and W. J. Tanski
Schlumberger-Doll Research
P. O. Box 307
Ridgefield, CT 06877

T. Lukaszek and A. Ballato
U.S. Army Electronics Technology and Devices Laboratory
USAERADCOM
Fort Monmouth, NJ 07703

ABSTRACT

It is known that the propagation characteristics of elastic waves are influenced by the nature of biasing stress distributions in the propagating medium. A knowledge of the magnitude of such effects plays an important role in improving the performance and long term aging characteristics of SAW devices. Generally, SAW devices employ thin plate-like structures as the propagating medium of surface waves. The biasing deformation of such structures due to externally applied forces can be decomposed into extensional and flexural deformations which can have substantially different effects on the change in the time delay between two observation points. This paper describes a study of the fractional change in the time delay of surface waves due to the stress components in terms of which a general deformation can be expressed. The information on the biasing state of thin discs has been employed in a previously reported perturbation procedure for equations of motion for small dynamic fields superposed on a static bias. Computational results have been obtained for fractional changes in the time delay of surface waves in crystalline quartz of various orientations and propagation directions due to many simple stress systems in thin discs. We have considered the cases of flexural deformation of rectangular plates due to cylindrical bending and discs subjected to normal pressures, as well as extensional deformation due to a pair of forces acting diametrically on a circular disc along various azimuthal angles. Experimental results have been obtained for the moment-frequency effect on circular ST-cut quartz discs and the force-frequency effects on both singly, as well as doubly rotated quartz discs. Good agreement has been obtained between the theoretical predictions and experimental measurements.

INTRODUCTION

There are various sources of biasing stresses in the propagating medium of Surface Acoustic Wave (SAW) devices. The influence of such stresses on the propagation characteristics of surface waves can significantly affect the performance and long term aging of high precision quartz devices¹⁻¹⁴. For instance, it is known that sawing, lapping and mechanical polishing processes produce a damaged surface layer of thickness approximately 50 to 100 Å with crystalline imperfection and large intrinsic surface stresses. The relaxation of such stresses with time is responsible for some of the drift in device characteristics. Other sources of such stresses are the intrinsic and electrode induced thermal stresses, mounting and acceleration stresses as well as stresses resulting from the externally applied forces and pressures. While most of the applications of high precision quartz resonators in frequency control and signal processing systems require those

orientations and propagation directions which are largely insensitive to both temperature and stress induced effects, there are also applications of quartz resonators in the development of sensors where one seeks to work with those orientations which have large sensitivity to the stresses produced by the externally applied load. In addition, a general study of stress induced effects on the propagation of elastic waves is of central importance in the area of non-destructive testing of structural materials. In particular, stress induced frequency shifts can also be employed to estimate the magnitude of intrinsic, surface stresses and thereby characterize the fabrication process.

In this paper we describe some analytical and experimental results on the stress induced effects on the propagation of piezoelectric surface waves in crystalline quartz. The analysis is based on the equations of motion for small dynamic fields superposed on a linear, static bias which have been derived from the rotationally invariant equations of nonlinear elasticity. When the solid is subject to a homogeneous bias, the equations of motion and boundary conditions can be solved in the usual manner after replacing the linear constants with the effective elastic constants referred to the reference (undeformed) configuration of the solid. However, when the biasing state is inhomogeneous as is the case with flexural deformation of thin plates, the effective elastic constants are position dependent and a direct solution of the boundary value problem is not possible. In the latter case, a perturbation procedure can readily treat spatially varying states and small changes in the time delay of surface waves between two observation points.

We have considered extensional and flexural deformations of thin plates. Any stress distribution through the thickness of the plate can be expressed in terms of such symmetric and anti-symmetric deformations. Computational results for the fractional change in frequency of SAW resonators on thin, circular discs subjected to a pair of diametrical forces acting at various azimuthal angles have been compared with experimental values obtained for Y-cut, ST-cut and a doubly-rotated cut of quartz. Good agreement has been obtained between the theoretical predictions and experimental measurements. The influence of flexural deformation of thin plates due to cylindrical bending along and normal to the propagation direction of surface waves on the fractional change in the time delay has been analytically studied for several orientations of quartz. The experimental set up for this case consists of a thin, circular disc of ST-cut quartz, subjected to a radial moment along its periphery which produces a bi-axial stress state at the surface of the disc where the SAW resonator is placed. The aforementioned bending stress components are along and normal to the propagation direction of surface waves. The total shift in the SAW resonator frequency is obtained by a linear of superposition of results due to the two cylindrical bending deformations considered separately. Again the agreement with

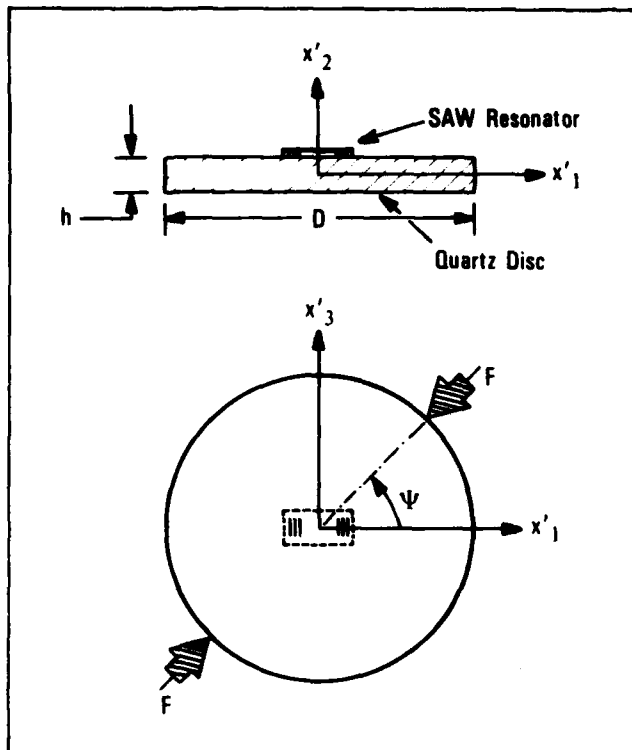


FIGURE 1 - Schematic diagram of a thin circular disc with a SAW delay line (or resonator) subjected to a pair of diametrical forces acting at an azimuthal angle ψ .

the experimental results from two SAW devices on ST-cut quartz is good and the principle of linear superposition of results due to different linear biasing states is validated.

STRESS INDUCED FREQUENCY SHIFTS IN SAW DEVICES

When a surface acoustic wave delay line in the form of a plate-like structure is subject to an externally applied load or pressure, the wave travel time generally changes as a result of pressure (stress) induced changes in the path length, elastic stiffnesses and mass density of the propagating medium. Most of the elastic deformations of thin, anisotropic plates can be described in terms of extensional and flexural deformations of such plates.

Extensional Stresses in Thin Plates

A schematic diagram of a circular plate subject to a pair of diametrical forces is shown in Figure 1. The change in the resonant frequency can be computed from the perturbation equation discussed in the next Section provided the biasing state and eigensolution for the piezoelectric surface waves are known. In the case of isotropic plates, the stresses at the center of the plate can be obtained from the solution of biharmonic equation in the Airy's stress function. The magnitude of such stresses at the center due to a pair of diametrical forces acting along the X_1 -axis is given by

$$T_{11} = -\frac{6}{\pi} \frac{F}{hD}, T_{33} = \frac{2}{\pi} \frac{F}{hD}, \text{ and } T_{13} = 0, \quad (1)$$

where T_{11} , T_{33} and T_{13} are the compressive, tensile and shear stress components, respectively; F is the magnitude of force, h and D are the thickness and diameter of the disc, respectively.

The appropriate plane stress constitutive relations for thin plates under extensional loading are given by^{7,8}

$$K_R^{(0)} = \gamma_{RS} E_S^{(0)}, \quad R, S = 1, 3, 5, \quad (2)$$

where

$$\gamma_{RS} = c_{RS} - c_{RW} c_{WV}^{-1} c_{VS}, \quad W, V = 2, 4, 6, \quad (3)$$

and the compressed matrix notation for stresses $K_R^{(0)}$ and strains $E_S^{(0)}$, have been employed. This implies that

$$(K_1^{(0)}, K_3^{(0)}, K_5^{(0)}) \equiv (T_{11}, T_{33}, T_{13}),$$

and

$$(E_1^{(0)}, E_3^{(0)}, E_5^{(0)}) \equiv (E_{11}, E_{33}, E_{13}). \quad (4)$$

The γ_{RS} are Voigt's anisotropic plate elastic constants for generalized plane stress conditions. Since the normal stress components are set equal to zero, the corresponding strain components are given by

$$E_W^{(0)} = -c_{WV}^{-1} c_{VS} E_S^{(0)}. \quad (5)$$

It is known that the solution to a boundary-value problem in static linear elasticity is unique only to within a static, homogeneous, infinitesimal rigid (body) rotation. Furthermore, it has been shown⁷ that a homogeneous infinitesimal three-dimensional rotation does not induce any change in the time delay of elastic waves in a propagating medium subject to a homogeneous bias. For the case of extensional deformation due to a pair of diametrical forces, we have assumed that the SAW device is occupying a region of homogeneous stress bias.

Therefore, without any loss in generality, we select the homogeneous rigid rotation to vanish which yields

$$\Omega_{KL} = \frac{1}{2} (w_{K,L} - w_{L,K}) = 0. \quad (6)$$

Under the aforementioned condition, the biasing displacement gradients are simply given by

$$w_{L,K} = E_{KL}. \quad (7)$$

Flexural Stresses in Thin Plates

For flexural deformation of anisotropic plates (Figure 2), the constitutive relations between the components of bending moment and curvature are given by⁵

$$K_R^{(1)} = \frac{2}{3} h^3 \gamma_{RS} E_S^{(1)}, \quad R, S = 1, 3, 5, \quad (8)$$

where $2h$ is the thickness of the plate, $K_R^{(1)}$ and $E_S^{(1)}$ are planar components of moments and curvatures and are defined by

$$(K_1^{(1)}, K_3^{(1)}, K_5^{(1)}) = \int_{-h}^h z_2 (T_{11}, T_{33}, T_{13}) dz_2,$$

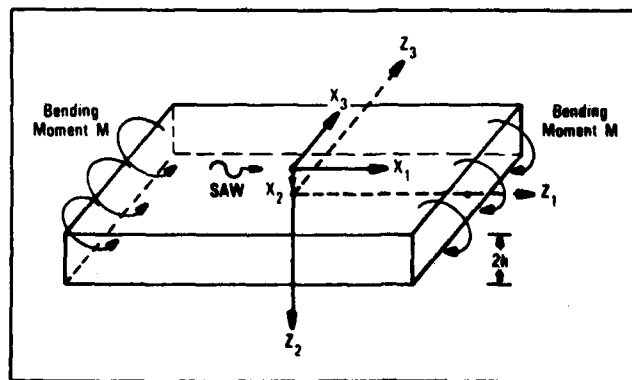


FIGURE 2 - Schematic diagram of a thin rectangular plate with a SAW delay line (or resonator) subjected to a pure bending moment along a pair of its edges.

and

$$(E_1^{(1)}, E_3^{(1)}, E_5^{(1)}) = - (w_{2,11}^{(0)}, w_{2,33}^{(0)}, w_{2,13}^{(0)}). \quad (9)$$

Since we allow for the free normal and shear components of strain along the thickness direction, the first order stress resultants

$$(K_6^{(1)}, K_2^{(1)}, K_4^{(1)}) \equiv \int_{-h}^h z_2 (T_{21}, T_{22}, T_{23}) dx_2, \quad (10)$$

are allowed to vanish and we can solve for the corresponding strains which are given by

$$E_w^{(1)} = -c_w^{-1} c_{vs} E_s^{(1)}, \quad w, v = 2, 4, 6, \quad (11)$$

where compressed matrix notation has been employed.

The three-dimensional strain field in this case of anisotropic flexure considered here, is given by

$$E_{KL} = z_2 E_{KL}^{(1)}, \quad (12)$$

and the associated three-dimensional displacement field w_K has the form

$$w_A = z_2 w_A^{(1)}, \quad w_2 = w_2^{(0)} + z_2^2 w_2^{(2)}. \quad (13)$$

For restrained motion normal to the length of the plate the three-dimensional displacement gradient field resulting from the flexural loading is given by

$$w_{1,1} = z_2 E_1^{(1)}, \quad w_{1,2} = z_1 E_1^{(1)} + z_2 E_6^{(1)}, \quad w_{K,3} = 0, \quad (14)$$

$$w_{2,1} = -z_1 E_1^{(1)}, \quad w_{2,2} = z_2 E_2^{(1)}, \quad w_{3,1} = 0, \quad (15)$$

$$w_{3,2} = z_2 E_4^{(1)}, \quad (16)$$

where

$$E_1^{(1)} = \frac{3M}{2h^3 \gamma_{11}}, \quad (17)$$

and M is the bending moment per unit width of the plate. Equation (15), has appropriately accounted for the vanishing of the slope of the deflection curve at $z_1 = 0$, and together with Eq. (14)₂ results in zero lowest order vertical shearing strain. The corresponding stresses are, respectively, given by

$$(T_{11}, T_{33}, T_{31}) = (\gamma_{11}, \gamma_{31}, \gamma_{51}) w_{1,1}. \quad (18)$$

For traction (stress) free conditions normal to the length of the plate, the three-dimensional displacement gradient field resulting from the flexural loading is given by

$$w_{1,1} = z_2 E_1^{(1)}, \quad w_{1,2} = z_1 E_1^{(1)} + z_2 E_6^{(1)}, \quad (19)$$

$$w_{2,1} = -z_1 E_1^{(1)}, \quad w_{2,2} = z_2 E_2^{(1)}, \quad w_{2,3} = 0, \quad (20)$$

$$w_{3,1} = 0, \quad w_{3,2} = z_2 E_4^{(1)}, \quad (21)$$

$$w_{1,3} = -\frac{(\gamma_{33}\gamma_{51} - \gamma_{53}\gamma_{31})}{(\gamma_{33}\gamma_{55} - \gamma_{35}\gamma_{53})} w_{1,1}, \quad (22)$$

$$w_{3,3} = -\frac{(\gamma_{55}\gamma_{31} - \gamma_{35}\gamma_{51})}{(\gamma_{33}\gamma_{55} - \gamma_{35}\gamma_{53})} w_{1,1}, \quad (23)$$

where γ 's are defined Eq. (3). The flexural stress T_{11} , under these conditions is given by

$$T_{11} = \left[\gamma_{11} + \frac{\gamma_{13}(\gamma_{35}\gamma_{51} - \gamma_{55}\gamma_{31}) + \gamma_{15}(\gamma_{53}\gamma_{31} - \gamma_{33}\gamma_{51})}{(\gamma_{33}\gamma_{55} - \gamma_{35}\gamma_{53})} \right] w_{1,1}. \quad (24)$$

We can in a similar manner, obtain the biasing stresses, strains and displacement gradients when the flexure is along the z_3 direction which is normal to the propagation direction of surface waves. The corresponding results are obtained by interchanging the indices 3 and 1 throughout in the expanded tensorial notation first, and then reverting back to the compressed

notation.

PERTURBATION PROCEDURE

When the solid is subject to a homogeneous bias, such as due to a uniform change in its temperature, the equations of motion and boundary conditions in terms of effective elastic constants can be solved in a straight-forward manner. However, when the biasing state is inhomogeneous as is the case in a flexural deformation or non-uniform temperature distribution, the effective elastic constants are position dependent and a direct solution of the boundary value problem is not possible. In the latter case, a perturbation procedure⁶ can readily treat spatially varying states and small changes in the elastic wave velocity can be calculated. We can write the effective elastic constant⁶⁻⁸ $G_{LYQ\alpha}$ as a sum of the ordinary elastic constants $c_{LYQ\alpha}$ and an additional term $\hat{c}_{LYQ\alpha}$ due to the bias, which takes the convenient form

$$G_{LYQ\alpha} = c_{LYQ\alpha} + \hat{c}_{LYQ\alpha}, \quad (25)$$

where

$$\hat{c}_{LYQ\alpha} = T_{LQ}^i \delta_{\gamma\alpha} + c_{LYQ\alpha AB} E_{AB}^i + c_{LYPQ} w_{\alpha,P} + c_{LPQ\alpha} w_{\gamma,P}. \quad (26)$$

In Eq. (26), T_{LQ}^i , E_{AB}^i and $w_{\alpha,P}$ are the biasing stress, strain and displacement gradient components, respectively; and the coefficients c_{LYPQ} and $c_{LYQ\alpha AB}$ are the second and third order elastic constants, respectively. Since we are interested only in the first perturbation of eigenvalue which has numerous applications, the perturbation equation can be derived in a straight-forward manner by assuming the existence of solutions in both the unperturbed and a nearby perturbed state. The Piola-Kirchhoff stress equations of motion in the presence of a bias may be written in the form⁶

$$K_{LY,L}^i + K_{LY,L}^n = \rho_0 \ddot{u}_\gamma, \quad (27)$$

where

$$K_{LY,L}^i = c_{LYQ\alpha} u_{\alpha,Q},$$

$$K_{LY}^n = \hat{c}_{LYQ\alpha} u_{\alpha,Q}, \quad (28)$$

and, K_{LY}^n and ρ_0 , are the mechanical perturbation term and mass density in the reference state, respectively. A harmonic, eigensolution u_γ^m , at frequency ω_m satisfies the equation of motion (27) in the absence of any perturbation terms, whereas the solution u_γ in the nearby perturbed state satisfies Eq. (27). The two aforementioned solutions can be employed to form the equation

$$\int_V \left[(K_{LY,L}^m + \rho_0 \omega_m^2 u_\gamma^m) u_\gamma \right. \\ \left. - (K_{LY,L}^i + K_{LY,L}^n + \rho_0 \omega^2 u_\gamma) u_\gamma^m \right] dV_0 = 0. \quad (29)$$

The small perturbation in the eigenfrequency ω can be obtained from Eq. (29) with the aid of Eqs. (25-27), and the divergence theorem in the usual manner. The resulting equation can be expressed in the form

$$(\omega_m^2 - \omega^2) \int_V \rho_0 u_\gamma^m u_\gamma dV_0 \\ - \int_S N_L \left[K_{LY}^i u_\gamma^m - K_{LY}^n u_\gamma \right] dS_0 \\ + \int_V K_{LY,L}^n u_\gamma^m dV_0. \quad (30)$$

Since the perturbations in the eigenvalue and eigenvector are small, the second- and higher-order terms can be neglected, and the first perturbation in the eigenfrequency is given by

$$\Delta_m = H_m/2\omega_m, \quad \Delta_m = \omega_m - \omega, \quad (31)$$

where for traction-free surface in the unperturbed state, we have

$$H_m = - \int_V K_{Ly}^n g_{y,L} dV_0. \quad (32)$$

In obtaining Eq. (32) we have employed the traction-free boundary conditions in the presence of the bias and the orthonormal eigensolution g_y of the homogeneous problem, which is defined by

$$g_y = \frac{u_y^m}{N_{(m)}}, \quad N_{(m)}^2 = \int_V \rho_0 u_y^m u_y^m dV_0. \quad (33)$$

Substitution of the appropriate constitutive relation Eq. (4.4) into (4.8) yields

$$H_m = - \int_V \hat{c}_{LyMy} g_{y,M}^m g_{y,L}^m dV_0. \quad (34)$$

It has been shown that the first perturbation in phase velocity and wave number can be obtained from the perturbation in eigenfrequency. The relations are

$$\Delta V = \Delta_m/\xi_m, \quad \Delta \xi = \Delta_m/V_m, \quad (35)$$

where

$$\Delta V = V_m - V, \quad \Delta \xi = \xi_m - \xi, \quad (36)$$

and V_m and ξ_m are the unperturbed phase velocity and wave number, respectively, of the m th eigensolution. In the foregoing the change in the natural velocity of elastic waves due to a biasing state at a constant temperature has been considered. The solution for surface waves propagating in arbitrarily anisotropic, piezoelectric substrates satisfies the differential equations and boundary conditions of linear piezoelectricity for piezoelectric surface waves and may be written in the form

$$u_j = \sum_{m=1}^4 C^{(m)} A_j^{(m)} \exp(i\beta_m \xi X_2) \exp[i\xi(X_1 - Vt)],$$

$$\phi = \sum_{m=1}^4 C^{(m)} \bar{A}_4^{(m)} \exp(i\beta_m \xi X_2) \exp[i\xi(X_1 - Vt)], \quad (37)$$

where u_j and ϕ are the complex representation of the real mechanical displacement and electric potential, respectively, for propagation in the natural direction X_1 with X_2 normal to the surface as shown in Figure 2. The quantities $C^{(m)}$, $A_j^{(m)}$, $\bar{A}_4^{(m)}$ and β_m are determined numerically for traction-free mechanical and open-circuit electrical boundary conditions. The normalized unperturbed surface wave eigensolution for the perturbation equation is obtained from Eq. (33). The normalization integral in Eq. (33) takes the form

$$N^2 = \rho \int_0^\infty dX_2 \int_{-\pi/\xi}^{\pi/\xi} u_j u_j dX_1$$

$$= \rho \frac{\pi i}{\xi^2} \sum_{m=1}^4 \sum_{n=1}^4 \frac{C^{(m)} A_j^{(m)} C^{(n)*} A_j^{(n)*}}{(\beta_m - \beta_n^*)}, \quad (38)$$

where the $*$ denotes complex conjugate and N^2 is, of course, real. Computational results for the fractional change in the resonant frequency of SAW oscillators which consist of a circular disc SAW resonator subject to a pair of diametrical in-plane forces at various azimuthal angles have been obtained and are shown in Figures 3, 4 and 5. It has been assumed in the present calculations that the stresses at the center of the disc are uniform over the entire area of SAW resonators or delay lines. This certainly requires the SAW resonators to occupy a very small area of the disc. It should be noted that the wave propagation direction is always along the diagonal axis for the results shown in Figures 3-5.

Computational results for the fractional change in the time delay per unit change in the maximum strain E_{max} on the free

FORCE-FREQUENCY
AZIMUTH CURVE FOR:
($\psi=0.0^\circ$)
($\theta=0.0^\circ$)
($\gamma=0.0^\circ$)

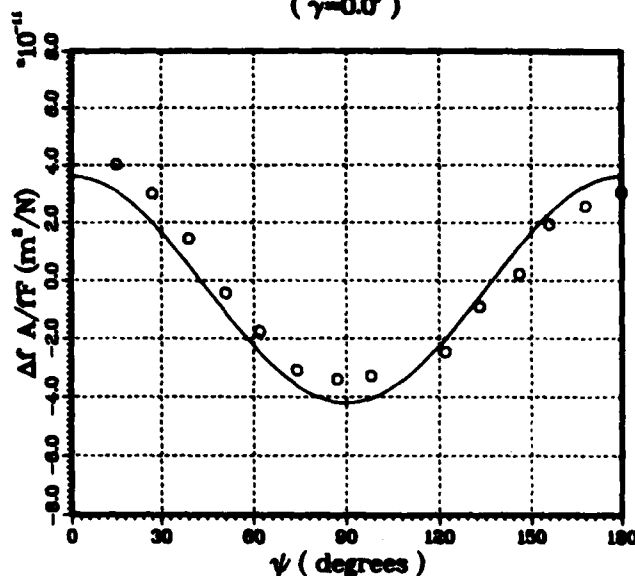


FIGURE 3 - Force-Frequency Coefficient, $\Delta fA/FF$, of a SAW device on Y-cut quartz disc as a function of azimuthal angle ψ . F is the magnitude of forces in Newton and A ($=\pi r^2$) is the cross-sectional area at the center of the disc. The dots represent the experimental points.

FORCE-FREQUENCY
AZIMUTH CURVE FOR:
($\psi=0.0^\circ$)
($\theta=43.75^\circ$)
($\gamma=0.0^\circ$)

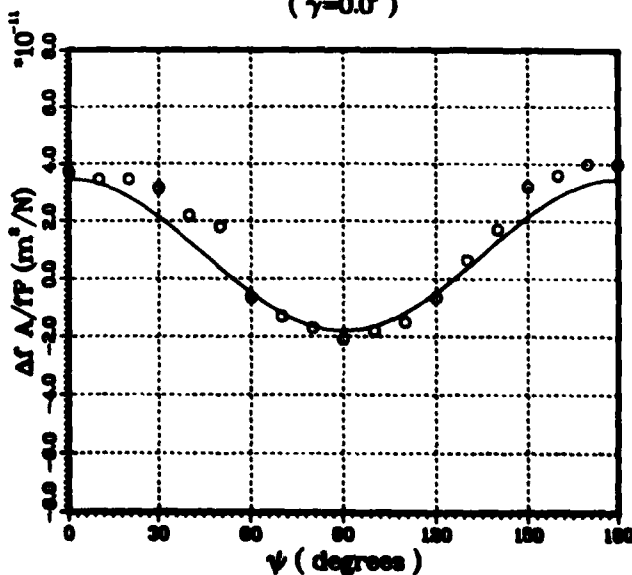


FIGURE 4 - Force-Frequency Coefficient, $\Delta fA/FF$, of a SAW device on ST-cut quartz disc as a function of azimuthal angle ψ . The notation is the same as in Figure 3.

FORCE-FREQUENCY AZIMUTH CURVE FOR:

$$\begin{cases} \psi = -1.05^\circ \\ \theta = 28.07^\circ \\ \gamma = 136.5^\circ \end{cases}$$

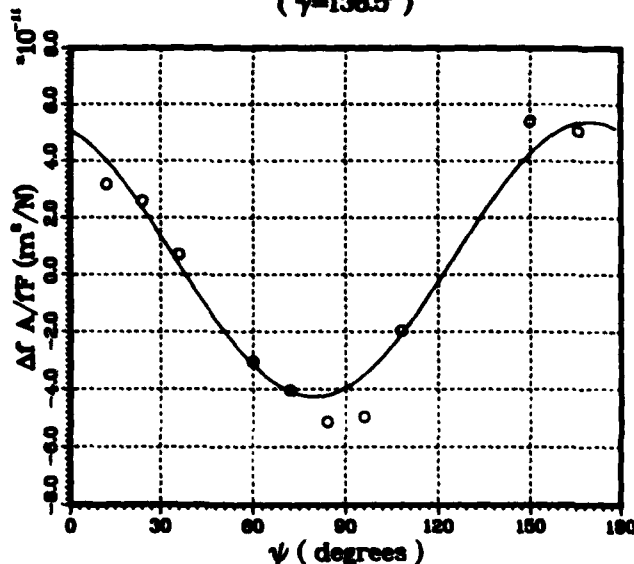


FIGURE 5 - Force-Frequency Coefficient, $\Delta f A/TF$, of a SAW device on a doubly rotated cut quartz ($\Phi = -1.05^\circ$, $\theta = 28.07^\circ$, $\gamma = 136.5^\circ$) disc as a function of azimuthal angle ψ . The notation is same as in Figure 3.

surface of the SAW device are shown in Figure 6 for the rotation angles $\theta = -50^\circ$. It is interesting to note that generally for substrates under flexure whose thicknesses are less than 50 wavelengths, the inhomogeneous bending stress through the thickness causes the surface wave propagation to become dispersive. However, for large thicknesses of substrates, the flexural stress is essentially homogeneous over a thin boundary layer close to the free surface and the surface wave propagation is nondispersive. The results have been obtained for the two cases of flexural biasing stresses being parallel and normal to the surface wave propagation direction. The other pair of edges normal to the direction of flexure of rectangular plates (or substrates) are subject to stress (or traction) free or constrained normal component of displacement. Curves labeled 1 and 2 are for cases when the normal component of displacement is constrained along edges normal to the direction of flexure whereas 3 and 4 are for stress-free conditions along those edges. The propagation direction in Figure 6 is fixed along the diagonal axis. Figures 7 and 8 show the fractional change in the time delay per unit change in E_{max} as a function of propagation direction measured from the diagonal axis for ST-cut and BT-cut quartz substrates. Results have been obtained for flexure parallel as well as normal to the propagation direction and for the traction-free boundary conditions along the edges normal to the flexure. The dotted lines in Figures. 7 and 8 represent temperature induced fractional changes in time delay as a function of propagation direction for surface waves on two rotated Y-cut substrates ($\theta = 42.75^\circ$ and -49.22°), respectively.

EXPERIMENTAL PROCEDURES

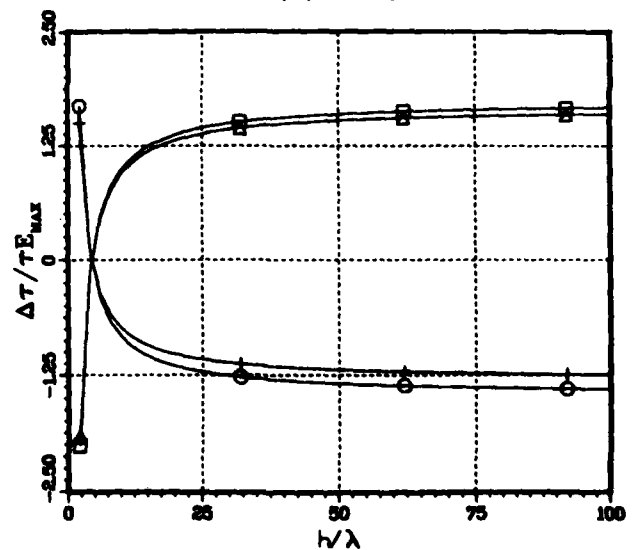
The force-frequency effect has been experimentally studied with SAW resonators on Y-cut and ST-cut quartz substrates and SAW delay lines on a doubly-rotated quartz plate. As part of an example, we describe the details of the experimental set up for Y-cut quartz. The Y-cut disc was of diameter 14mm and thickness 0.207 mm which results in a frequency constant of

1.950 km/s for the lowest thickness-shear (C-mode) mode of vibration. The SAW device employed a one-port resonator produced from a photomask for 60 MHz nominal center frequency on ST-cut substrates. A 40 finger-pair, apodized, central interdigital transducer (IDT) with 5 mm aperture and 100-plus finger-pair reflectors connected by a common bus resulted in a measured resonant frequency of 61.91 MHz. A gap of 1 finger-pair was provided between the central IDT and reflectors. The propagation direction was fixed along the diagonal axis. The electrode material was aluminum and 1 mil gold wires provided attachments to bonding pads. Electrode pattern was aligned with respect to score mark on plate; it is within $\pm 5^\circ$ of crystal axis. This inaccuracy results in an uncertainty in the azimuthal angle ψ , which implies that the abscissa of the force-frequency curve for Y-cut quartz may be shifted.

Force application angle ψ has a starting inaccuracy of $\pm 5^\circ$, but the precision of subsequent angles with respect to the starting angle is less than 1° . The uncertainty in the starting value of ψ produces another possible shift in the abscissa ψ scale. The substrate back surface is attached to a 6.5 mm diameter insulating dowel concentric with the substrate disc. The dowel has attached to it a collar for the electrical connections from the resonator to the double coaxial lines leading to the phase measurement apparatus. The other end of the dowel shaft attaches to a gearing mechanism for accurately turning the quartz plate in ψ angle while maintaining alignment

FLEXURAL BIASING STRESSES PARALLEL AND NORMAL TO THE PROPAGATION DIRECTION

$$\begin{cases} \psi = 0.0^\circ \\ \theta = -50.0^\circ \\ \gamma = 0.0^\circ \end{cases}$$



PARALLEL NORMAL
 $\square = 1$ $\circ = 2$
 $\Delta = 3$ $\pm = 4$

FIGURE 6 - Relative change in time delay per unit applied maximum extensional biasing strain E_{max} vs. thickness-wavelength ratio for a negatively rotated ($\theta = -50^\circ$) quartz plate. Curves 1, 3 and 2, 4 are for cylindrical flexure in and normal to the direction of propagation, respectively. Also note that curves 1 and 2 are for zero normal component of displacement along edges normal to the direction of flexure whereas 3 and 4 are for stress-free conditions along those edges. The thickness-wavelength ratio is 300.

FLEXURAL BIASING STRESSES
PARALLEL AND NORMAL
TO THE PROPAGATION DIRECTION
($\psi=0.0^\circ$)
($\theta=42.75^\circ$)
($(h/\lambda)=100.0$)

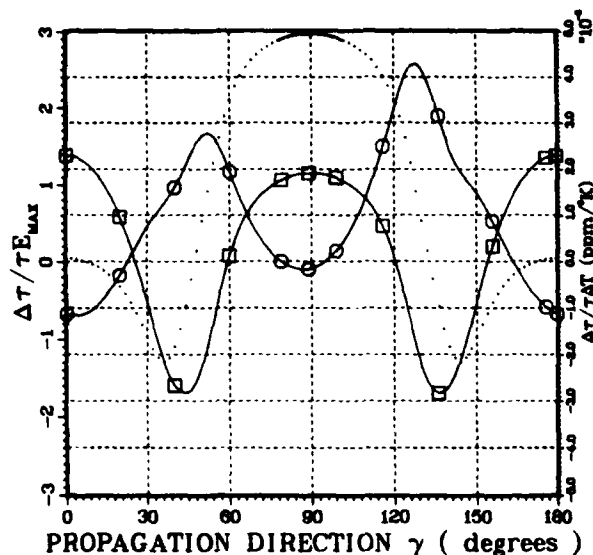


FIGURE 7 - Flexure and temperature induced fractional changes in time delay for surface waves on ST-cut quartz plate as a function of propagation directions ($h/\lambda=100$). The squares and circles represent curves for flexure along and normal to the propagation direction. Results are for stress-free conditions along edges normal to the direction of flexure.

with the force application jaws. Figure 9 shows the apparatus used for the application of edge forces. The picture shows a BAW plate and electrical leads attached to the left side. This portion has been removed and replaced with the aforementioned fixture for the SAW substrate. To the right of the photograph is the Ψ angle adjustment gearing. In the center of the picture is shown the force application portion of the apparatus. At bottom center, a micrometer head has been modified to provide a blunt knife-edge lower anvil. Aligned with this is the upper shaft with similar edge for contacting the crystal diametrically; the knife edges are parallel to the crystal plate thickness. A force of 100 gm weight is added to and removed from the upper shaft and the resulting frequency shift recorded; this eliminates the tare of the shaft and effects of friction and settling. The knife edges are coated with a thin plastic film to prevent scratching or chipping the quartz due to momentary high seating stresses. Because of the high temperature coefficient of Y cut quartz resonators (80-100 ppm/K), the entire apparatus was placed in a plastic shroud, and the measurements were taken in rapid succession, with the force on first, then with it abruptly removed. When taken in this manner it was observed that repeated readings for each psi value were consistent and no oven was required. Frequency shift measurements were made by using an HP Vector Voltmeter and frequency synthesizer. The phase was set to zero at the resonance frequency point when a 100 gm weight was applied to the crystal. The weight was then removed and the synthesizer used to bring the phase back to zero. This procedure was repeated between 5 and 10 times, and the frequency shifts averaged. Then the crystal was rotated to a new Ψ angle for further measurements. The relative frequency shift is the negative of the time delay variation introduced by the applied load, and the experimental values of fractional frequency shifts are compared with the theoretically computed values in Figs. 3, 4

FLEXURAL BIASING STRESSES
PARALLEL AND NORMAL
TO THE PROPAGATION DIRECTION
($\psi=0.0^\circ$)
($\theta=-49.2167^\circ$)
($(h/\lambda)=100.0$)

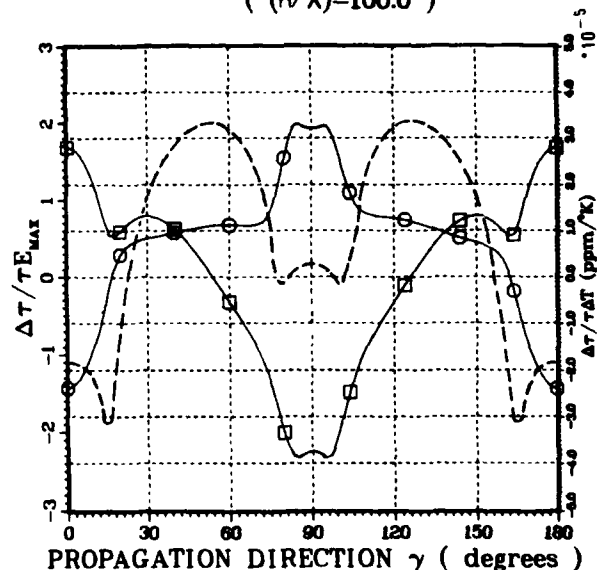


FIGURE 8 - Flexure and temperature induced fractional changes in time delay for surface waves on BT-cut quartz plate as a function of propagation direction ($h/\lambda=100$). The notation is same as in Figure 7.



FIGURE 9 - Force-Frequency Apparatus and Measurement Instrumentation.

and 5 for Y-cut, ST-cut and a doubly rotated cut quartz substrates, respectively.

The flexural deformation stress tests were conducted on 1.25" diameter ST-cut quartz discs. The resonators^{15,16} were centrally located with wave propagation along the X (digonal) axis. The frequency response of a typical device is shown in Figure 10, where we observe that three longitudinal resonances occur. For this design the transmission phase shift of the central resonance was zero degrees, while the phase shift for the other two modes was 180 degrees. The resonator was used to stabilize an

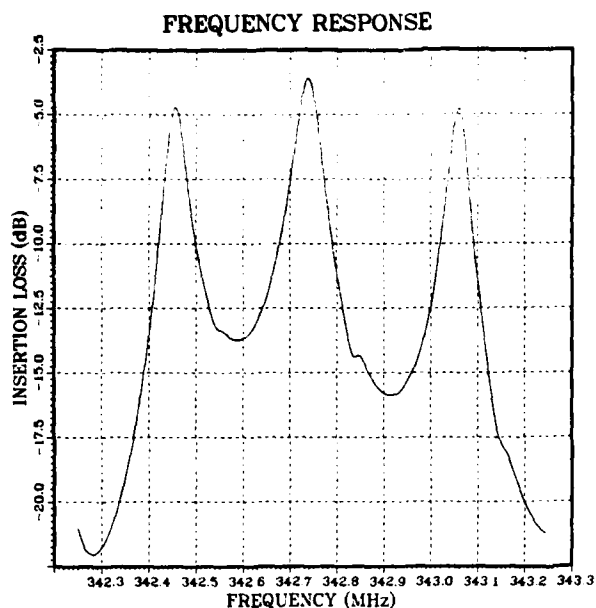


FIGURE 10 - Frequency response of a SAW resonator on ST-cut quartz disc employed in the study of moment-frequency effect.

oscillator which was designed to operate on the central resonance, and because of the 180 degree phase shift between modes there was no possibility of mode hopping. The circuitry consisted of a modular 15dB amplifier and a 10 dB directional coupler in a loop with the resonator. The effective cavity length of the resonator was 550 wavelengths, and the transducers were each 75 wavelengths long, yielding a fractional cavity metallization of about 25%. We endeavored to keep this fraction low in order not to stress (due to metallization) bias the substrate more than necessary. The frequency of oscillation was measured with no applied load, and for various flexural stresses applied using the apparatus shown in Figure 11. The resonator was housed in the disc-holder (lower left of Figure 11, which is shown in greater detail in Figure 12) and placed in the slot at the base of the column. The weight-rod (shown in the lower left of Fig. 11 with a 0.75" diameter ring) was used to load the substrate in the column with the calibrated weights shown. The resonator sensitivity to the applied stresses was determined from the slope of the frequency variation with applied load. The curves obtained for resonators utilizing both etched-groove and metal-line-segment reflectors were linear lending a high degree of



FIGURE 11 - Moment-Frequency Apparatus.



FIGURE 12 - Detailed view of the disc-holder for the flexure test.

confidence to our experimental results. The theoretically obtained value of stress sensitivity for the given substrate dimensions was .019 ppm/gm, and the experimental values were .020 and .025 ppm/gm. This agreement is considered good, and we plan further experiments on other material orientations.

In conclusion, we have obtained good agreement between theoretical and experimental results on force-frequency and moment-frequency effects in SAW discs of several orientations. The study has clearly demonstrated that such stress induced effects are strongly dependent on the type of loading, substrate geometry, orientation and propagation direction of surface waves.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. M. P. Ekstrom for support and encouragement; Mr. T. Muricchio for help with the calculations; Dr. E. J. Staples of Rockwell International for fabricating the SAW resonator patterns on Y-cut quartz; and Drs. J. Callerame and T. Parker of Raytheon for providing resonators on ST-cut discs for flexural tests.

REFERENCES

1. J. F. Dias, H. E. Karrer, J. A. Kusters, and C. A. Adams, "Frequency/Stress Sensitivity of SAW Resonators," *Elect. Lett.* 12 (22) 580-582 (1976).
2. D. F. Weirauch, R. J. Schwartz, and R. C. Bennett, "SAW Resonator Frit-Bonded Pressure Transducer," 1979 Ultrasonics Symposium Proceedings, 874-877 (1979).
3. D. Hauden, S. Rousseau, G. Jaillet, and R. Coquerel, "Pressure and Temperature Measurements with SAW Sensors," *Proc. 36th Ann. Symp. Freq. Control, USAERADCOM, Ft. Monmouth, N.J. 07703, May 1982, p. 284-289.*
4. J. F. Dias and H. E. Karrer, "Stress Effects in Acoustic Surface-Wave Circuits and Applications to Pressure and Force Transducer," 1974 IEEE International Solid-State Circuits Conference Proc. p. 166-167 (1974).

5. B. K. Sinha and H. F. Tiersten, "On the Influence of a Flexural Biasing State on the Velocity of Piezoelectric Surface Waves," *Wave Motion*, *1*, p. 37-51 (1979).
6. H. F. Tiersten, "Perturbation Theory for Linear Electroelastic Equations for Small Fields Superposed on a Bias," *J. Acoust. Soc. Am.*, *64*, p. 832-837 (1978).
7. H. F. Tiersten, B. K. Sinha and T. R. Meeker, "Intrinsic Stress in Thin Films Deposited on Anisotropic Substrates and its Influence on the Natural Frequencies of Piezoelectric Resonators," *J. Appl. Phys.*, *52*, p. 5614-5624 (1981).
8. B. K. Sinha, "Elastic Waves in Crystals Under a Bias," *Ferroelectrics*, *41*, p. 61-73 (1982).
9. D. Hauden, S. Rousseau, and J. J. Gagnepain, "Sensitivities of SAW Oscillators to Temperature, Forces and Pressure: Application to Sensors," *Proc. 34th Ann. Symp. Freq. Control*, USAERADCOM, Ft. Monmouth, N.J. 07703, May 1980, p. 312-319.
10. T. Lukaszek and A. Ballato, "What SAW Can Learn From BAW: Implications for Future Frequency Control, Selection and Signal Processing," *Proc. IEEE Ultrasonics Symposium*, p. 173-183 (1980).
11. P. Das, C. Lanzl, and H. F. Tiersten, "A Pressure Sensing Acoustic Surface Wave Resonator," *Proc. IEEE Ultrasonics Symposium*, p. 306-308 (1976).
12. E. J. Staples and J. Wise, "Surface Acoustic Wave Underwater Sound Sensors," *Proc. IEEE Ultrasonics Symposium Proc.*, p. 870-874 (1979).
13. B. K. Sinha and P. Das, "Determination of Thin Film Residual Stresses Using SAW Oscillators," *Ultrasonics International 81 Proceedings*, (IPC Science and Technology), p. 366-371 (1981).
14. T. M. Reeder and D. E. Cullen, "Surface Acoustic Wave Pressure and Temperature Sensors," *Proc. IEEE*, *64*, p. 754-756 (1976).
15. W. J. Tanski, "Surface Acoustic Wave Resonators on Quartz," *IEEE Trans. on Sonics and Ultrasonics*, SU-26, p. 93-104, (1979).
16. D. T. Bell and R. C. M. Li, "Surface Acoustic Wave Resonators," *Proc. IEEE*, *64*, p. 711-721, (1976).



SURFACE WAVE RESONATORS ON SILICON

S. J. Martin, R. L. Gunshor, T. J. Miller, S. Datta, R. F. Pierret
School of Electrical Engineering, Purdue University, West Lafayette, IN 47907

and

M. R. Melloch
Central Research Labs, Texas Instruments Inc., Dallas, TX 75265

INTRODUCTION

Over the past few years a development effort has been undertaken to adapt the surface acoustic wave (SAW) resonator to a silicon based configuration. The motivation for using silicon as the substrate lies in the possibility of incorporating SAW resonators in monolithic integrated circuits. The ZnO-on-Si two port resonator configuration, shown in Fig. 1, utilizes a thermally oxidized silicon substrate on which a piezoelectric ZnO layer is deposited to permit surface wave excitation. Surface features include aluminum interdigital transducers as well as distributed reflector arrays which are formed from grooves ion beam etched in the ZnO layer. An aluminum underlayer is deposited prior to ZnO sputtering. This underlayer serves to enhance the electromechanical coupling, in addition to isolating charge carriers in the semiconductor substrate from electric fields originating in the ZnO layer.

Two port SAW resonators fabricated on silicon have achieved Q-values of 12,000 at 100 MHz. Reflector arrays, formed by modulating the ZnO layer thickness with etched grooves, have been characterized by a linear change in reflectivity and wave velocity with groove depth¹. Rejection levels of 30 dB have been obtained in devices with an

untuned resonant insertion loss of 9 dB. An attractive feature results from the fact that the thermal expansion coefficient of SiO₂ is opposite in sign to that of Si and ZnO; it is possible to use the thermal oxide layer to temperature compensate the layered acoustic medium. The temperature stability of ZnO-on-Si SAW resonators has been shown to be comparable to those fabricated on ST quartz when the SiO₂ thickness is properly chosen². To complement the temperature stability, a preliminary ageing study of hermetically packaged resonators has indicated that ageing rates of less than 5 ppm per month are obtainable.

→ In this paper ~~we will discuss~~ two topics new to the ZnO-on-Si resonator development:

- 1) The use of SAW resonators to determine the effect of a laser anneal on layered medium propagation loss;
- 2) A conceptually new device, called the mode conversion resonator, which utilizes two propagating normal modes of the layered medium in order to gain enhanced out-of-band signal rejection. ←

LASER ANNEALING OF THE LAYERED MEDIUM

A fundamental limitation on the Q-value achievable by a SAW resonator is set by the surface wave propagation loss. In the ZnO-SiO₂-Si layered medium, propagation loss is believed to dominate in the polycrystalline ZnO layer as evidenced by an increase in attenuation with ZnO thickness. A correlation has been found between the density of defects in the sputtered ZnO film and the propagation loss observed³. These defects, while much smaller than an acoustic wavelength, apparently act as Rayleigh wave scattering centers.

By using a ZnO layer on an oxidized Si substrate as an optical waveguide, Hickernell has suggested that optical propagation loss is dominated by defects located at the ZnO-SiO₂ interface. It was conjectured that these interface defects, arising during the initial stages of ZnO sputter-deposition, may also contribute significantly to surface wave attenuation. Attempts were made to reduce defects in the ZnO film through furnace annealing treatments. A decrease in tensile stress, as well as increased crystallographic ordering

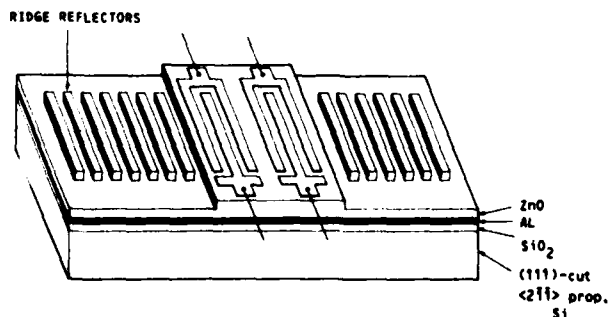


Fig. 1. SAW two-port resonator configuration using a piezoelectric ZnO layer on a thermally oxidized Si substrate.

were observed in films annealed at 400° C⁴. However, no change was obtained in either optical or acoustic propagation loss.

On the other hand, annealing methods which provide localized heating at the ZnO-SiO₂ interface at a level which reduces defects were found to be successful in decreasing optical propagation loss. By exposing the substrate side of ZnO-on-Si optical waveguides to the radiant energy of a strip heater, in a process called "rapid isothermal annealing," Hickernell was able to reduce optical propagation loss by a factor of two⁵.

An alternative means of providing localized heating was to use laser illumination in the infra-red region of the spectrum. At a laser wavelength of 10.6 μm ZnO and Si are nearly transparent, while the SiO₂ layer absorbs strongly. As shown by the laser absorption profile in Fig. 2, front side IR illumination of the layered medium provides localized heating at the ZnO-SiO₂ interface, precisely at the location where the defect density is believed highest. Dutta *et al.*⁶ found that by laser scanning optical waveguides at power densities of approximately $2 \times 10^5 \text{ W/cm}^2$ at a scan rate of 1 cm/sec, one could achieve a reduction in optical propagation loss of one to two orders of magnitude.

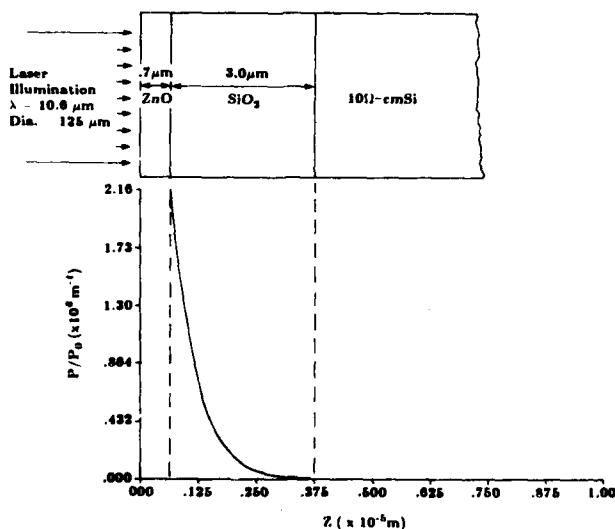


Fig. 2. Absorption profile of the incident CO₂ laser energy in the ZnO-SiO₂-Si layered medium. Defects at the ZnO-SiO₂ interface are reduced by the localized heating.

In both the rapid isothermal annealing experiment and the laser annealing experiment outlined above, experiments in which significant reductions in optical propagation loss were found, the experimenters did not determine the effect on acoustic propagation loss. It was our intention to use resonators fabricated on the ZnO-SiO₂-Si layered medium as a sensitive means to detect changes in acoustic loss caused by the laser anneal. The resonators employed in the experiment were fabricated on (111) Si, using a 3 μm SiO₂ layer

and 0.75 μm ZnO layer. The 3 μm SiO₂ layer provides sufficient electrical isolation of the substrate from the ZnO electric fields so that an aluminum underlayer is not required for this purpose. In addition to eliminating transducer losses from the cavity, placing the transducers outside the resonant cavity enables one to eliminate aluminum from the region to be scanned. (Aluminum has proven detrimental in high temperature steps, diffusing into the ZnO layer and increasing the bulk conductivity.) By minimizing all sources of loss, these externally coupled resonators exhibited Q values of up to 14,000 at 100 MHz.

A CO₂ laser was used to scan the surface of each externally coupled resonator. A scanning apparatus containing an xy-translator was constructed to enable the stationary laser beam to raster-scan the surface of the device at a rate of 1 cm/sec. The scanner was capable of accommodating the device mounting case so that center frequency f_c and Q could be measured immediately before and after the laser scan at each power level. Placing the device 39 cm behind the lens (having a focal length of 40 cm) produced a spot size with a diameter calculated to be 0.125 mm.

The variation in resonant frequency and Q-value found after the laser anneal of a single device at successively higher laser power densities is shown in Fig. 3. We note a 3% maximum increase in Q-value after annealing at a laser power

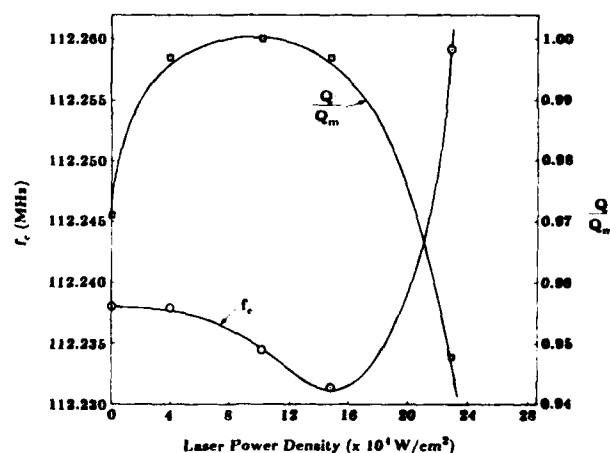


Fig. 3. Variation in resonant frequency and Q-value after laser annealing a single externally coupled resonator at successively higher power densities.

density of $9 \times 10^4 \text{ W/cm}^2$. This power level is 42% of the value reported for optimum optical loss reduction. As our devices used a greater thickness of SiO₂ (3 μm as opposed to 1 μm), a greater efficiency is expected for localized heating. Thus the slight improvement in Q-value is attributed to a decrease in surface wave propagation loss through a reduction in interface defects. At higher laser power densities we note a rolloff in Q-value and a marked increase in resonant frequency. Microscopic examination of the device surface

revealed microfractures formed during the anneal. The accumulation of microfractures in successive scans at laser power densities above 10^5 W/cm^2 seems to cause wave scattering leading to Q-degradation. This observed effect of a laser anneal at excessive power levels might have been anticipated, since Hickernell has reported bulk recrystallization and micro-fracture formation in ZnO films subjected to furnace anneals at temperatures between 500-700° C.

At the optimum laser annealing power density, the decrease in acoustic propagation loss is less dramatic than that observed for optical propagation loss by Dutta *et al.* It is conjectured that optical waveguiding, which relies on reflection at the upper and lower ZnO surfaces to confine light energy, may be more sensitive to defects at the ZnO-SiO₂ interface than is SAW propagation. SAW devices operating at higher frequencies and having a greater proportion of acoustic energy in the vicinity of the ZnO-SiO₂ interface may be more sensitive to interface defects. Such devices may benefit more from a laser anneal than those tested. Alternatively, propagation loss may not be the dominant source of loss in the resonant cavity. If other sources dominate losses, the improvement in propagation characteristics would not be fully reflected in the Q-value increase.

MODE CONVERSION RESONATOR

The second new development in ZnO-on-Si resonators which we address here is the operation of the mode conversion resonator. The devices discussed previously are constructed on a layered medium having a ZnO layer which is made as thin as possible in order to minimize surface wave attenuation. Consequently, only the lowest order sagittal mode (the Rayleigh mode) will propagate, as is the case at the surface of a semi-infinite homogeneous medium. When the ZnO layer is made sufficiently thick in relation to acoustic wavelength, it is possible to excite a second order sagittal mode or Sezawa mode.

In Fig. 4 the solid lines indicate the dispersion characteristics for Rayleigh and Sezawa modes propagating in a forward direction in a layered medium designed to support both modes. By using a periodic array of metal strips or grooves, it is possible to couple these forward modes into backward modes (indicated by dashed lines in Fig. 4) at points where the forward and backward branches intersect. The dispersion characteristics indicate the possibility of constructing a mode-converting reflector array. At ω_{RS} an incident Rayleigh mode R_+ is coherently scattered into a backward Sezawa mode S_- , and *vice versa*. The wavenumbers of the Rayleigh and Sezawa waves which participate in this reflection between modes must satisfy the relation

$$k_R + k_S = \frac{2\pi}{d} \quad (1)$$

in which d is the mode conversion reflector periodicity. A resonant cavity may be formed between two mode-converting reflector arrays. A Rayleigh wave traversing the resonant cavity in one direction will be reflected as a Sezawa

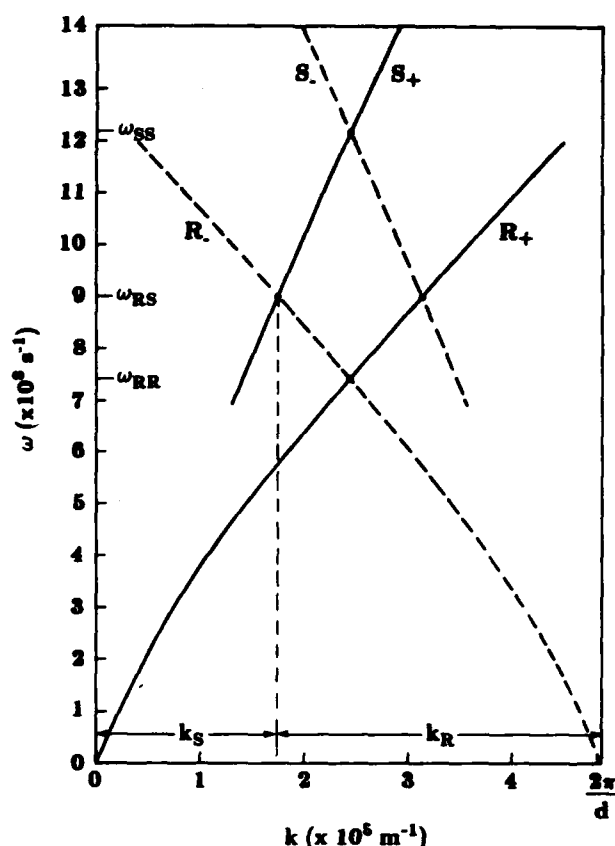


Fig. 4. Dispersion diagram illustrating Rayleigh and Sezawa mode propagation characteristics in a layered medium consisting of 6.5 μm ZnO on 1.0 μm SiO₂ on (100)-cut, <010>-propagating Si.

wave at the same frequency, provided Eq. 1 is satisfied. After traversing the resonant cavity in the opposite direction as a Sezawa wave, it will be re-reflected as a Rayleigh wave. A resonance condition is satisfied whenever the round trip phase shift is a multiple of 2π , i.e.,

$$(k_R + k_S)L_{\text{eff}} = 2n\pi \quad (2)$$

where L_{eff} is the effective cavity length of the device and n is the longitudinal mode index. The transducers forming input and output ports are designed to couple energy into the resonant cavity in one propagating mode and out of the cavity from the other. As indicated in Fig. 5, one transducer having periodicity λ_R is positioned to couple optimally to the Rayleigh wave, while a second transducer, having periodicity λ_S , is placed to couple optimally to the Sezawa wave. The difference in transducer periodicities ensures a significant reduction in direct acoustic coupling between transducers, thereby increasing the rejection level of the resonator. In designing transducers to couple to each mode, it is necessary to have appreciable electromechanical coupling to both modes with a given ZnO thickness. This condition is best achieved by placing transducers at the ZnO-SiO₂ interface,

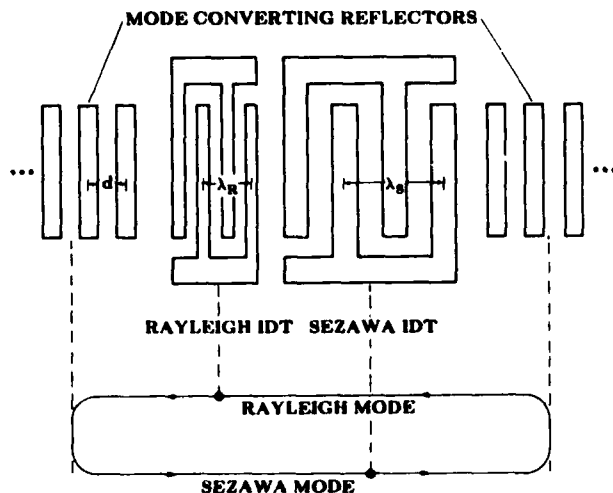


Fig. 5. Schematic of SAW mode conversion resonator. Input and output ports are tuned to different propagating modes, with distributed arrays coherently scattering between modes.

and employing a ground plane on top of the ZnO. Using a ZnO thickness of $6.5 \mu\text{m}$ and an SiO_2 thickness of $1.0 \mu\text{m}$ on (100)-cut $\langle 010 \rangle$ -propagating Si, the following values of electromechanical coupling were calculated: $\Delta_R = 0.016$, $\Delta_S = 0.024$.

In order to use a single masking step to simultaneously define IDT's and mode conversion reflector arrays, the array grooves were etched to a depth of $0.1 \mu\text{m}$ in the SiO_2 layer prior to the deposition of ZnO. The frequency responses for

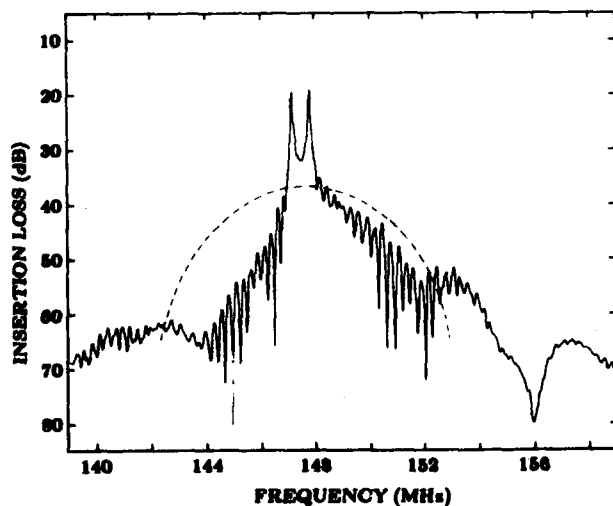


Fig. 6. Frequency response of a two-port mode conversion resonator. The calculated direct coupling level between transducers tuned to a single propagating mode is indicated by the dashed curve.

two prototype mode conversion resonators are shown in Fig. 6 and Fig. 7. The device whose characteristics are shown in Fig. 6 exhibits two resonances. The large separation between mode converting arrays ($52\lambda_R$ or $29\lambda_S$), along with the detuning of resonance from the mode-conversion center frequency, causes Eq. 2 to be satisfied twice within the mode conversion bandwidth of the arrays. The out-of-band acoustic coupling level is seen to fall off more rapidly than that calculated for transducers tuned to the same propagating mode. Employing a smaller separation between mode-converting arrays ($32\lambda_R$ or $18\lambda_S$), the device whose characteristics are shown in Fig. 7 exhibits a single resonance peak. Due to the smaller number of transducer fingers, however, the rolloff in out-of-band acoustic coupling between transducers is less dramatic in this case.

Mode conversion resonator Q-values ranging from 2300 to 3000 have been obtained, indicating that power conversion efficiency between Rayleigh and Sezawa modes can exceed 90% with a 400-period array. In fact, device Q is thought to be limited primarily by propagation loss occurring in the $6.5 \mu\text{m}$ -thick ZnO film.

ACKNOWLEDGEMENT

This work was sponsored jointly by the Air Force Office of Scientific Research under grant AFOSR810214 and by the National Science Foundation under grant 8020249.

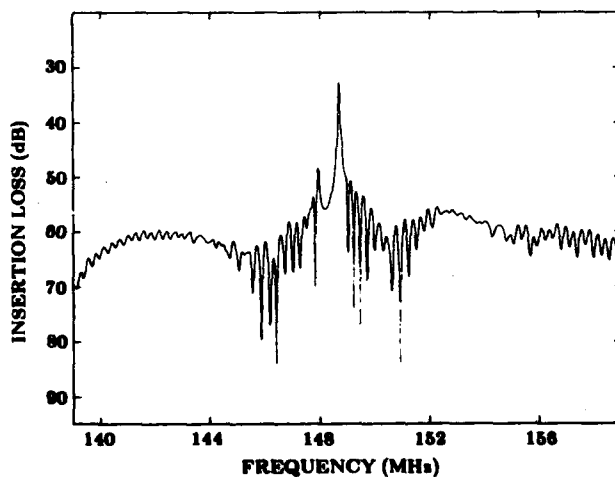


Fig. 7. Frequency response of a two-port mode conversion resonator exhibiting a single resonance peak.

REFERENCES

1. S. J. Martin, S. S. Schwartz, R. L. Gunshor, and R. F. Pierret, "Surface Acoustic Wave Resonators on a ZnO-on-Si Layered Medium," *J. Appl. Phys.*, **54**, 561 (1983).
2. S. J. Martin, R. L. Gunshor, and R. F. Pierret, in *1980 Ultrasonics Symposium Proceedings*, (IEEE, New York, 1980) 113.
3. F. S. Hickernell, "Surface Acoustic Wave Propagation Loss in Zinc Oxide Films," in *1982 Ultrasonics Symposium Proceedings*, (IEEE, New York, 1982).
4. Ibid.
5. F. S. Hickernell, "Post Deposition Annealing of Zinc Oxide Films," in *1982 Ultrasonics Symposium Proceedings*, (IEEE, New York, 1981) 489.
6. S. Dutta, H. E. Jackson, J. T. Boyd, F. S. Hickernell, and R. L. Davis, "Scattering Loss Reduction in ZnO Optical Waveguides by Laser Annealing," *Appl. Phys. Lett.*, **39**, 206 (1981).



HIGH-FREQUENCY SHALLOW BULK ACOUSTIC WAVE QUARTZ FREQUENCY SOURCES

K. V. Rousseau, K. H. Yen, K. F. Lau, and A. M. Kong

TRW Electronic Systems Group
Redondo Beach, California

Summary

This paper describes the design and fabrication of high-frequency shallow bulk acoustic wave (SBAW) delay lines on quartz. Direct generation of frequency tones in the GHz range is desirable because it avoids the phase noise increase inherent in the multiplication of low frequency bulk-wave sources. The close-in phase noise can then be improved if desired by locking the GHz source with a bulk crystal. The devices reported in this paper used crystal cuts near AT because of its high velocity and BT because of its excellent temperature stability.

Design parameters are chosen using an in-line model. This model uses the aperture width, number of fingers, wave velocity and coupling constant to predict radiation resistance, capacitance, and conversion loss. When used in conjunction with experimental propagation loss data, this model accurately predicts the total insertion loss of the device. Recent refinements of this model indicate that, at GHz frequencies, a significant trade-off exists between insertion loss and Q.

Electron-beam lithography is used to form 0.4 μm patterns in a PMMA resist. Resists are typically single-layer, but more exotic schemes (several layers) have also been examined. Criteria have been developed to decide whether or not the metal transducer fingers should be embedded in the quartz. Usually the fingers are not embedded, and the devices still achieve low insertion loss, high Q, and excellent temperature stability. Fundamental and fifth-harmonic devices on AT quartz and fifth-harmonic devices on BT quartz were fabricated using the lift-off method. Frequencies range from 3 to 5 GHz, with the best devices having insertion losses in the low twenties. Q-values greater than 2000 have been observed.

Low-feedthrough packaging has also been devised using a TO-8 can that is cold-welded under high vacuum. Organics are not used; the device is held in place with a stainless-steel clip that is relatively stress-free.

Phase noise and preliminary aging data are also presented.

Introduction

Shallow bulk acoustic waves (SBAW) in quartz are preferred for high frequency applications because of their high wave velocity, lower propagation loss, and small temperature coefficient of delay.^{1,2} Previous papers³⁻⁵ have shown the practicality of direct frequency generation at a few gigahertz. This paper discusses the design considerations and fabrication techniques that yield oscillator delay lines over 3 GHz with high Q and good temperature stability.

Design Considerations

Material properties of +35.5° (near AT) and -50.5° (near BT) rotated Y-cut quartz are presented in Table 1. The AT cut has much higher velocity and coupling coefficient, while the BT quartz is less sensitive to mass

Table 1. Comparison of SBAW Properties in +35.5° and -50.5° Rotated Y-Cut Quartz

	SUBSTRATE ANGLE	
	+35.5° (AT)	-50.5° (BT)
WAVE VELOCITY	5100 M/sec	3331 M/sec
COUPLING COEFFICIENT	1.44×10^{-2}	0.4×10^{-2}
MASS LOADING EFFECT FOR $H/\lambda = 0.01$	$\Delta v/v = 0.16\%$	$\Delta v/v = 0.01\%$
WAVE ATTENUATION AT 1 GHz	0.87 dB/ μsec	2.0 dB/ μsec
UNIFORM ISOTROPIC STATIC STRESS ($\times 10^{-12} \text{ M}^2/\text{N}$)	2	-7
TEMPERATURE STABILITY (-55°C TO +85°C)	$\pm 127 \text{ ppm}$	$\pm 55 \text{ ppm}$

loading and has better temperature stability. Thus, one might choose to fabricate delay lines on either cut, depending upon the application. For a given lithographic constraint on linewidth, AT would be selected to produce higher frequency devices, but applications where a wide temperature variation is expected would probably be better satisfied using BT quartz.

One can use a multi-electrode configuration to extend the frequency of operation on either species of quartz.⁶ Utilizing the fifth harmonic response increases the frequency of operation significantly, so devices on BT quartz are typically designed with that configuration. We also used electron beam lithography in order to extend the frequency of operation beyond that accessible with conventional photolithography.

Electrodes are typically embedded to reduce reflections. On BT quartz, however, the SBAW velocity is so close to that of aluminum that this is a relatively small effect. On AT quartz, the average difference in fundamental mode insertion loss between embedded and non-embedded devices was found⁵ to be about 1 dB, but is significantly higher for harmonic-mode devices--on the order of 3-5 dB at gigahertz frequencies.

Propagation loss must also be taken into account. As discussed elsewhere,^{5,7} we empirically model the propagation loss on AT quartz as

$$\text{Loss (in dB)} = 9.0 \text{ dB}/\mu\text{sec} \cdot \nu^2, \quad (1)$$

where ν is the frequency in gigahertz. Thus, even as a device's dimensions (typically) scale inversely with frequency, the ν^2 growth of the propagation loss ensures greater insertion loss at higher frequencies. This is incorporated into our standard in-line model for SBAW insertion loss, enabling us to predict the loss of a given design with reasonable accuracy. Figure 1 shows a plot of expected insertion loss as

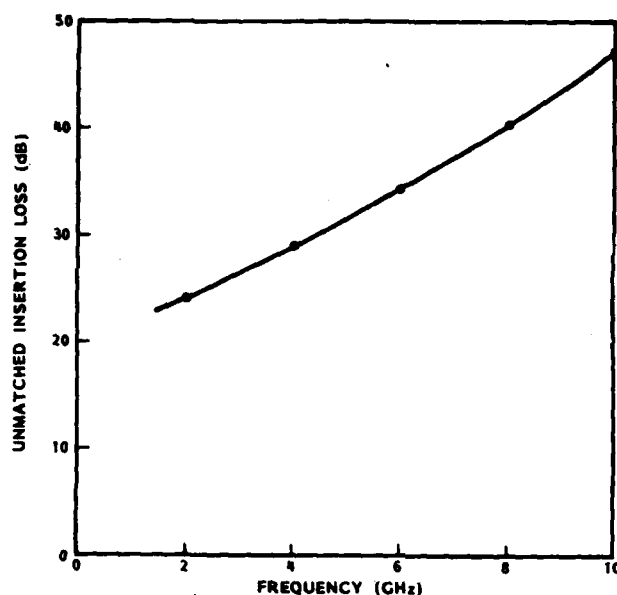


Figure 1. Scaling of Insertion Loss as a Function of Frequency

a function of frequency for a scaled, fundamental mode device on AT quartz with an aperture of 75λ , identical transducers of 200 finger pairs each, and a gap of $50 \mu\text{m}$ between transducers. The increase in insertion loss is rapid, and almost linear in dB. Also, the Q of these devices is relatively low--ranging from 750 to 930. In oscillator applications, one typically wants the Q as high as possible, which implies a long delay time, which in turn increases the propagation loss. At 2 GHz, propagation loss is about 10% of the total loss in dB; at 10 GHz, it is over one-half. Therefore, it should be clear that a fundamental trade-off exists between insertion loss and Q at such frequencies.

Fabrication

The devices reported here are all fabricated with electron beam lithography. A thin (100 \AA) layer of metal is evaporated onto the quartz to serve as an adhesion/conduction layer, then a layer of PMMA electron resist is spun on and baked. Exposure is accomplished with a Cambridge EBMF-2 system. The PMMA is then developed, which is followed by a plasma de-scum step, metal deposition and lift-off. The electrodes are not embedded.

A multi-level resist scheme has been devised that would enable us to embed the electrodes in the quartz. The beginning configuration is shown in Figure 2. Patterns exposed in the PMMA are replicated in the polychrome resist, which withstands ion-milling better than the PMMA. The quartz is then etched with reactive ion-etching, then metal is deposited and lifted off as usual. This method is still under development.

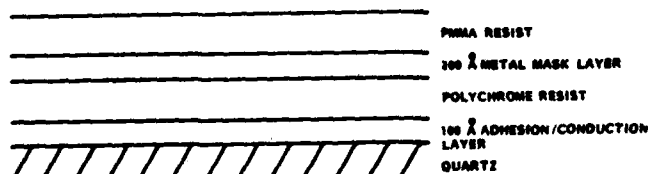


Figure 2. Multi-Level Resist

After the patterns are produced, the wafers are diced and packaged in TO-8 cans. Devices are held by a stainless-steel clip that is spot-welded to the package base. This clip places very little stress on the substrate. The packages are then cold-welded in a vacuum that is typically 10^{-8} or 10^{-9} torr. It is believed this packaging approach will minimize long-term aging.

Delay Line Results

We report results of four devices here--two 3.1 GHz AT fundamental mode devices (Devices A and B), a 3.5 GHz fifth harmonic BT device (Device C), and a 5.25 GHz fifth-harmonic AT device (Device D). Design parameters are summarized in Table 2.

Table 2. Device Design Parameters

	DEVICE A	DEVICE B	DEVICE C	DEVICE D
CRYSTAL	AT	AT	BT	AT
FINGERWIDTH/ GAP WIDTH	9.4 μm	9.4 μm	9.4 μm	9.4 μm
APERTURE WIDTH	81 λ_1	81 λ_1	50 λ_5	75 λ_5
HARMONIC MODE	FUNDAMENTAL	FUNDAMENTAL	FIFTH	FIFTH
CENTER FREQUENCY	3.1 GHz	3.1 GHz	3.5 GHz	5.25 GHz
CENTER-TO-CENTER TRANSDUCER SEPARATION	210.4 μm	210.4 μm	868 μm	210.4 μm
METALLIZATION	400 \AA AL	400 \AA AL	400 \AA AL	400 \AA AL
TRANSDUCER LENGTH	100 λ_1	250 λ_1	898 λ_5	165 λ_5

Figure 3 shows device A, a 3.1 GHz AT fundamental device, unmatched. Before packaging in the TO-8 can, the insertion loss was 22 dB. Electromagnetic feed-through changes the packaged loss to 20 dB, and also causes the high-frequency "tail" to appear. Figure 4 shows device B unmatched and unpackaged. It appears as if a stopband/resonator effect is occurring.

A loss analysis of these devices is given in Table 3. The measured insertion loss of device A does not agree very well with the predicted untuned loss. This is because the inductance of the several bond wires used partially tuned the device--this was observed on a Smith chart. Device B had better agreement, but was still off. We believe this is due to the resonator effect one can observe from the plot. Alternatively, perhaps our model overestimates the propagation loss.

Table 3. Loss Analysis for 3.1 GHz AT Devices (in dB)

DEVICE LOSS	DEVICE 1		DEVICE 2	
	UNTUNED	TUNED	UNTUNED	TUNED
CONVERSION LOSS (1)	11.2	6.6	9.4	8.2
CONVERSION LOSS (2)	11.2	6.6	9.4	8.2
RESISTIVE LOSS	0.5	0.5	0.5	0.5
SPREADING LOSS	1.7	1.7	1.0	1.0
PROPAGATION LOSS	3.3	3.3	6.4	6.4
TOTAL LOSS	27.9	18.7	27.9	25.0
MEASURED	22.0	---	24.5	---

Device C, shown in Figure 5, is a fifth-harmonic device on BT quartz that operates at 3.5 GHz. The passband shape and rejection are quite good, but insertion loss is rather high. However, the measured insertion loss agrees quite nicely with the predicted untuned loss, shown in Table 4. When tuned, this

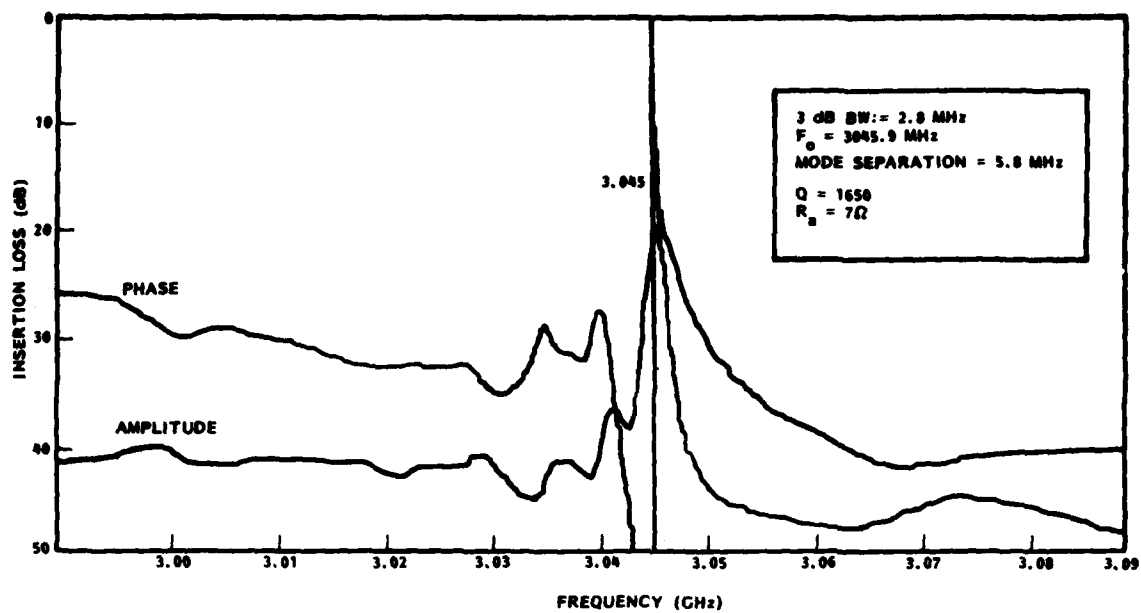


Figure 3. 3.1 GHz AT Delay Line (Device A)

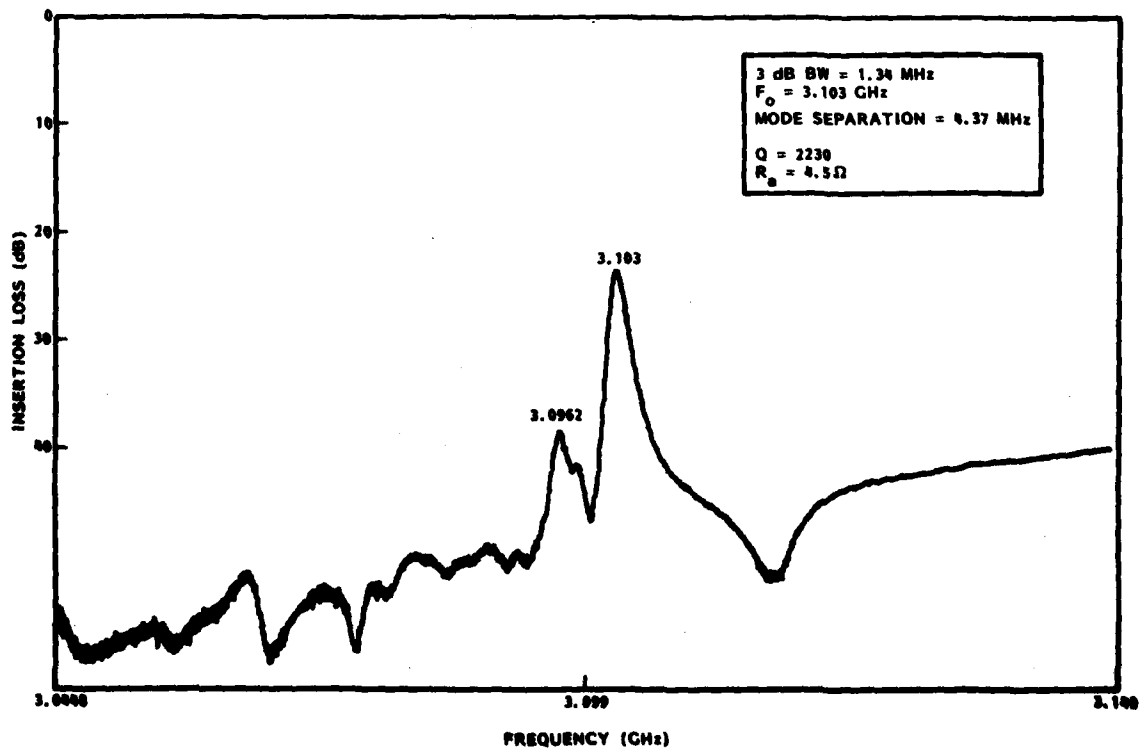


Figure 4. 3.1 GHz AT Delay Line (Device B)

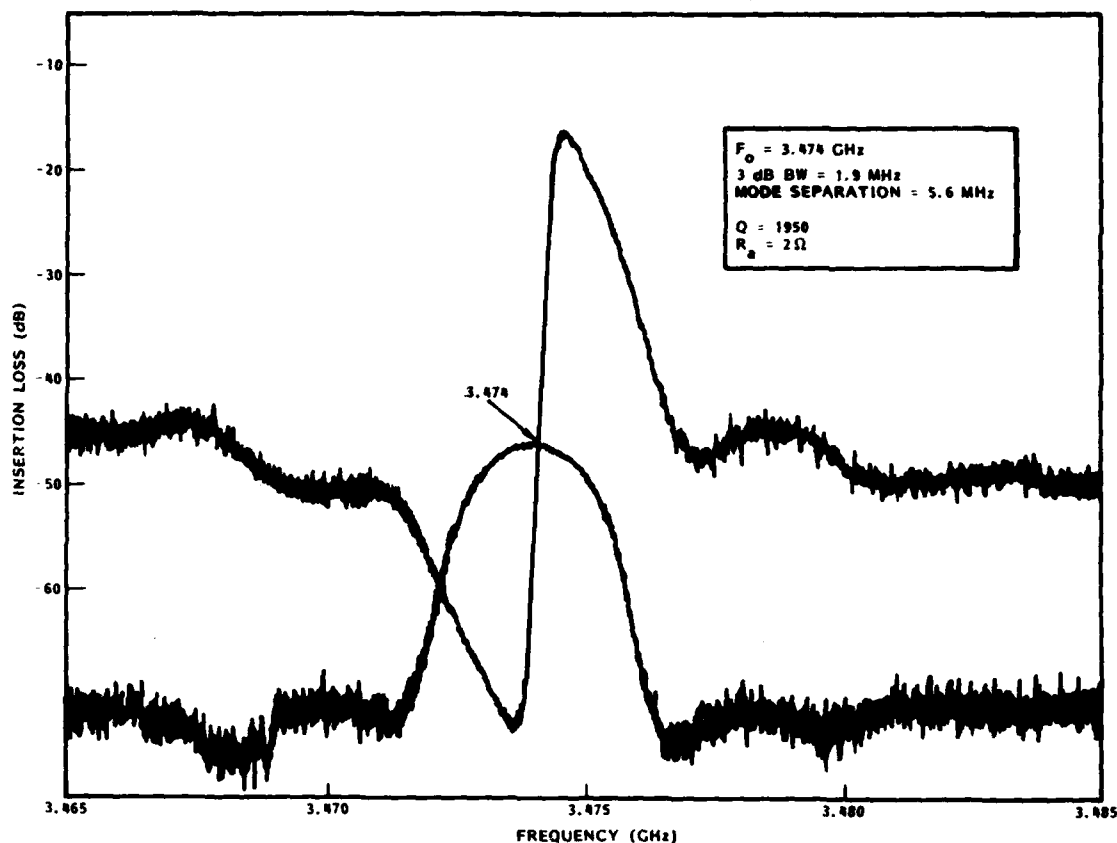


Figure 5. 3.5 GHz BT Delay Line (Device C)

Table 4. Loss Analysis for 3.5 GHz BT Device (in dB)

DEVICE LOSS	UNTUNED	TUNED
CONVERSION LOSS (1)	16.1	8.7
CONVERSION LOSS (2)	16.1	8.7
RESISTIVE LOSS	0.5	0.5
SPREADING LOSS	.8	.8
PROPAGATION LOSS	11.2	11.2
TOTAL LOSS	44.8	29.9
MEASURED LOSS	46.0	36.0

device reached an insertion loss of 36 dB, but the Q was reduced about 25%.

Device D, a 5.25 GHz, fifth-harmonic device, was fabricated with parameters previously shown. Predicted insertion loss was 46 dB (shown in Table 5), and a response at about 50 dB was barely observable due to feedthrough at the same level. The discrepancy is due to the several dB increase in insertion loss (mentioned earlier) caused by not embedding the electrodes of a harmonic AT device.

Oscillator Design and Results

As shown schematically in Figure 6, the SBAW oscillator consists of a SBAW delay line and an

Table 5. Loss Analysis for 5.25 GHz AT Device (in dB)

DEVICE LOSS	UNTUNED	TUNED
CONVERSION LOSS (1)	17.2	4.7
CONVERSION LOSS (2)	17.2	4.7
RESISTIVE LOSS	0.5	0.5
SPREADING LOSS	1.7	1.7
PROPAGATION LOSS	9.6	9.6
TOTAL LOSS	46.1	21.3
MEASURED LOSS	50.0	-----

amplifier in a feedback loop circuit. The conditions for oscillations are (1) gain around the loop must exceed all losses, and (2) phase shift around the loop must be equal to a multiple of 2π radians. For single mode operation, the SBAW delay line must have a sufficiently narrow passband so that it contains only one oscillating mode. All other modes satisfying condition (2) will lie in the rejection band of the delay line and will not oscillate due to lack of gain.

Discrete components are used in the construction of the oscillators. This allows maximum flexibility in interchanging oscillator components and thus permits

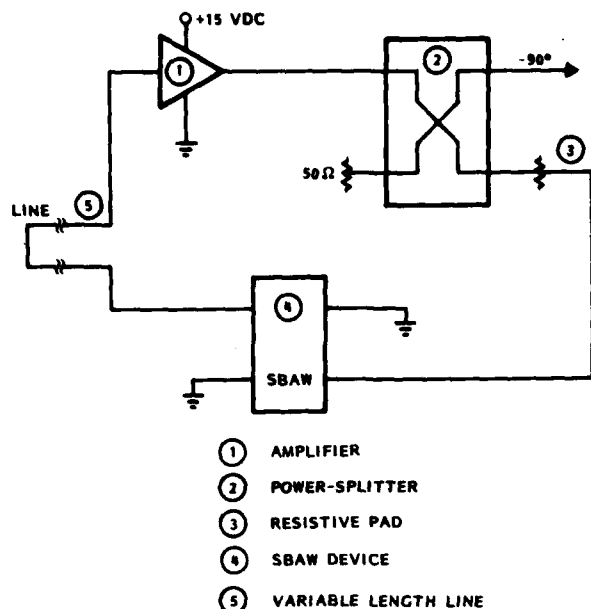


Figure 6. SBAW Oscillator Schematic

many delay lines to be incorporated and evaluated. Each oscillator consists of a SBAW delay line, amplifier, attenuator, some external delay, and an output coupler. Typically, a wideband (2-4 GHz) GaAs FET amplifier is used.

Figures 7 and 8 show phase noise for oscillators using devices A and B, respectively. The straight lines on each plot represent predicted phase noise of a 10 MHz commercially available bulk crystal⁸ oscillator multiplied to 3.1 GHz. The floor of the SBAW oscillators is lower, but close-in phase noise is worse than that of the bulk crystal. A circuit combining the two would be optimal if weight and power penalties could be afforded.

Aging

The two oscillators for which phase noise plots were given recently started long-term aging tests. Only one month of data is available, so any conclusions are extremely tentative. One has shifted about 10 ppm and appeared to level off, and the other has shifted less than 2 ppm to date.

Conclusions

Oscillators with good phase noise characteristics can be fabricated using direct generation of high-frequency tones. There is a trade-off between Q and insertion loss, but one might reasonably expect to produce devices with Q's of around 2000 and an untuned insertion loss in the low twenties. In addition, aging results are encouraging, but not yet conclusive.

Acknowledgements

The authors wish to thank Drs. J. L. Rogers and P. H. Liu for the electron beam exposure, S. T. Lo for the remaining device fabrication steps, and W. T. Akiyama for oscillator design and phase noise measurements.

This work was partially supported by ERADCOM Contract DAAK20-81-C-0415.

References

1. Lewis, M., "Surface Skimming Bulk Waves," Proc. IEEE Ultrasonics Symposium, October 1977, pp. 744-752.
2. Yen, K. H., K. F. Lau, and R. S. Kagiwada, "Recent Advances in Shallow Bulk Acoustic Waves Devices," Proc. IEEE Ultrasonics Symposium, September 1979, pp. 776-785.
3. Lau, K. F., K. H. Yen, R. S. Kagiwada and A. M. Kong, "A Temperature Stable 2 GHz SBAW Delay Line Oscillator," Proc. 34th Annual Frequency Control Symposium, U. S. Army Electronics Technology and Device Laboratory, Fort Monmouth, NJ, May 1980.
4. Lau, K. F., K. H. Yen, R. S. Kagiwada and A. M. Kong, "High Frequency Temperature Stable SBAW Oscillators," Proc. IEEE Ultrasonics Symposium, October 1980, pp. 240-244.
5. Rousseau, K. V., K. H. Yen, K. F. Lau, and A. M. Kong, "High Q, Single Mode S-Band SBAW Oscillators," Proc. IEEE Ultrasonics Symposium, October 1982, pp. 279-283.
6. Engan, H., "Surface Acoustic Wave Multielectrode Transducers," IEEE Trans., SU-22, 1975, pp. 395-401.
7. Weglein, R. D., and O. W. Otto, "Microwave SAW Oscillators," Proceedings, IEEE 1977 Ultrasonics Symposium, pp. 913-922.
8. Hewlett-Packard 10 MHz Bulk Crystal.

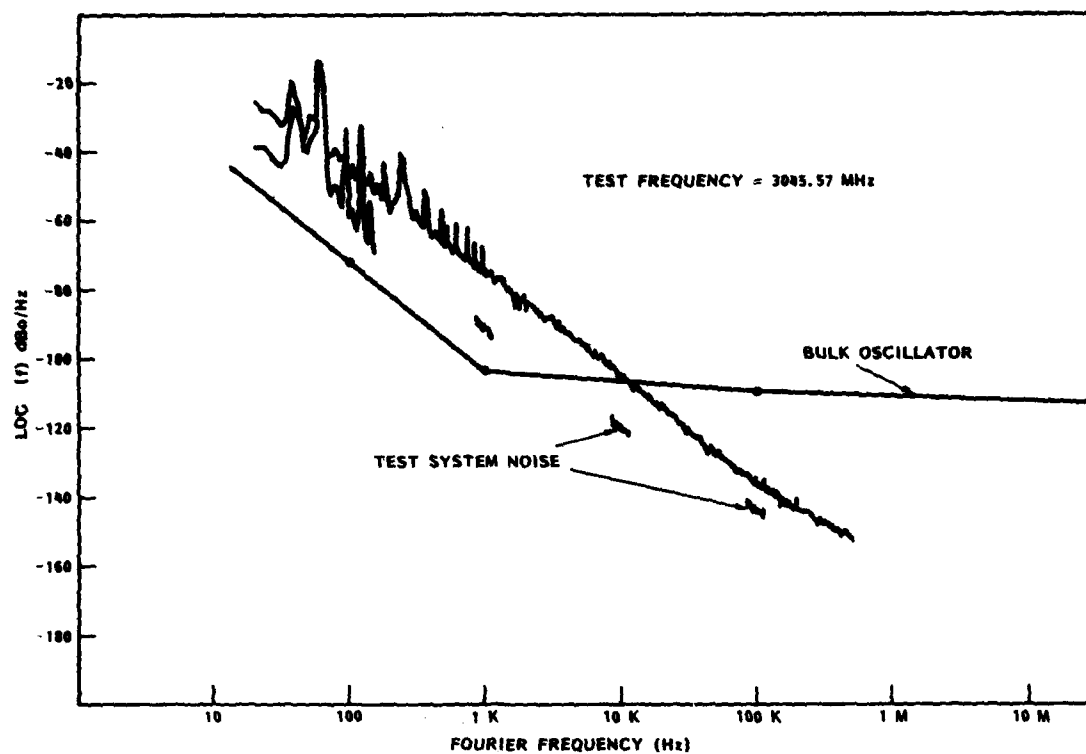


Figure 7. Single-Sideband Phase Noise of Oscillator Using Device A

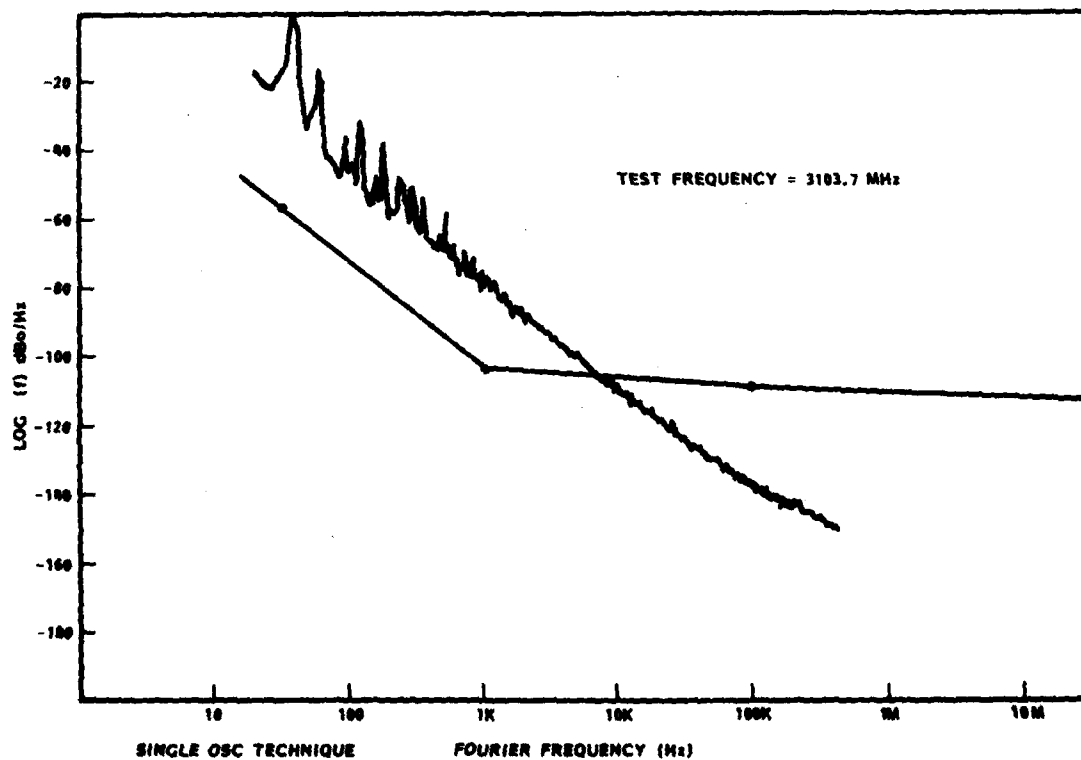


Figure 8. Single-Sideband Phase Noise of Oscillator Using Device B

A New Digital TCXO Circuit Using a Capacitor-Switch Array

Takehiko Uno and Yoshio Shimoda

Musashino Electrical Communication Laboratory,
Nippon Telegraph and Telephone Public Corporation
Musashino-shi, Tokyo, 180, JAPAN

Summary

Recently, Digital TCXOs have been studied extensively, because of their high stability and rapid start-up time. However, they have the disadvantages of large circuit scale and high power dissipation. To solve these problems, it is necessary to develop an LSI circuit for TCXOs. This paper describes a new digital TCXO structure suitable for a C-MOS integrated circuit. A new and simple temperature compensation algorithm is also proposed.

The TCXO circuit proposed here has a new crystal oscillation circuit directly controlled by a digital signal. This makes it possible to eliminate D/A converter and varactor diode used in conventional digital TCXOs. The new oscillation circuit is composed of an amplifier, a quartz resonator and a capacitor-switch array. The capacitor-switch array operates as a digital controlled variable capacitor. Both the capacitors and switch elements can be formed by a C-MOS LSI process.

To stabilize the oscillation frequency against supply voltage change, a C-MOS voltage stabilizing circuit composed of three MOS transistors and a voltage divider is proposed.

A simple control algorithm obtaining good compensation effect is considered. In this algorithm, frequency temperature characteristic of a resonator is expressed by a 3rd order polynomial of the temperature. The temperature coefficient for each order A_i ($i=1,2$ or 3) is represented by the summation of a standard value and a deviation δ_i . This deviation value is coded in a few bits and externally adjusted. This makes it possible to eliminate memory cells from the control circuit.

A basic experiment of the new TCXO using

a bread-board circuit was conducted and the results were successful.

Introduction

Temperature compensated crystal oscillators (TCXOs) have been widely used in telecommunication systems, especially in mobile equipments. The frequency stability of TCXOs has been limited in a range of a few ppm. Recently, more stable TCXOs became to be required and digital TCXOs (D-TCXOs) have been studied extensively.¹⁾⁻³⁾ However, they have the disadvantages of large circuit scale and high power dissipation. To solve these problems, it is necessary to develop an LSI circuit for TCXOs. Conventional D-TCXO circuit is unsuitable for one chip LSI, because a D/A converter and a varactor diode are used in the circuit.

In this paper, we describe a new D-TCXO configuration suitable for C-MOS LSI technology. The features of the configuration are as follows:

- 1) New digital controlled crystal oscillation circuit using a capacitor-switch array.
- 2) C-MOS voltage stabilizing circuit in order to stabilize the oscillation frequency against supply voltage change.
- 3) Simple compensation algorithm without any memory cells.

AD P002511

Oscillation Circuit

1. Crystal Oscillation Circuit

A block diagram for a conventional D-TCXO is shown in Fig.1 (a). Temperature information from the temperature sensor is coded into a digital signal by the temperature coder. The programmable ROM or micro-processor generates a compensation data when it receives the coded temperature signal.

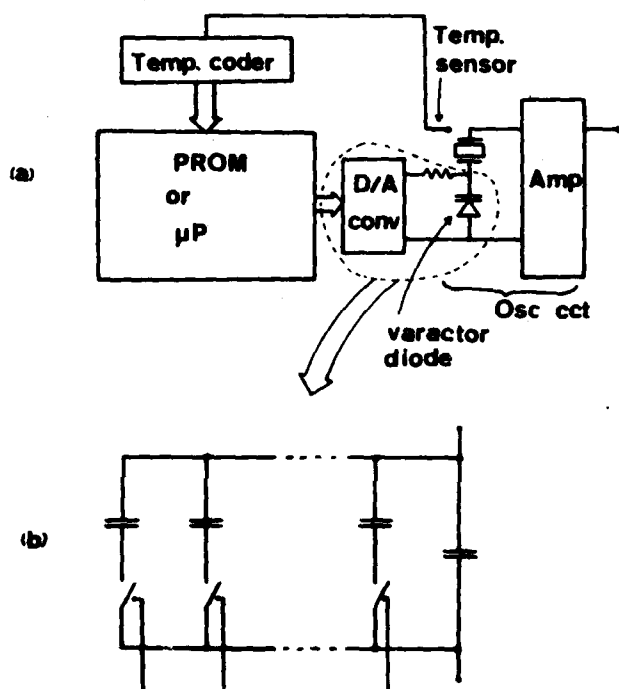


Fig.1. Digital TCXO Block Diagram.

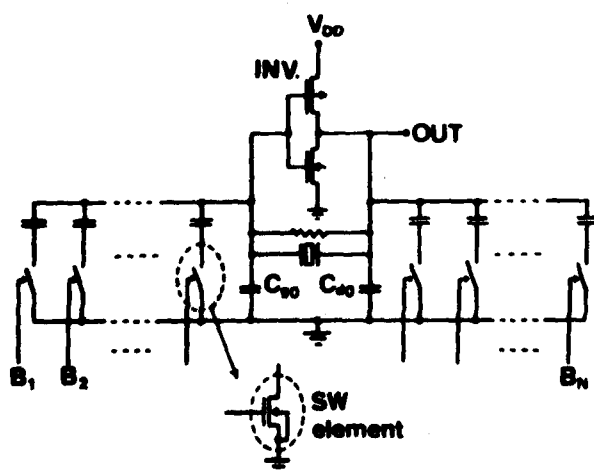


Fig.2. Crystal Oscillation Circuit Using a Capacitor-Switch Array.

The D/A converter controls the capacitance value of the varactor diode. In such a set-up, a reference voltage is needed for the D/A converter, and a varactor diode cannot be formed on an LSI chip.

These problems can be solved by using a capacitor-switch (C-SW) array as shown in Fig. 1 (b). For switch elements, transistors are useful.

Figure 2 shows the new crystal oscillation circuit using a C-SW array. The C-MOS inverter operates as an amplifier, and the capacitance value is controlled by digital signals, B_1 , B_2 and so on. For the switch elements, MOS transistors can be used. As the capacitor elements can be formed on an LSI chip by a C-MOS process, this oscillation circuit except for crystal unit will be built on such a chip.

In the oscillation circuit using a C-SW array, it must be considered that on-resistance R_{on} of an MOS transistor is relatively high. This is solved by connecting fixed capacitors C_{g0} and C_{d0} parallel to the C-SW array. Figure 3 (a) shows an equivalent circuit for the C-SW array considering R_{on} of the switch, where C_0 corresponds to the fixed capacitance and C_x corresponds to capacitance value connected to the switch element. The equivalent capacitance C_{eq} and the equivalent

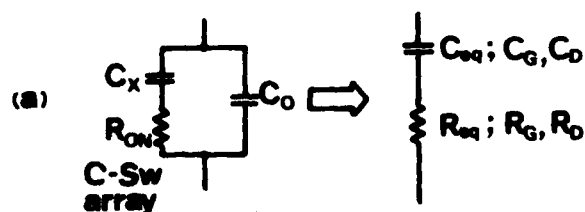


Fig.3. Equivalent Circuit for C-SW Array (a) and Oscillation Circuit (b).

resistance R_{eq} are given by

$$C_{eq} = \frac{(C_0 + C_x)^2 + \omega^2 C_0^2 C_x^2 R_{on}^2}{C_0 + C_x + \omega^2 C_0^2 C_x^2 R_{on}^2} \quad (1)$$

$$R_{eq} = \frac{C_x^2 R_{on}}{(C_0 + C_x)^2 + \omega^2 C_0^2 C_x^2 R_{on}^2} \quad (2)$$

respectively.

Therefore, if C_0 is much larger than C_x , equivalent resistance R_{eq} is reduced enough as compared with R_{on} value.

Oscillation condition of the circuit is obtained from an equivalent circuit. Figure 3 (b) shows the equivalent circuit for the C-MOS crystal oscillation circuit using a C-SW array. C_G and R_G represent equivalent capacitance and equivalent resistance for the C-SW circuit connected to the input (gate) of the C-MOS inverter, and C_D and R_D correspond to them connected to the drain.

Mutual conductance g_m and drain-source resistance r_{ds} of the inverter are given by

$$g_m = g_{mN} + g_{mP} \quad (3)$$

$$1/r_{ds} = 1/r_{dsN} + 1/r_{dsP} \quad (4)$$

respectively, where suffixes N and P correspond to N-channel MOS transistor and P-channel MOS transistor of the inverter.

From Fig.3 (b), load capacitance C_L for the crystal unit and negative resistance $-R$ of the circuit are obtained as

$$C_L = \frac{C_G C_D}{C_G + C_D} \times \left\{ 1 - \frac{g_m r_{ds} [1 + \omega^2 C_D r_{ds} (C_G R_G + C_D R_D)] - C_G / C_D}{1 + g_m r_{ds} + \omega^2 r_{ds}^2 C_D (C_D + C_G)} \right\} \quad (5)$$

$$-R = \frac{g_m r_{ds}^2 C_D / C_G - r_{ds} (1 + g_m R_G) (1 + \omega^2 C_D^2 r_{ds} R_D)}{1 + \omega^2 C_D^2 r_{ds}^2} - R_G \quad (6)$$

respectively, where $r_{ds} \gg R_D$ is assumed.

Figure 4 shows an example of the load capacitance characteristics of the oscillation circuit. A series circuit of a capacitor and an MOS transistor is connected to both the input and output terminal of the inverter.

Table 1. Parameter values of the circuit.

	parameter	specification
inverter	g_m	14.6 mS
	r_{ds}	4.78 k Ω
switches	R_{on} (SW.1, SW.2)	220
capacitors	C_{g0}	90 pF
	C_{d0}	60
	C_{x1}	9
	C_{x2}	6
	$C_{s1}^{(a)}$	8
	$C_{s2}^{(a)}$	8

(a) stray capacities of the printing wired board

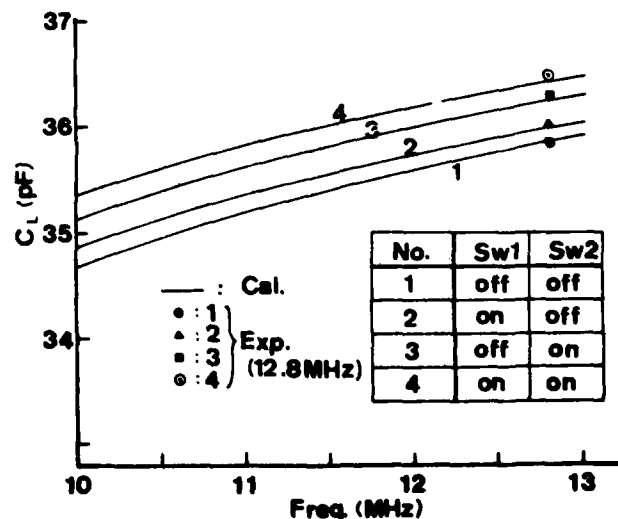
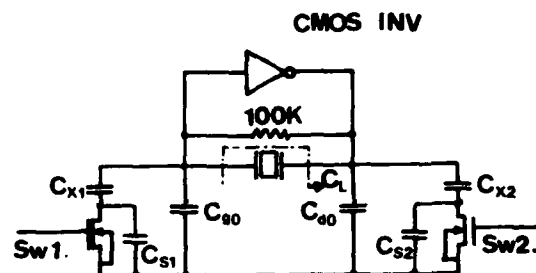


Fig.4. Load Capacitance Characteristics.

Specific values of the circuit parameters are listed in Table 1. Solid line in Fig.4 show load capacitances for various switching states calculated from Eq.(5). It is found that load capacitance C_L depends on both frequency and switching state.

Experimental values at about 12.8 MHz are plotted in the figure. Good agreement with the calculated values was obtained. Here, experimental C_L values were estimated from the oscillation frequency f_{osc} by using Eq.(7):

$$f_{osc} = f_0 \left\{ 1 + \frac{1}{2\gamma(1 + C_L/C_0)} \right\} \quad (7)$$

where f_0 and γ are series resonance frequency and capacitance ratio of the resonator.

From this result, we can accurately design a C-MOS crystal oscillation circuit using a C-SW array.

A bread-board crystal oscillation circuit controlled by a 6-bit digital signal was fabricated in tests. The frequency change per digit has been designed to be equal to -0.3 ppm. The control characteristic of the oscillation circuit is shown in Fig.5. Solid line represents -0.3 ppm/digit slope, and black

circles plot experimental results. In spite of slight deviations from the solid line, experimental values agree well with designed values.

2. Stabilizing Circuit.

A serious problem of a C-MOS crystal oscillation circuit is that the load capacitance depends on the inverter supply voltage. This causes oscillation frequency dependency on inverter supply voltage. Therefore, a stabilizing circuit against supply voltage change is necessary.

Figure 6 shows the voltage stabilized C-MOS oscillation circuit. The stabilizing circuit is composed of three MOS transistors, MP2, MP3 and MN2, and two resistances, R_1 and R_2 .

The condition to stabilize the effective supply voltage, V_E , for the inverter is derived in the following:

Operational point of inverter (MP1, MN1) is determined by drain current I_D of transistor MP2 which is given by

$$I_D = \beta_{P2} (|V_{GS}| - |V_{thP}|)^2 (1 + \lambda_{P2} V_{DS}) \quad (8)$$

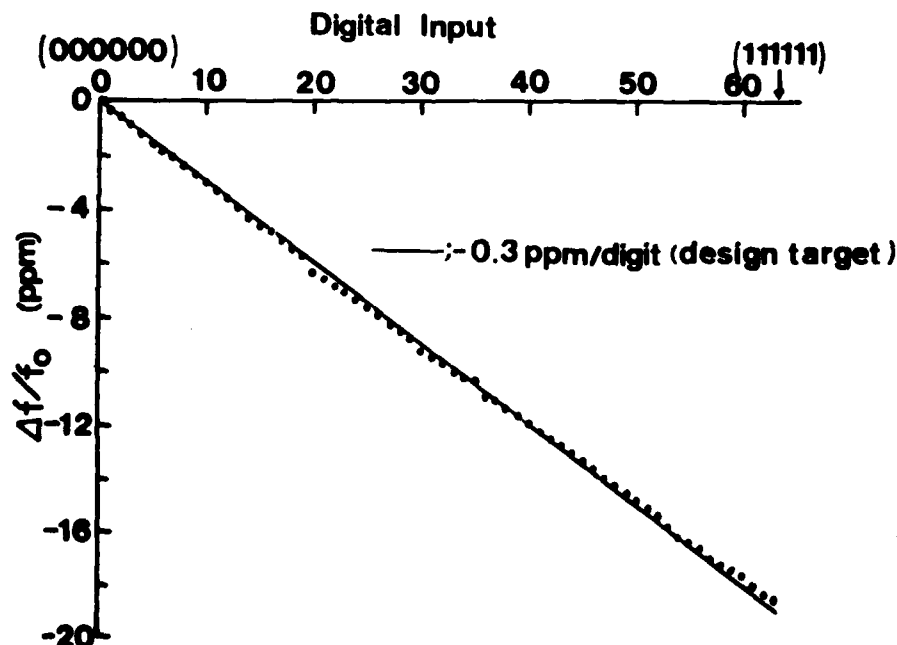


Fig.5. Control Characteristic of the Oscillation Circuit.

If $\lambda_{P2} V_{DS} \ll 1$ is satisfied, Eq.(8) is rewritten as

Furthermore, for transistors MP3 and MN2,

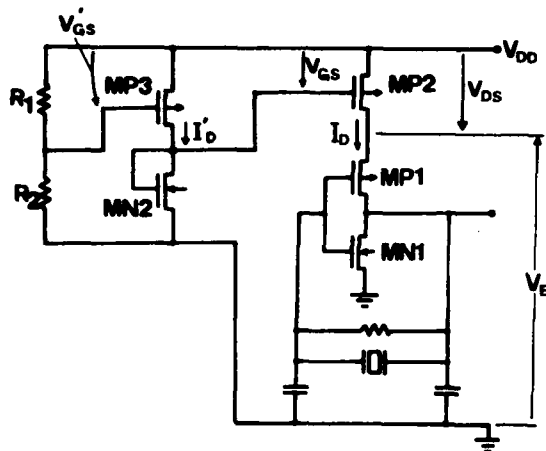


Fig.6.Voltage Stabilized Crystal Oscillation
Circuit.

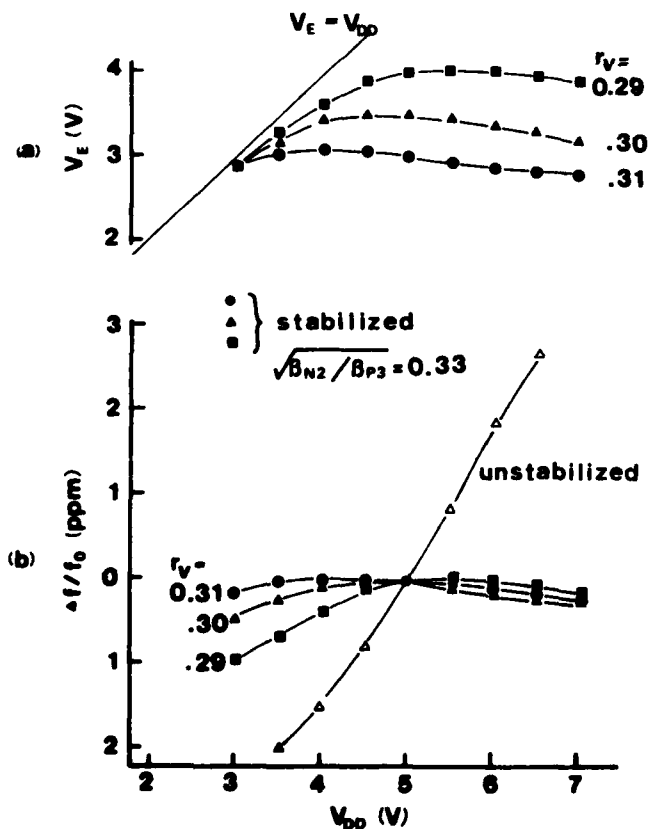


Fig.7. Effect of the Stabilizing Circuit.

$$I_D' \approx \beta_{p3} (|V_{GS}'| - |V_{thP}|)^2 \approx \beta_{n2} (V_{DD} - |V_{GS}'| - V_{thN})^2 \quad (10)$$

is satisfied.

In equations (8) - (10), it is assumed that transistors MP2, MP3 and MN2 operate in saturation region

From Eq. (10), gate-source voltage V_{GS} for MP2 is obtained as

$$-V_{GS} = V_{DD} \left(1 - \frac{R_1}{R_1 + R_2} \sqrt{\frac{\beta_{P3}}{\beta_{N2}}} \right) + \sqrt{\frac{\beta_{P3}}{\beta_{N2}}} |V_{thP}| - V_{thN} \quad (11)$$

Therefore, if

$$\frac{R_1}{R_1 + R_2} = \sqrt{\frac{\beta_{N2}}{\beta_{P3}}} \quad (12)$$

is satisfied, drain current I_D of MP2 will become to be independent from supply voltage V_{DD} . The constant supply current keeps the operational point of the inverter at a fixed point. As a result, the inverter supply voltage V_E and oscillation frequency are stabilized.

Figure 7 shows the effect of the stabilizing circuit. Horizontal axis is the supply voltage V_{DD} , (a) and (b) correspond to effective supply voltage V_E and oscillation frequency change, respectively. The value of $\sqrt{\beta_{N2}/\beta_{P3}}$ was adjusted at 0.33 and the voltage divide ratio $r_V = \{R_1/(R_1 + R_2)\}$ was allowed to vary as a parameter. Voltage V_E is effectively stabilized when $r_V = 0.31$. This r_V value is nearly equal to $\sqrt{\beta_{N2}/\beta_{P3}}$.

The oscillation frequency change corresponds very well to the variation of voltage V_E . In the unstabilized case, the oscillation frequency depends strongly on V_{DD} , while in the stabilized case, that depends on it very little. The best stabilizing effect for this frequency was, like V_E , at $r_V = 0.31$.

Compensation Algorithm and D-TCXO Experiment

Figure 8 shows block diagram of the new digital TCXO. Temperature signal from temperature sensor is coded into a digital signal through temperature coder. Signals δ_1 , δ_2 and δ_3 are coded temperature coefficients of the resonator corresponding to 1st, 2nd and 3rd order. These δ signals are externally adjusted by fixing the input terminals to low or high logic level. The processor generates compensation data from δ_1 , δ_2 , δ_3 and the coded temperature signal. The oscillation frequency is controlled by this compensation data through controlling the switching state of the C-SW array.

This processor circuit can be built by a logic circuit without any memory cells. Therefore, all TCXO circuit except for the crystal resonator and temperature sensor element, will be formed on a C-MOS LSI chip.

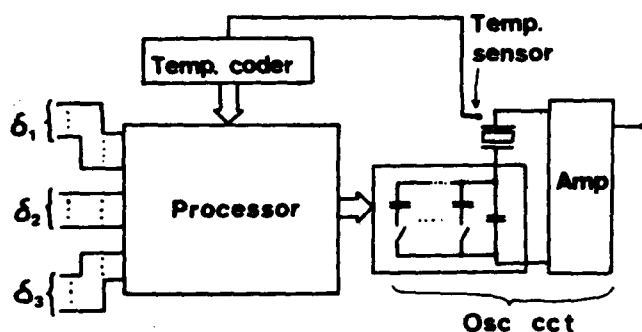


Fig.8. New Digital TCXO Block Diagram.

Compensation algorithm for Fig.8 is explained in the following:

It is well known that the frequency temperature characteristic of an AT-cut quartz resonator can be represented by a third order polynomial of the temperature,

$$\Delta f/f_0 = A_1(T - T_0) + A_2(T - T_0)^2 + A_3(T - T_0)^3 \quad (13)$$

Temperature coefficients A_1 , A_2 and A_3 are mainly determined by the cut angle of the resonator. From investigations of many AT-cut quartz resonator samples, temperature coeffi-

cients can be represented by

$$\left. \begin{aligned} A_1 &= -\delta_1 \times 10^{-7} \\ A_2 &= -(6 + \delta_2) \times 10^{-10} \\ A_3 &= (8 + \delta_3) \times 10^{-11} \end{aligned} \right\} \quad (14)$$

where δ_1 , δ_2 and δ_3 are limited to ranges given by

$$\left. \begin{aligned} 0 \leq \delta_1 \leq 4 \\ 0 \leq \delta_2 \leq 14 \\ 0 \leq \delta_3 \leq 3 \end{aligned} \right\} \quad (15)$$

From this fact, only δ values are needed in order to specify the frequency temperature characteristic of a resonator. These δ values can be coded into digital signals by a few bits. Table 2 shows quantization error of frequency deviation $\Delta f/f_0$ for various coding bit numbers of each order. If we code δ_1 in 7 bits, and both δ_2 and δ_3 in 5 bits, oscillation frequency deviation of the TCXO is limited to the region - 0.2 to + 0.2 ppm for the temperature range of -20 to + 70 °C.

Table 2. Quantization error of frequency temperature characteristic.

coding bit number \ order	1st	2nd	3rd
3	1.12	0.195	0.18
4	0.56	.10	.09
5	.28	.05	.045
6	.14	.025	.023
7	.07		
8	.035		
9	.028		
	ppm	ppm	ppm

Figure 9 shows the limitation of the frequency temperature characteristics corresponding to the deviation given by Eq.(15). Cut angle error for this limitation is about 6 minutes.

The control algorithm has been experimentally studied using a bread-board circuit.

Figure 10 shows the frequency temperature characteristic of the oscillator. For the uncompensated case, oscillation frequency devia-

tes from - 12 to + 6 ppm. Here, normalization frequency f_0 is the oscillation frequency at 25°C. On the other hand, in the compensated case, the frequency deviation is markedly reduced to the range of - 0.2 to + 0.4 ppm.

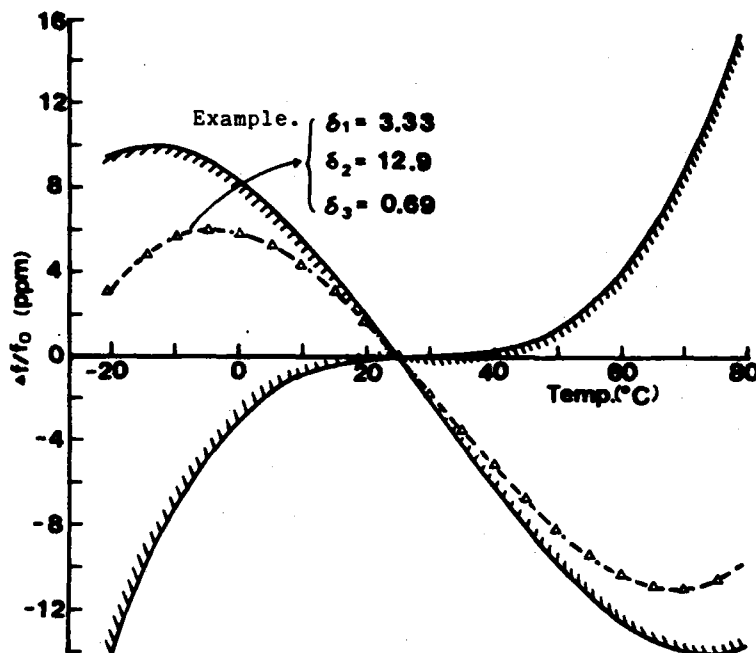


Fig.9. Limitation of the Frequency Temperature Characteristics Corresponding to Eq.(15).

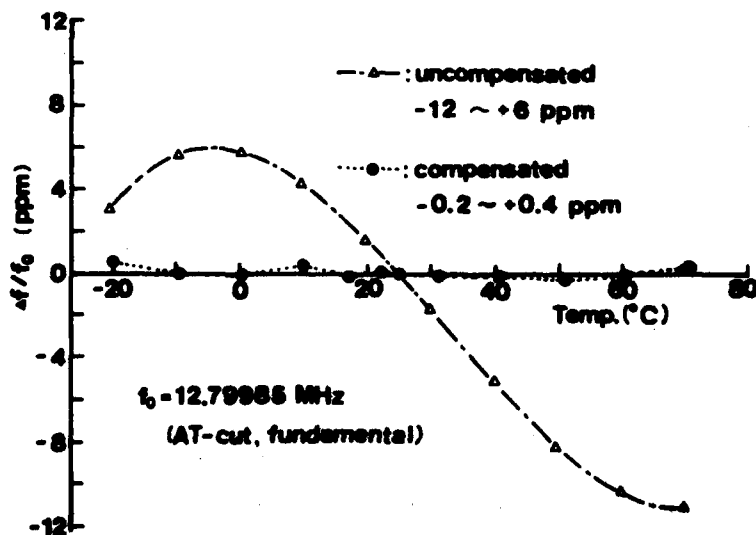


Fig.10. Frequency Temperature Characteristics of the D-TCXO.

Conclusion

We have proposed a new C-MOS crystal oscillation circuit using a capacitor-switch array. This oscillation circuit can be directly controlled by a digital signal without using a D/A converter. It was shown that control characteristic of the oscillation circuit can be accurately estimated from an equivalent circuit.

A C-MOS stabilizing circuit that stabilize the oscillation frequency against supply voltage change was presented.


Furthermore, a compensation algorithm for the coding of the frequency temperature coefficients were proposed.

Experimental results using a bread-board circuit agreed well with expected values. These results strengthen the possibility that a C-MOS LSI can be obtained for digital TCXOs.

Acknowledgements

The authors wish to thank Masakazu Yamada, Director of Electronic Equipment Development Division, and Hidekichi Sato, Director of Engineering Division in Musashino Electrical Communication Laboratory, for their encouragements. They also acknowledge Hiromichi Jumonji, Hideki Ishio, Kota Onuki and Atsushi Iwata in Electrical Communication Laboratories for their useful discussions.

References

- 1) A. B. Mroch et.al., Proc. 30th AFCS, pp.292 - 300, 1976
 - 2) M. Onoe et.al., Proc. 32nd AFCS, PP.398 - 402, 1978
 - 3) M. E. Frerking, Proc. 31st AFCS pp.431 - 435, 1977
- 

AD P002512

AN IMPROVED METHOD OF TEMPERATURE COMPENSATION

OF CRYSTAL OSCILLATORS

James S Wilson

STC Components plc
Harlow England

University of Aston in Birmingham
(Interdisciplinary Higher Degrees Scheme)
Birmingham England

ABSTRACT

A new method of temperature compensation of crystal oscillators has been developed. It is primarily intended for AT-cut crystal oscillators and offers advantages of higher stability, lower power consumption and smaller size compared with oscillators in current production. Compensation is achieved by generating a power series representation of the control voltage for a voltage controlled crystal oscillator. A bipolar integrated circuit has been designed which generates four voltages, each varying as a function of temperature. One voltage is constant while the others have a linear, quadratic and cubic variation with temperature. These voltages are similar in shape to the first four Chebyshev polynomials and the compensation voltage is obtained by summing the voltages in an operational amplifier. Prototype oscillators have been assembled which have a frequency tolerance of less than 1 ppm over the temperature band 85°C to -40°C. The current consumption of these oscillators is 1mA.

Over the last few years there have been market demands for TCXOs which have a higher stability, smaller size and lower power consumption compared with oscillators in current production. An improved method of temperature compensation has been developed enabling TCXOs to be manufactured to satisfy these market requirements.

The most significant factor is frequency stability. Higher stability requires a more complex system of compensation which is best realized using an integrated circuit. The other demands of low power and small size are also compatible with this approach. A bipolar integrated circuit has been designed which generates a temperature-varying control voltage for a voltage controlled crystal oscillator. The method of compensation is primarily intended for AT-cut crystal oscillators and compensation is achieved by generating a power series representation of the desired control voltage. It is also possible to incorporate the voltage controlled crystal oscillator substantially within the integrated circuit, and to create a TCXO in which the majority of the circuit elements are contained within a single integrated circuit.

COMPENSATION VOLTAGE

Conventionally a voltage controlled crystal oscillator consists of a crystal oscillator having a variable reactance in the feedback path. This variable reactance is normally a variable capacitance diode and the frequency of the oscillator is controlled by the d.c. bias voltage across the diode. If this voltage is varied in the correct manner as the ambient temperature varies, the frequency of oscillation is kept constant. The bias or compensation voltage is determined to a large extent by the frequency/temperature coefficient of the crystal and using the Weierstrass theorem, the voltage can be expressed as:

INTRODUCTION

A temperature compensated crystal oscillator (TCXO) consists of two sections. The first section is a voltage controlled crystal oscillator whose frequency of oscillation is controlled by a d.c. voltage. The second section is a circuit which generates a temperature-varying voltage. This voltage is fed to the crystal oscillator and compensates for the frequency changes that would have occurred due to a change in ambient temperature.

$$V = a_0 + a_1(T - T_1) + a_2(T - T_1)^2 + a_3(T - T_1)^3 + \dots + a_n(T - T_1)^n \quad (1)$$

where V = bias or compensation voltage

a_n = weighting coefficient

T = temperature

T_1 = reference temperature

Some non-linearities are introduced by the variable capacitance diode but these are usually small, and a good approximation to the desired compensation voltage is obtained using the first four terms of the power series. Precision compensation is achieved by adding higher order terms such as the fourth and fifth order.

The integrated circuit which has been designed, generates four voltages, each varying as a function of temperature. One voltage is constant while the others have a linear, quadratic and cubic variation with temperature respectively. The compensation voltage is obtained by summing these voltages in an operational amplifier using suitable coefficient-setting resistors.

The integrated circuit does not generate voltages of the form $a_n(T - T_1)^n$. These terms are difficult to generate accurately and are subject to measurement errors. The integrated circuit generates voltages corresponding to the Chebyshev polynomials. The equations of the first four Chebyshev polynomials are:

$$T_0(X) = 1 \quad (2)$$

$$T_1(X) = X \quad (3)$$

$$T_2(X) = 2X^2 - 1 \quad (4)$$

$$T_3(X) = 4X^3 - 3X \quad (5)$$

Figure 1 shows these polynomials plotted as a graph. Equation 1 can be re-written as:

$$V = a_0 T_0(X) + a_1 T_1(X) + a_2 T_2(X) + a_3 T_3(X) + \dots + a_n T_n(X) \quad (6)$$

Apart from the constant, the dominant terms are the linear and the cubic terms, and $T_3(X)$ alone is a good approximation to the desired compensation voltage. The Chebyshev polynomials also have the following advantages:

- (i) the range of values of the weighting coefficients is smaller since the Chebyshev polynomial $T_3(X)$ is a better initial approximation to the desired compensation voltage.
- (ii) the sensitivity of the synthesized compensation voltage to changes in the Chebyshev polynomials is reduced, since the weighting coefficients are smaller in value.

- (iii) the Chebyshev polynomials have an equi-oscillation property. The error between the synthesized compensation voltage and the desired compensation voltage must have this equi-oscillation property and it is more easily obtained using Chebyshev polynomials.

INTEGRATED CIRCUIT DESIGN

A block diagram of the elements of the compensation system is shown in figure 2. The constant voltage is generated using a voltage stabilizer and generates a temperature invariant voltage using a bandgap reference circuit. The linear-varying voltage with temperature is generated using the temperature coefficient of the base-emitter junction voltage of a transistor, operated at constant emitter current. Two currents are generated: one (I_A) has a negative

linear temperature coefficient and the other (I_B) has an equal and opposite positive linear temperature coefficient. Figures 3 and 4 show circuits which could be used to generate the currents I_A and I_B

respectively. Current I_A is used to generate a voltage which varies in a linear manner with temperature.

The currents I_A and I_B are then used to generate other currents with quadratic and cubic variations with temperature. An analogue multiplier as shown in figure 5 is used, and an analysis of this circuit shows:

$$I_0 = \frac{(I_A - I_1)(I_2 - I_3)}{I_1 + I_4} \quad (6)$$

In the first analogue multiplier, used to generate a current with a quadratic variation with temperature, $I_1 = I_2 = I_A$ and $I_3 = I_4 = I_B$.

$$I_0 = \frac{-(I_B - I_A)^2}{I_A + I_B} \quad (7)$$

Since I_A and I_B have equal and opposite temperature coefficients, the sum of I_A and I_B is constant and so current I_0 has a quadratic variation with temperature. This current can then be used to generate a voltage with a quadratic variation with temperature.

A second analogue multiplier is used to generate a current with a cubic variation with temperature. This uses the circuit as shown in figure 5, and equation 6 is used to determine the output current. I_1 is set equal to I_A and I_4 set equal to I_B . I_2 is the current generated by the first multiplier and has a quadratic variation with temperature. I_3 is a constant current. With these values, the output current from the second multiplier is:

$$I_0 = \frac{(I_B - I_A)(I_0 - I_{\text{constant}})}{I_A + I_B} \quad (8)$$

$$I_0^1 = \frac{(I_B - I_A) \cdot I_0}{I_A + I_B} - \frac{(I_B - I_A) \cdot I_{\text{constant}}}{I_A + I_B} \quad (9)$$

I_0^1 is composed of two terms, one which has a cubic variation with temperature and the other which has a linear variation with temperature.

I_0^1 therefore varies with temperature in a manner similar to the Chebyshev polynomial $T_3(X)$. A voltage with a cubic variation with temperature can easily be generated using the current I_0^1 .

The summing amplifier is a high-gain operational amplifier and uses a differential input stage followed by a class AB output stage. The range of compensation voltage which can be generated is to within 0.7V of either supply voltage of the integrated circuit.

Figure 6 shows a typical set of voltages generated by the integrated circuit over the temperature range 120°C to -55°C.

COMPENSATION METHOD

In order to compensate a crystal oscillator using this method of compensation, the values of five resistors have to be determined. These resistors set the values of the coefficients and they are determined from data measured at a number of temperatures over the operating temperature range. At each of these temperatures the compensation voltage needed to keep the oscillator at the correct frequency is measured. The output voltages which represent the linear, quadratic and cubic components are also measured. A computer program is then used to calculate the coefficients. 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 and adjusting the coefficients until the errors between the synthesized compensation voltage and the required compensation voltage are of a specified form. This approximation minimizes the maximum error between the synthesized compensation voltage and the desired compensation voltage. The coefficient-setting resistors have a tolerance of 2%. Once the resistors are fitted the performance of each oscillator is verified by measuring the frequency variation of the unit over the operating temperature range.

EXPERIMENTAL RESULTS

Several prototype oscillators have been assembled using this method of compensation. The integrated circuit is capable of operating from a supply voltage between 4.5V and 15V. The current consumption of a compensated oscillator is less than 1mA. The frequencies of the prototype oscillators are between 4MHz and 10MHz and the tolerances achieved have varied between ±0.5ppm and ±0.9ppm over the temperature range 85°C to -40°C.

Figure 7 shows the frequency variation of one of these oscillators and the graph shows the equi-oscillatory behaviour of the frequency variation. The typical range of compensation voltage generated by the integrated circuit lies between 1V and 5.5V, and the error in the synthesis of the compensation voltage is from ±40mV to ±80mV.

CONCLUSIONS

An improved method of temperature compensation has been described. It is intended for the compensation of oscillators employing AT-cut crystals, although compensation of oscillators with other temperature coefficients is possible. The performance of the prototype oscillators indicates that large volume production of oscillators employing this method of compensation can be achieved. These oscillators would typically have a frequency tolerance of ±1ppm over the temperature range 85°C to -40°C, a current consumption of 1mA, and a package volume of six cubic centimetres.

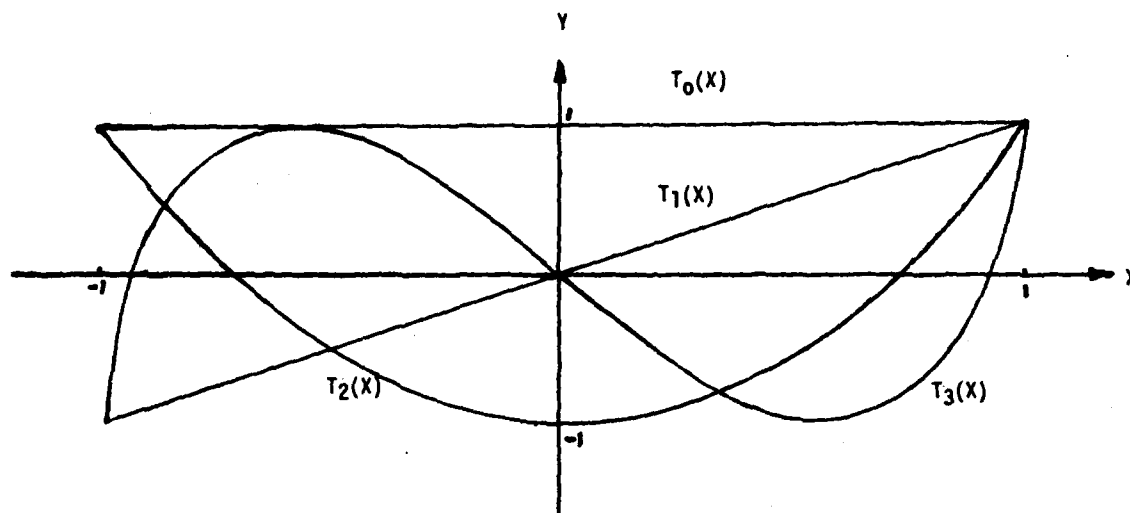


FIGURE 1

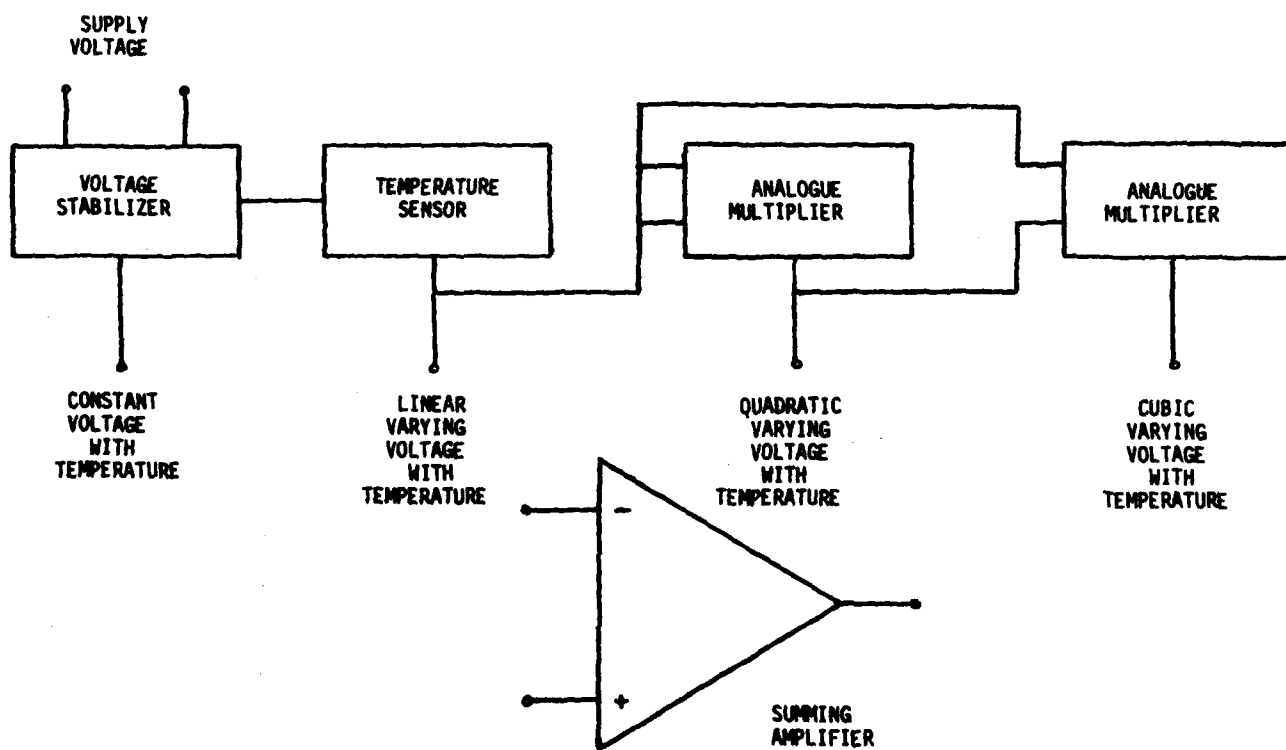


FIGURE 2

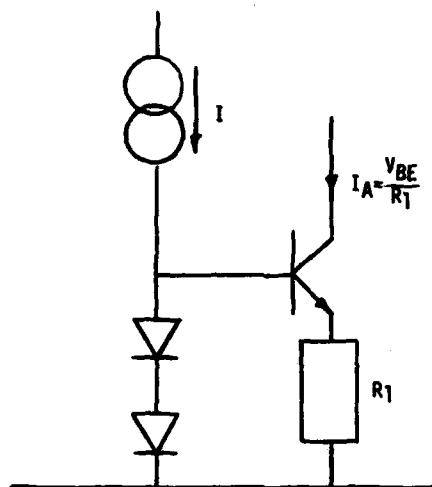


FIGURE 3

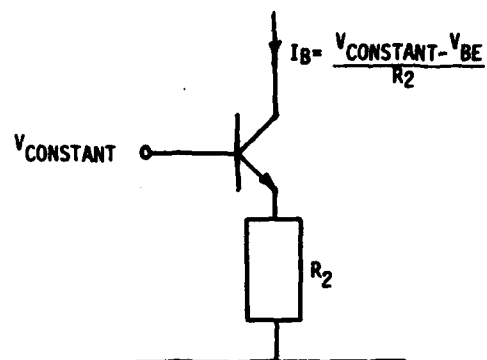


FIGURE 4

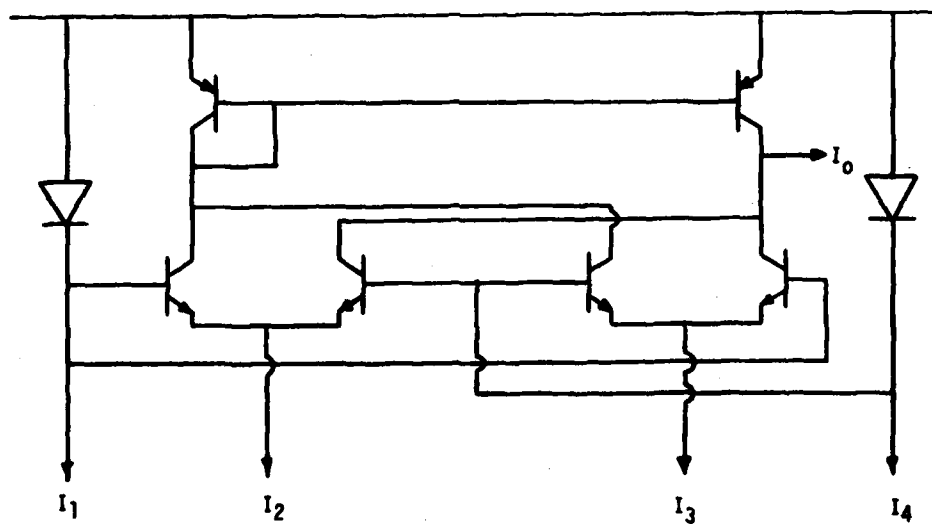
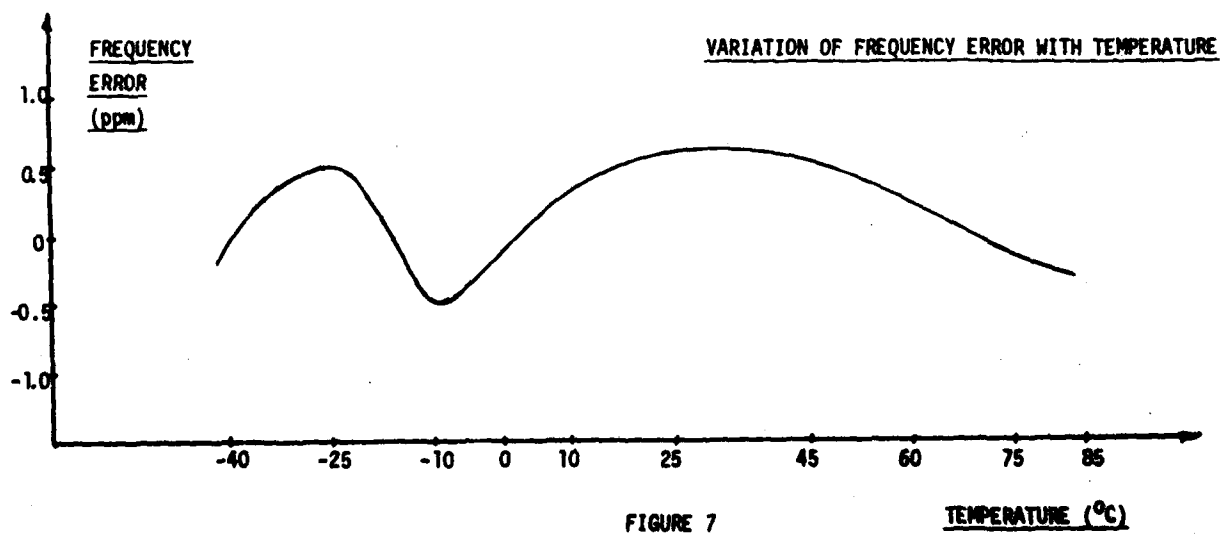
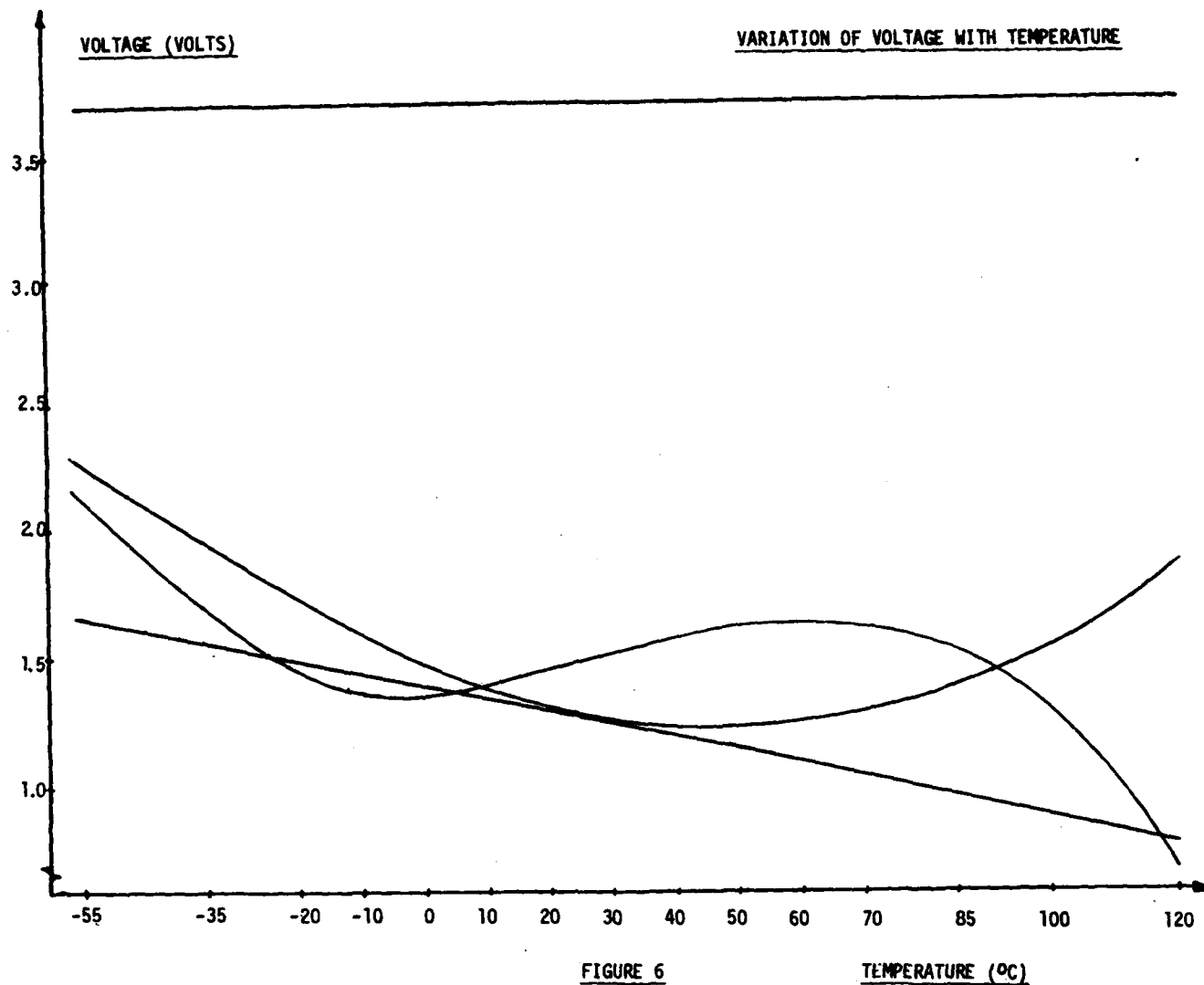


FIGURE 5



Tetsuo Kudo, Syuzo Fujii and Satoshi Nanamatsu

Transmission Division
NEC Corporation

Kawasaki, 211, Japan

AD P 002513

Summary

This paper describes a new frequency temperature compensation method for crystal oscillators. The method is applicable to resonators with second order or third order frequency temperature characteristics. Hence, an oscillator with an AT cut, a BT cut crystal unit or a surface acoustic wave resonator (SAWR) can be efficiently compensated.

A conventional temperature compensated crystal oscillator (TCXO) uses an expensive variable capacitance diode (varactor) with good linearity as well as many thermistors.

Considerable efforts have been devoted to developing a new reactance circuit to replace the varactor. It was found that a series circuit, involving a capacitor and a silicon diode, had the same characteristics as a varactor, and its linearity could be easily adjusted.

In the case of an AT cut crystal unit with third order frequency temperature characteristics, compensation over a wide temperature range can be achieved by using two series circuits.

The compensation in the low temperature range can be accomplished independently from the compensation in the high temperature range. Component values calculation is simple and requires no computer. The same method can be applied to a high frequency oscillator which uses a SAWR.

Several experimental results are shown, which include a 12 MHz AT cut crystal unit, a 600 kHz BT cut crystal unit and a 150 MHz SAWR. It was confirmed that the one order compensation is possible over a wide temperature range.

Introduction

Temperature compensated crystal oscillators (TCXOs), which include crystal units or SAWRs, have been widely used in communication systems. Most of TCXOs use thermistors as temperature sensors and a varactor as a variable reactance element.¹⁻⁶

At least, three thermistors are necessary to compensate for the change in the AT cut crystal unit, whose frequency temperature characteristics are approximated by a third order function.⁷

A SAWR or BT cut crystal unit, whose frequency temperature characteristics are approximated by a second order function, requires at least two thermistors⁸. The varactor should have good linearity. The calculation of values for each component is rather complicated and usually requires a computer.

These requirements for an expensive varactor with good linearity, many thermistors, a selected crystal unit with good frequency temperature characteristics and the complicated calculation to determine the individual component values make an oscillator considerably complicated and hence expensive to manufacture.

With progress in integrated circuit (IC) technology, a digitally controlled TCXO has been developed. The adjustment is easy, but high power consumption and poor short term frequency stability are problems.

This paper presents a new frequency temperature compensation method. The method requires only one thermistor and no varactor. It can be easily applied to compensate for changes in frequency temperature characteristics for an AT cut crystal unit, a BT cut crystal unit or a SAWR.

Compensation Principle

A TCXO consists of two basic sections, as shown in Fig. 1, a voltage controlled oscillator (VCO) section, which must be compensated for, and a frequency temperature compensating section.

A conventional VCO uses a varactor. The relationship between junction capacitance C_j and varactor control voltage V_R can be expressed as:

$$C_j = \frac{K}{(V_R + \phi)^m} \quad (1)$$

where

K : constant
 ϕ : contact potential
 m : capacitance variable ratio

A major problem in the use of a varactor is that only the tuning ratio across a particular voltage range is specified. A considerable variation exists in the capacity-voltage function curve.

Measured data for several varactors are shown in Fig. 2. It is difficult to obtain a good linearity in oscillating frequency versus control voltage characteristics for a VCO with a varactor. The present method replaces the varactor with a series circuit employing a capacitor and a forward operational diode, as shown in Fig. 3.

Figure 4 shows an equivalent circuit for Fig. 3. Equivalent capacitance C_e is given :

$$C_e = \frac{r^2 + \frac{1}{\omega^2} \left(\frac{1}{C_a} + \frac{1}{C_b} \right)^2}{\frac{r^2}{C_a} + \frac{1}{\omega^2 C_a C_b} \left(\frac{1}{C_a} + \frac{1}{C_b} \right)} \quad (2)$$

where

ω : $2\pi f$
 r : forward operational diode resistance
 C_a : capacitance as part of the crystal load capacitance
 C_b : capacitance of connected in series with a diode

The relationship between diode current I and diode voltage V is given:

$$I = I_s (e^{\frac{qV}{kT}} - 1) \quad (3)$$

where

T : absolute temperature
 I_s : saturation current
 q : magnitude of charge of the electron
 k : Boltzmann's constant

Then, forward operational diode resistance r is obtained by Eq. (3) differentiation:

$$r = \frac{\partial V}{\partial I} = \frac{1}{\frac{q}{kT} (I + I_s)} \quad (4)$$

From Eqs. (2) and (4), the relationship between equivalent capacitance C_e and diode current I can be derived, as follows:

$$C_e = \frac{\left(\frac{kT}{q(I+I_s)} \right)^2 + \frac{1}{\omega^2} \left(\frac{1}{C_a} + \frac{1}{C_b} \right)}{\frac{\left(\frac{kT}{q(I+I_s)} \right)^2}{C_a} + \frac{1}{\omega^2 C_a C_b} \left(\frac{1}{C_a} + \frac{1}{C_b} \right)} \quad (5)$$

Figure 5 shows the calculated C_e results as functions of diode current I . The maximum C_e value can be easily adjusted by changing capacitor C_b .

The relationship between load capacitance change for a crystal unit or a SAWR and the oscillating frequency change is expressed as:

$$\Delta(\Delta f/f) = - \frac{1}{2\gamma(1 + \frac{C_L}{C_O})^2} \cdot \frac{C_L}{C_O} \cdot \frac{\Delta C_L}{C_L} \quad (6)$$

where

γ : capacitance ratio
 C_O : parallel capacitance
 C_L : load capacitance

From Eqs. (2) ~ (6), required forward diode current I , to compensate for the frequency change, can be obtained as follows:

$$\frac{\Delta I}{I} = \frac{-C_O \gamma \left(1 + \frac{C_L}{C_O} \right)^2}{1} \cdot \left(\frac{\omega}{P} + \frac{P(I+I_s)^2}{\omega C_b} \left(\frac{1}{C_a} + \frac{1}{C_b} \right) \right)^2 \Delta(\Delta f/f) \quad (7)$$

where

$$P = \frac{q}{kT}$$

It should be noted that in Fig. 5, the change in the equivalent capacitance diminishes at a low as well as at a high current.

In other words, the capacitance change is prominent only in the diode transition range from turn off to turn on. Hence, several series circuits can be incorporated in parallel and operated independently, if on-off voltages for individual diodes do not overlap.

Design Procedure

Designing the compensating section for an AT cut crystal unit with third order frequency temperature characteristics proceeds in three steps:

- (1) Compensation at an intermediate temperature range.
- (2) Compensation at a low temperature range.
- (3) Compensation at a high temperature range.

Figure 6 shows a circuit diagram for the oscillator. Temperature variable capacitance C_6 is for compensation in the intermediate temperature range. The series circuit including capacitance C_3 and diode RC_1 and resistances R_2 , R_3 , R_4 are for the compensation in the low temperature range, whereas C_4 , RC_2 , R_5 , R_6 and R_7 are for compensation in the high temperature range.

Figure 7 (a) shows the frequency temperature characteristics of the VCO section to be compensated for. First, the slope at the intermediate temperature from T_2 to T_3 is compensated for by selecting capacitance value of C_5 and C_7 . The balance shown in Fig. 7 (b) exhibits prominent variation only at low and high temperature ranges.

Compensation at the low temperature range is accomplished first by selecting C_3 , which determines maximum deviation, and diode threshold voltage, determined by R_3 and R_4 .

Then, minor R_3 adjustment is carried out to match characteristics in transition. Diode RC_2 can be disregarded, because it is cut off at a low temperature range.

Similarly, compensation at the high temperature range is accomplished by selecting C_4 , R_5 , R_6 and R_7 . Although diode RC_1 is on at this temperature range, its effect on frequency is saturated and hence negligible.

Diode conditions are shown in Fig. 8. The diode RC_1 operates as variable resistance in the low temperature range. The diode RC_2 operates as variable resistance in the high temperature range.

Designing of compensation for a BT or a SAWR requires only compensation at low and high temperature ranges.

An oscillator circuit, similar to that shown in Fig. 6, is used. Diode RC_2 , however, is reversely connected, so that the diode is gradually back biased when temperature increases.

All these procedures are programmed in a desk-top calculator.

Compensations in low, intermediate and the high temperature ranges are carried out almost independently. Experiments have yielded excellent results.

Experimental Results

First compensation experiments, using an AT cut 12 MHz crystal unit are presented. Figure 9 is an overall compensated result. The less than ± 1 ppm frequency variation is achieved over the from -20°C to 70°C temperature range.

Figure 10 shows the C_3 capacitance change effect in the compensation in the low temperature range. The parameter, a is the present change in capacitance.

Figure 11 shows the R_3 effect on compensation in a high temperature range. Fine tuning for temperature frequency characteristics in the low temperature range can also be done by R_3 selection.

It is also noted that the variation in characteristics in the high temperature range is very small.

Hence, the adjustment is almost independent.

Figure 12 shows the C_4 effect on characteristics in the high temperature range. There is no effect on characteristics in the low temperature range.

The frequency temperature compensation result, for a 600 kHz BT cut crystal unit is shown in Fig. 13. A ± 8 ppm frequency stability was achieved across the 0°C to 60°C temperature range.

The frequency temperature compensation result using a 150 MHz SAWR, is shown in Fig. 14. A ± 10 ppm frequency stability was achieved across the -20°C to 70°C temperature range.

External and internal views of TCOs, using an AT cut crystal unit and a SAWR, are shown in Fig. 15. In the case of the crystal unit, its volume is about 4 cm^3 and the weight is about 9 g. And in the case of the surface acoustic wave resonator, its volume is about 9 cm^3 and the weight is about 15 g.

Conclusion

A new frequency temperature compensation method has been described which can be used to compensate for resonators with second order or third order frequency temperature characteristics.

The merits of the present method are as follows:

- (1) No varactor is needed.
- (2) Only one thermistor is necessary.
- (3) Separated, simple compensation is used in the low and high temperature ranges.
- (4) Wide temperature range compensation can be possible.
- (5) It is suitable for a small size TCO.

Acknowledgment

The authors express their thanks to Prof. M. Onoe, University of Tokyo, for his valuable discussion. They also wish to express their thanks to Mr. M. Oyama and Dr. S. Nonaka, for their encouragement.

References

1. S. Fujii and H. Uchida, "An Analysis of Frequency Stability for TCXO", Proc. 29th Annu. Symp. on Frequency Control, pp.294-299, 1975.
2. S. Fujii and H. Uchida, "Improvement of Frequency Stability for TCXO", NEC R&D No. 43, pp.75-80, 1976.
3. S. Fujii, T. Kudo, H. Takado and T. Yamamoto, "A New Frequency Temperature Compensation Method for TCXO", NEC R&D No. 63, pp.74-82, 1981.
4. K. Hirama and M. Onoe, "A Temperature Compensated Crystal Oscillator Utilizing Three Crystals", Proc. 26th Annu. Symp. on Frequency Control, pp. 132-139, 1972.
5. M. Onoe, I. Yamagishi and H. Nariai, "Temperature Compensation of Crystal Oscillator by Microprocessor", Proc. 32nd Annu. Symp. on Frequency Control, pp. 398-402, 1978.
6. S. Schodowski, "A New Approach to a High Stability Temperature Compensated Crystal Oscillator", Proc. 24th Annu. Symp. on Frequency Control, pp.200-208, 1970.
7. R. Bechmann, A. D. Ballato and T. J. Lukaszek, "Frequency Temperature Characteristics of Quartz Resonators Derived from the Temperature Behavior of the Elastic Constants", Proc. 20th Annu. Symp. on Frequency Control, pp.78-109, 1962.
8. R. G. Kinsman, "Temperature Compensation of Crystals with Parabolic Temperature Coefficients", Proc. 32nd, Annu. Symp. on Frequency Control, pp.102-107, 1978.

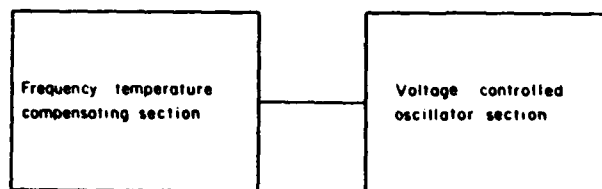


Fig. 1 Block diagram of TCO

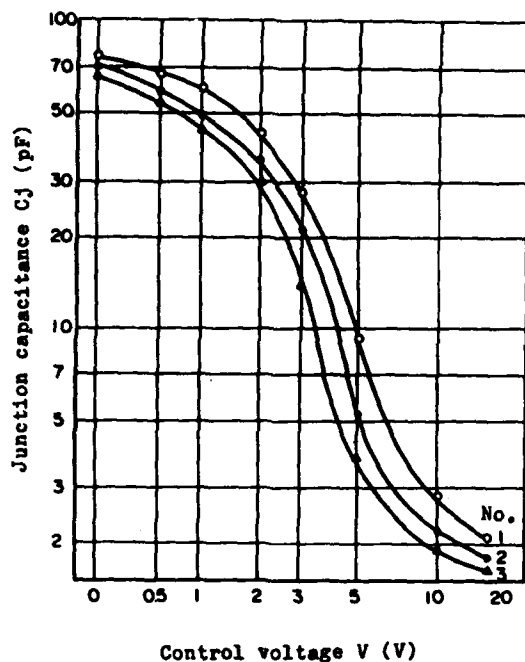


Fig. 2 Junction capacitance vs. control voltage characteristics of several varactors

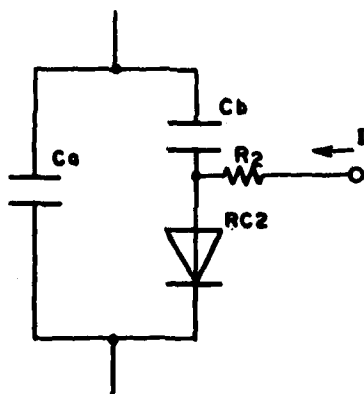


Fig. 3 Reactance circuit

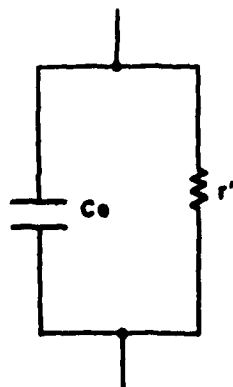


Fig. 4 Equivalent circuit of a new circuit

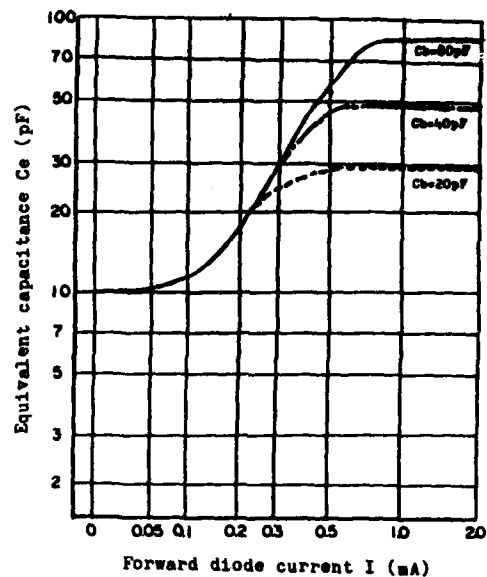


Fig. 5 Calculated curves of equivalent capacitance vs. forward diode current characteristics

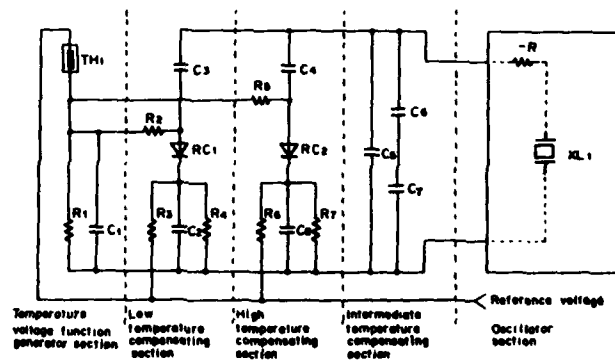


Fig. 6 One TCO block

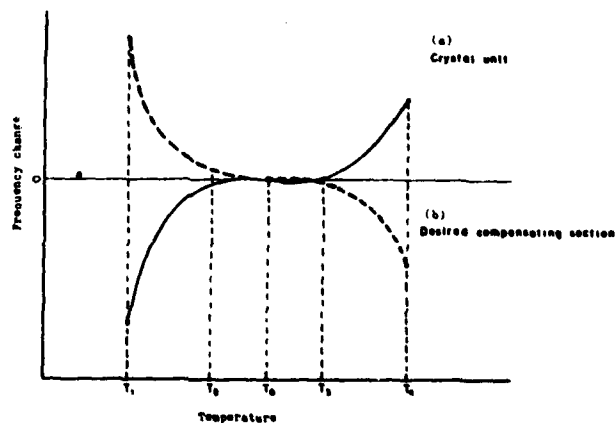


Fig. 7 Relationship between crystal unit frequency vs. temperature characteristics and desired compensating section characteristics

Mode Temperature Range	RC1	RC2
High	ON	VARIABLE RESISTANCE
Intermediate	ON	OFF
Low	VARIABLE RESISTANCE	OFF

Fig. 8 Diode conditions

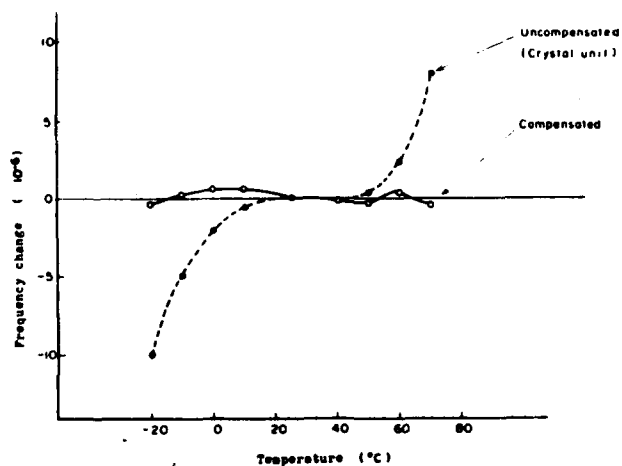


Fig. 9 Frequency vs. temperature characteristics (12MHz AT cut crystal unit)

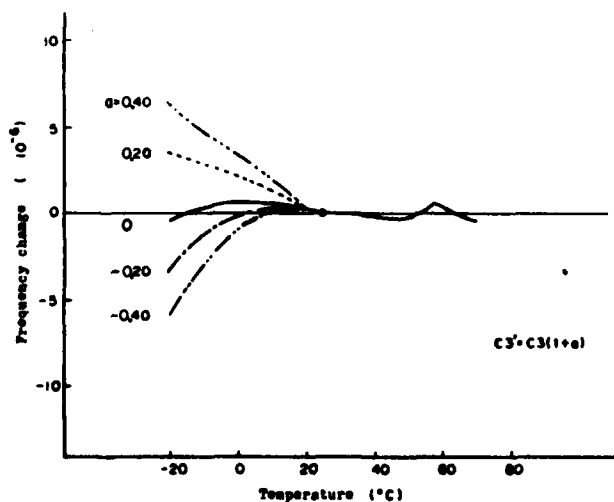


Fig. 10 Calculated frequency vs. temperature characteristics due to capacitance change in capacitor C3

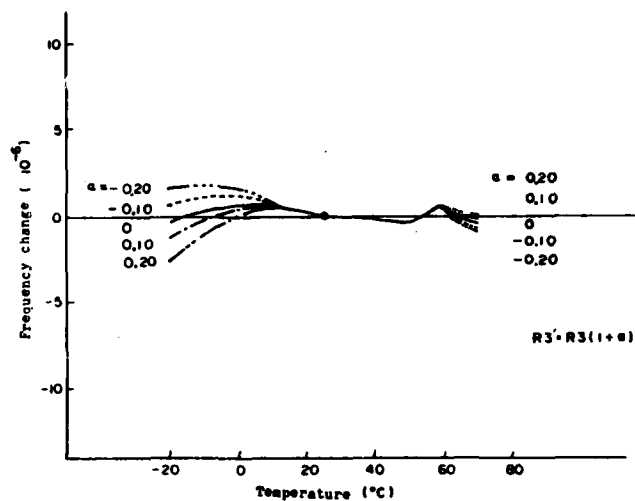


Fig. 11 Calculated frequency vs. temperature characteristics due to resistance change in resistor R3

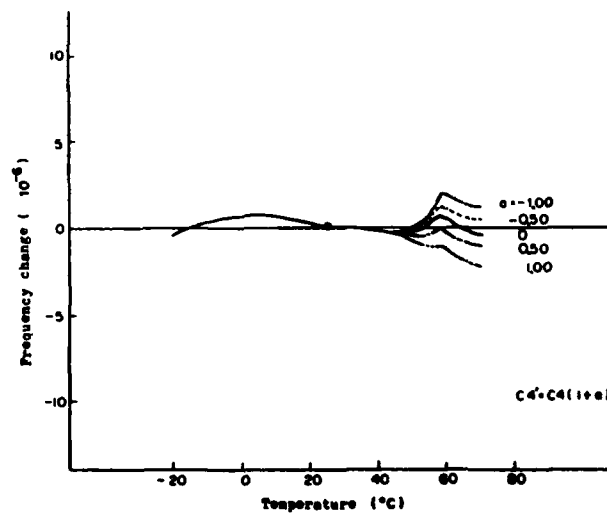


Fig. 12 Calculated frequency vs. temperature characteristics due to capacitance change in capacitor C4

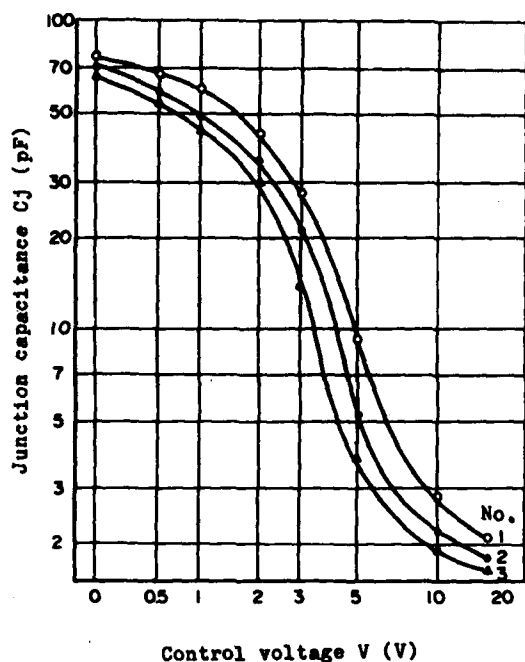


Fig. 2 Junction capacitance vs. control voltage characteristics of several varactors

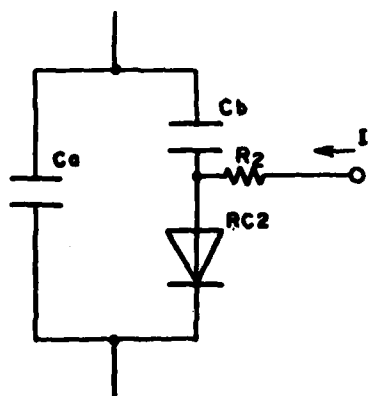


Fig. 3 Reactance circuit

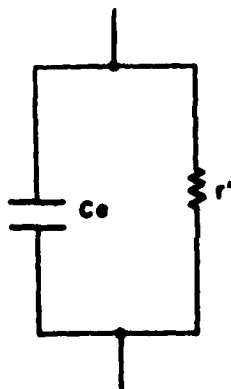


Fig. 4 Equivalent circuit of a new circuit

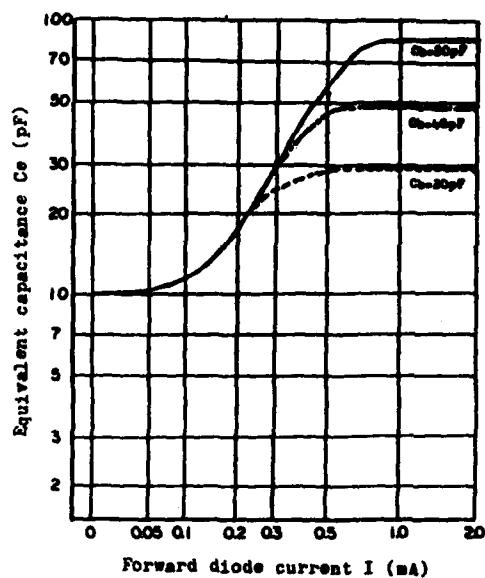


Fig. 5 Calculated curves of equivalent capacitance vs. forward diode current characteristics

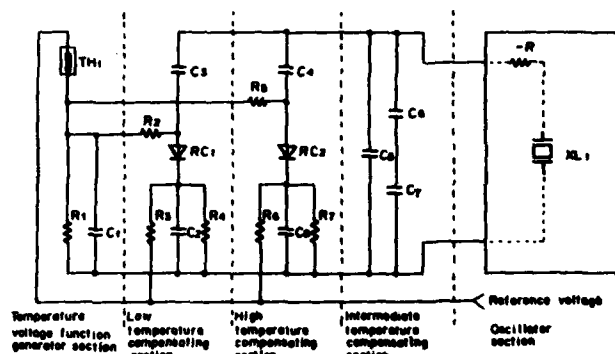


Fig. 6 One TCO block

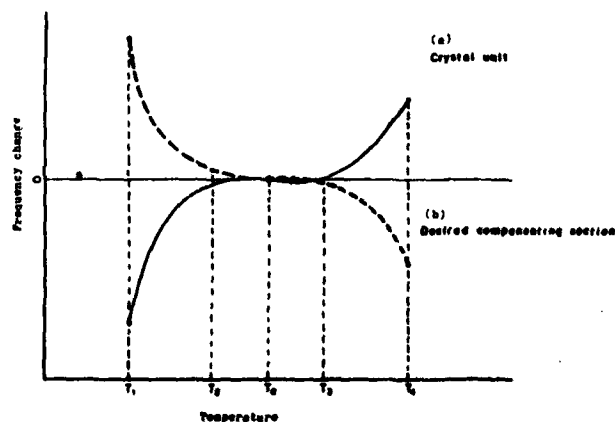


Fig. 7 Relationship between crystal unit frequency vs. temperature characteristics and desired compensating section characteristics

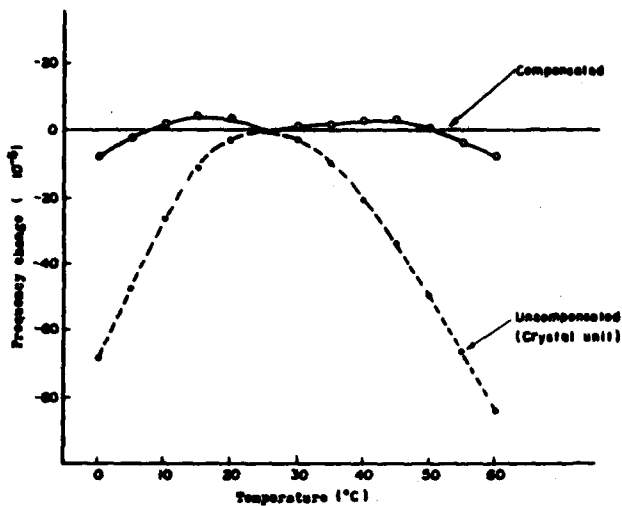


Fig. 13 Frequency vs. temperature characteristic (600kHz HF cut crystal unit)



Fig. 15 External and internal views of new TC06s (The left uses SAWR and the right uses AT cut crystal unit)

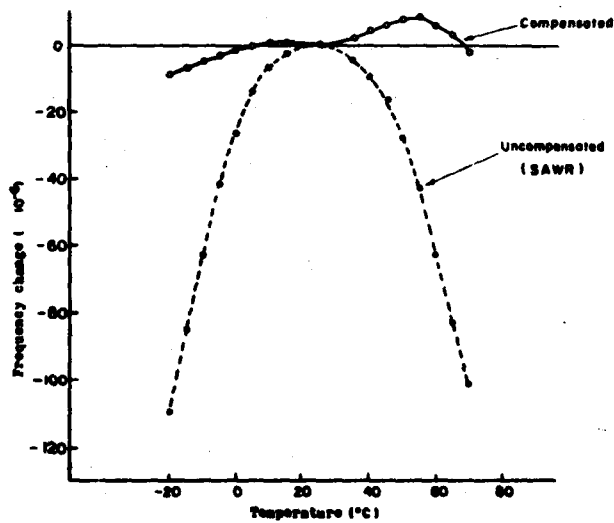
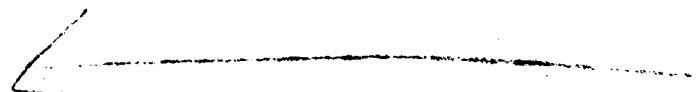


Fig. 14 Frequency vs. temperature characteristic (130MHz SAWR)



AD P002514

AN ULTRASTABLE CRYSTAL OSCILLATOR
FOR BEACONS (PROGRAM SARSAT)

A. STAHL
QUARTZ ET ELECTRONIQUE-ALCATEL
1, rue d'Anjou - 92602 ASNIERES - FRANCE

M. BRINET
CNES, 18 av. E.Belin - 31055 TOULOUSE - FRANCE

S U M M A R Y

An oven controlled oscillator has been developed by Q.E.A in collaboration with the C.N.E.S. to equip beacons for the SARSAT project.

The aim of this project is to speed up rescue operations by pinpointing targets through precise satellite observation.

The means of location are :

- on board equipment assuring reception and pre-processing of the collected data.
- ground stations for processing data sent by satellite in real time to local security stations.
- distress beacons at 406 MHz adapted to the space system. These beacons contain an ultra stable oscillator the characteristics of which condition the performance of the whole system.

Improvements over other similar oscillators have been made with regard to size, power consumption, resistance against shock and vibration.

The SARSAT Program

The SARSAT/COSPAS program issues from a process developed by the U.S.A., Canada and France, and comprises two aspects : first an improvement on the search and rescue system based on the use of the presently existing 121.5/243 MHz distress beacons, and second the demonstration of the advantage offered by a new concept using 406 MHz beacons.

This project will lead to equipping NOAA meteorological satellites of the TIROS with repeaters for the 121.5/243 and 406 MHz frequency bands and 406 MHz processors. These satellites are low-orbiting and cover the whole globe in less than 12 hours.

The first satellite was launched in April 1982 and two USSR Spacecraft are also in orbit now with the same function as the American one.

The 406 MHz system constitutes an original contribution by the C.N.E.S. to the SARSAT program and is the result of experience acquired in the field of location and data collection from the ARGOS system currently exploited.

In case of distress the 406 MHz beacons will transmit every 50 seconds a short message of about 500 ms containing information to be processed in the spacecraft. The beacon includes a message coding system, an ultrastable oscillator (U.S.O.) and a modulator transmitter controlled by the U.S.O. Power supply is provided by a battery of cells ensuring operation for at least 48 hours.

The precision in locating a target is directly related to the performances of the U.S.O. which will now be described.

DESIGN DETAILS

General considerations

A stringent requirement is frequency stability in the presence of a thermal gradient : that

is to say no more than $5 \cdot 10^{-10}/^{\circ}\text{C}$ of frequency deviation is allowed for in the presence of a temperature variation of $1^{\circ}\text{C}/\text{Min}$. For this reason a TC oscillator was eliminated and choice was made for an oven controlled one, despite its larger power consumption. In fact the biggest difficulty encountered was the reduction of power consumption, so the primary objective was to develop a technology able to preserve a high vacuum inside the package for several years.

For these reasons the oscillator is mounted inside a TO8 package hermetically sealed and suspended in a vacuum box ; (only the voltage regulator and adjustment device are located on a PC board outside the vacuum package).

This box (Fig.1) is composed of a header with glass feed through and of an outer casing of stainless steel welded by the TIG process to the header and provided with a pumping pipe at its top.

Figure 2 shows how the oscillator package is fixed : the upper side of the TO8 case is attached by a stretched stainless steel wire at the top of a frame work, while the lower part is provided with 3 hooks on which 3 stainless steel loops are fixed, maintaining the entire package suspended.

This mechanism yields a very high resistance against shocks and vibrations.

The electrical connection wires are also made of stainless steel, in the form of a coil to minimize thermal conduction.

Thermal considerations

The diagram on Figure 3. gives us an idea of what the calculations are :

The total power consumption is :

$$W = W_{\text{reg}} + W_{\text{el}} + W_{\text{heat}}$$

$$W_{\text{reg}} = 50 \text{ mW}$$

$$W_{\text{el}} = 30 \text{ mW taken up by the electronics}$$

$$W_{\text{heat}} = \text{power to maintain the TO8 case at } 70^{\circ}\text{C}.$$

The thermal equilibrium of the TO8 is obtained when :

$$W_{\text{heat}} + W_{\text{el}} = W_{\text{rad}} + W_{\text{con}}$$

where

$$W_{\text{rad}} = \text{radiated power (} \sim 30 \text{ mW)}$$

$$W_{\text{con}} = \text{thermal loss through the suspending wires and electrical conducting wires. (Heat conduction through the residual gas is neglected.)}$$

The measured power consumption at different temperatures is as follows :

at	- 40°C200 mW
	+ 25°C130 mW
	+ 55°C100 mW

The vacuum problem

Figure 4 shows the power consumed in accordance with the residual gas pressure. We can see that a pressure of at least $1 \cdot 10^{-4}$ torr has to be maintained to assure good thermal isolation. This means that the average leakage rate (or outgassing rate) must lie under $2,5 \cdot 10^{-14}$ torr l/s if the pressure is not to be exceeded over a period of 3 years (for instance). Such a low rate is not directly measurable but experience has shown that this can be achieved.

In fact there are 3 ways in which the vacuum might be impaired :

- permeation through the case walls, which is negligible.
- Degassing from materials inside the case ; to minimize this effect no organic materials were allowed and mainly stainless steel was used. Baking of each part and very clean processing is of paramount importance.
- Leakage through welding defects or inefficient feedthrough. Last but not least a low temperature activatable getter (from SAES) is placed inside the vacuum space, which greatly improves the vacuum. An especially interesting fact about this getter is the possibility of reactivating it several times during the life time of the oscillator.

About 20 oscillators have been tested,

some having a 12 month shelf life, either in our factory or in the C.N.E.S. Laboratories, and the following observation has been made : either the pressure goes up along with power consumption, and in this case all we can do is recover the TO8 package with the oscillator and mount it in an other case, or the power consumption remains constant and the unit retains its initial characteristics.

Electrical design features

To achieve low consumption the choice was made of a Pierce type oscillator using a fundamental mode quartz. As shown in figure 5. this oscillator is composed of 4 parts :

- an oscillating circuit and means of adjustment
- an HF amplifier and buffer for 50 ohm operation
- a heating device, that is a transistor soldered on to the TO8 header

These 3 devices use hybrid circuit technology (Figure 6)

- finally a voltage regulator located outside the outer box.

To avoid contamination of the quartz which would lead to a poor aging rate a very clean process is needed and each part is baked at a temperature in accordance with the nature of the material.

The table 1 shows the principal characteristics of the U.S.O., both short and medium term stability are necessary to ensure a proper functioning of the phase lock loop of the on-board processor as well as precise location : the aim is to have not more than 1 to 2 kms error. And to achieve this the mean slope of the frequency shift has to be no more than $8 \cdot 10^{-10}$ /min over 15 min. (Figure 7).

CONCLUSION

An ultrastable oscillator with low power consumption has been developed to meet the requirements of the SAREAT program. In view of its very characteristics other uses for this oscillator may also be envisaged.

TABLE 1

<u>FREQUENCY</u>	5 076 kHz
<u>POWER CONSUMPTION</u>	at - 40°C : 200 mW (140) + 25°C : 140 mW " + 50°C : 100 mW "
<u>FREQUENCY STABILITY</u>	Between - 40°C and + 55°C ... $\frac{\Delta F}{F} < 5 \cdot 10^{-8}$
<u>MEAN SLOPE OF FREQUENCY DRIFT</u>	$\frac{\Delta F}{F} < 8 \cdot 10^{-10}/\text{min}, 15 \text{ min.}$
<u>WARM UP TIME</u> at - 40°C : 4 min. ;	$\frac{\Delta F}{F} = 2 \cdot 10^{-7}$
<u>SHORT TERM STABILITY</u>	$5 \cdot 10^{-10}, \tau = 1 \cdot 10^{-1} \text{ sec.}$
<u>OUT PUT</u>	140 mV/50 Ω (V = 14 - 20V)
<u>HOUSING</u>	ϕ 32 mm H = 37 mm
<u>WEIGHT</u>	60 g
<u>SHOCK</u>	100 g 6 ms 3 axis
<u>VIBRATIONS</u>	<div> <div> 5 - 2 000 Hz 7 g. 1 oct/min. 2 000 Hz 10 min. </div> <div> } operating material </div> </div>

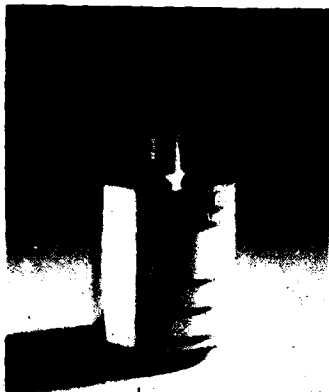
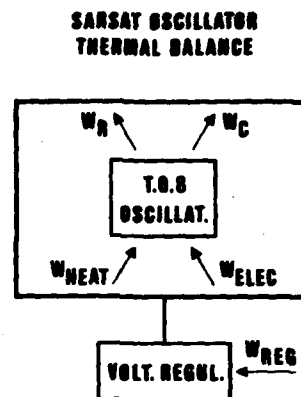


Fig. 1



Fig. 2



$$W_{TOT} = W_{REG} + W_{ELEC} + W_{HEAT}$$

$$W_{HEAT} + W_{ELEC} = W_R + W_C$$

Fig. 3

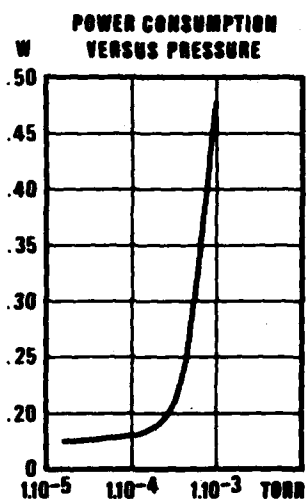


Fig. 4

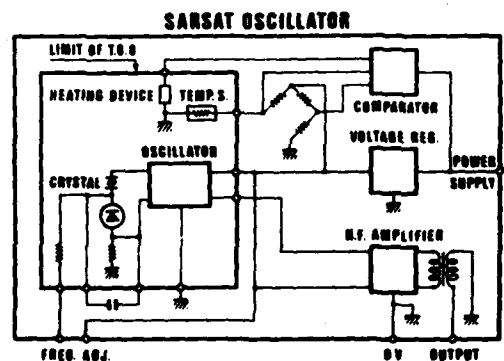


Fig. 5



Fig. 6

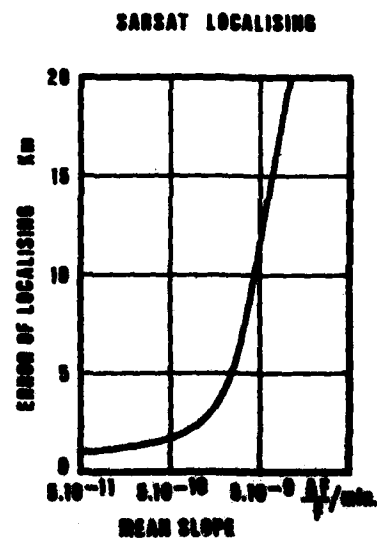


Fig. 7

AD P 002515

MODELING OF TTL INVERTER AMPLIFIERS AS APPLIED TO OVERTONE CRYSTAL CLOCK OSCILLATORS

Dan Nehring
CTS Knights
400 Reimann Ave.
Sandwich, IL 60548

Summary

Methods have been developed for analysis of clock oscillator circuits which utilize TTL inverter gates as active devices. Equivalent circuit parameter values have been measured for Schottky TTL inverter amplifiers and used in computer analysis of oscillator circuits. Computer optimization of third overtone oscillators has been done using the models developed for the TTL inverters. The basic circuit consists of a biased TTL inverter driving a notch filter with a coupled dual monolithic resonator completing the feedback loop back to the amplifier input. Tuning the notch filter to the fundamental crystal frequency inhibits fundamental mode oscillations and allows third overtone mode oscillations to occur. Several variations of this circuit have been computer analyzed and empirically tested yielding good performance for third overtone mode oscillators.

oscillating loop, information was needed as to the characteristics of the TTL inverter amplifiers and the twin-tee filter.

Introduction

Quartz crystal clock oscillators have been produced for many years with the miniaturized thick film hybrid type being the dominant form of construction utilized. These oscillators are produced in large quantities using techniques that yield devices with appropriate electrical performance at absolute minimum cost. A circuit that has been utilized as a fundamental mode oscillator with success is shown in Figure 1.¹ This is a simple circuit using few parts and giving excellent electrical performance. The amplifier is a Schottky TTL inverter which is biased in a linear mode via two resistors. This creates an amplifier which has significant gain over a range of input voltages with 180 degrees phase shift. Oscillations are generated by the feedback of the monolithic resonator on its symmetric mode of resonance.²

Figure 2 is a third overtone adaptation of the fundamental circuit. This circuit is proposed as a simple, inexpensive third overtone oscillator.³ Advantages are that no tuned circuits are needed for overtone selection and only one semiconductor die is used for all active devices. The twin-tee notch filter is tuned to the fundamental frequency of the crystal and blocks fundamental operation. At the third overtone frequency, the filter has much less loss and third overtone oscillations occur.

A second important function performed by the notch filter is that of lead compensating the amplifier. Since the TTL inverter is made up of several transistor stages, an inherent propagation delay exists across the amplifier. Schottky inverters typically exhibit a 2-5 ns delay which translates into a phase lag whose magnitude depends on the frequency of operation. If the frequency of operation is not too great, as in the case of a fundamental oscillator, the phase lag will not be a problem. However, a third overtone clock operating at 40 Mhz encounters a 30 to 40 degree loop phase error due to the amplifier phase lag. Since a twin-tee notch filter acts as a lead network at a frequency three times its notch frequency, (Figure 3) it can offset the phase lag of the amplifier and allow the monolithic resonator to operate at its proper 180 degree phase frequency and point of maximum transmission. In order to properly optimize the

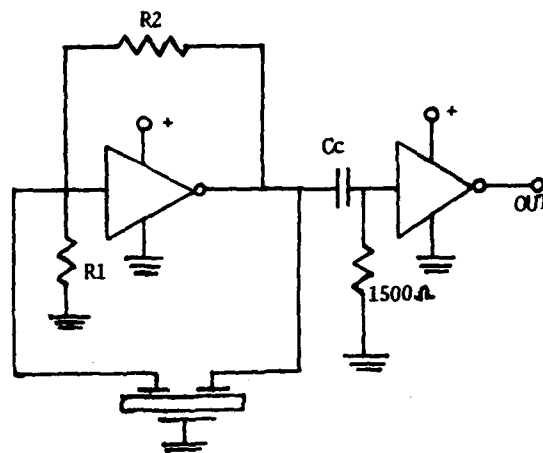


FIGURE 1

DUAL MONOLITHIC INVERTER OSCILLATOR

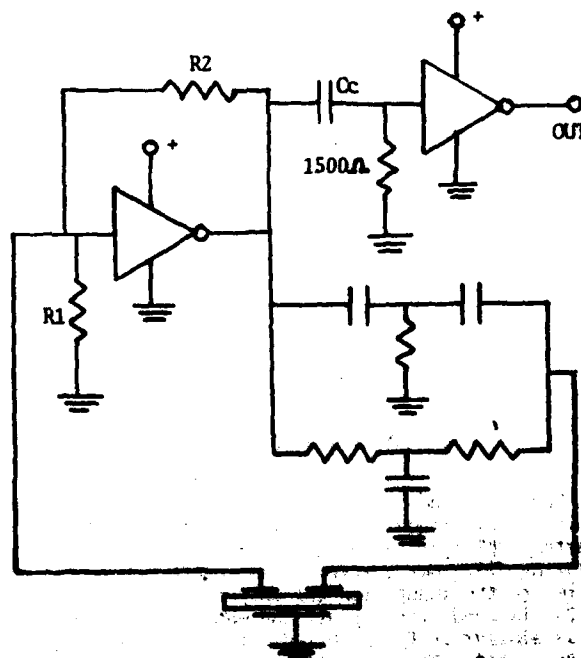


FIGURE 2

THIRD OVERTONE INVERTER OSCILLATOR

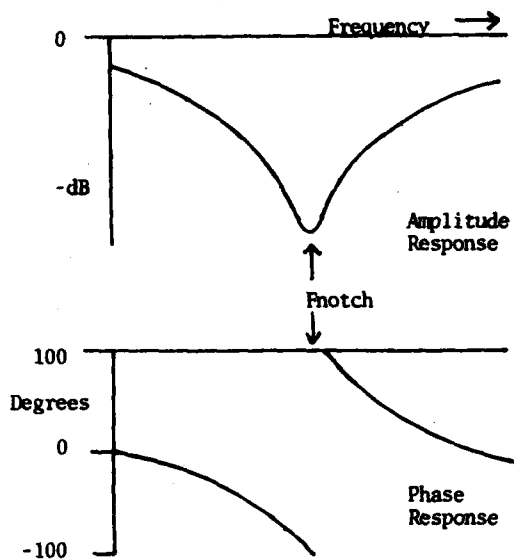
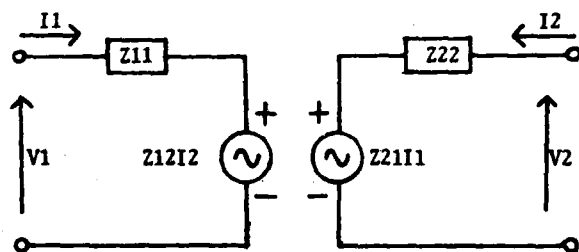


FIGURE 3

TWIN-TEE NOTCH FILTER CURVES

Characterization of TTL Inverter Amplifiers

The biasing scheme for the TTL amplifier was illustrated in Figure 1. The R1 and R2 resistors will place the input and output voltages near the mid-point of their active region providing a reasonably linear inverting amplifier. Modeling of this amplifier was accomplished by utilizing a Z-parameter equivalent network as shown in Figure 4.



$$V1 = Z11I1 + Z12I2$$

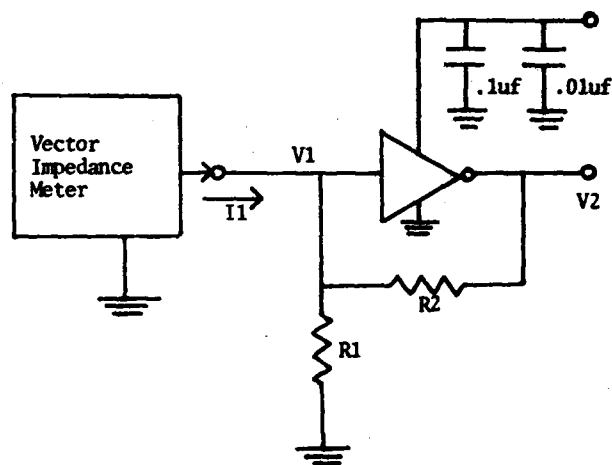
$$V2 = Z21I1 + Z22I2$$

FIGURE 4

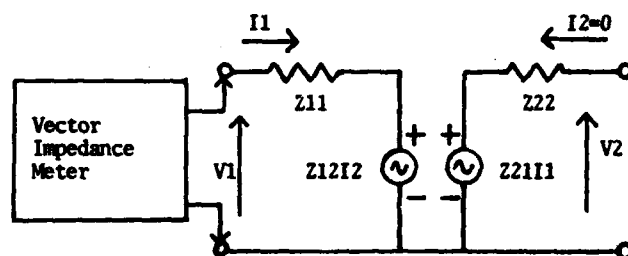
Z-PARAMETER EQUIVALENT CIRCUIT

Measuring the four Z-parameters values provides a complete small signal ac equivalent circuit for the inverter amplifier. Figure 5 shows the method for measuring Z11. Here, a vector impedance meter is applied to the input with the output open circuited and Z11 is read directly off the meter. Measurement of Z22 is similar as Figure 6 indicates. By applying the impedance meter to the amplifier output, Z22 is determined. Measurement of Z12 and Z21 is somewhat more complicated. Figure 7 shows the method for measurement of Z12. Network theory indicates that Z12 is obtained by setting I1 to zero and determining the ratio

of V1 to I2. However, the presence of the voltmeter probe on the amplifier input will allow a significant I1 to flow and must be included in the calculations. Writing equations for the Figure 7 equivalent network yields the equation for Z12 shown in Figure 7. Measurement of Z21 is accomplished as in Figure 8. Here again the voltmeter probe impedance interferes with the measurement of I1 through the sampling resistor Rs. Including this in the calculations gives the Z21 equation of Figure 8. These methods were applied to amplifiers and Z-parameter empirical values were obtained.



Physical Circuit

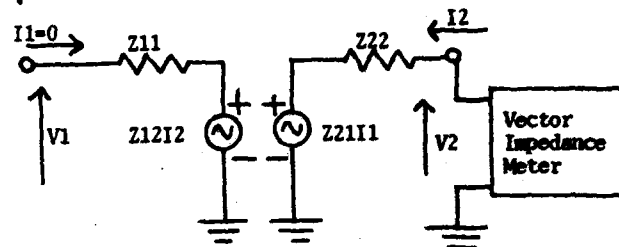
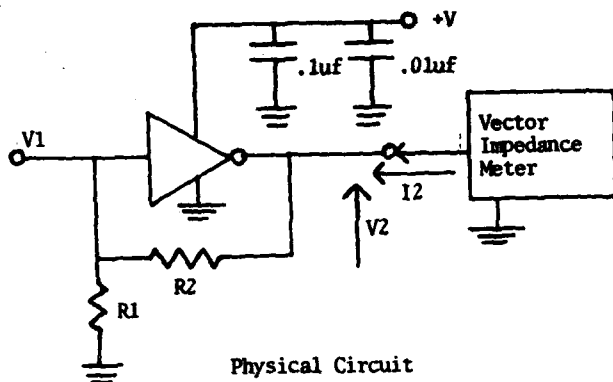


Equivalent Network

$$Z11 = \frac{V1}{I1} \Big|_{I2=0}$$

FIGURE 5

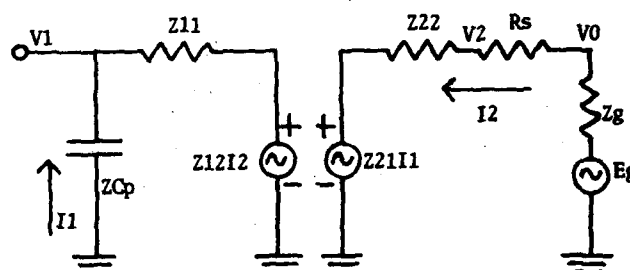
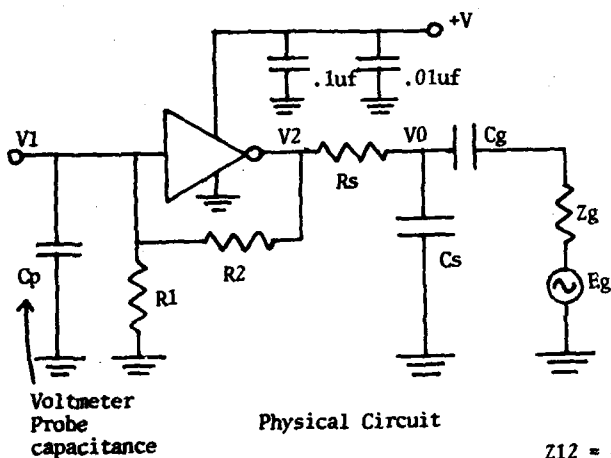
TEST SETUP FOR Z11 MEASUREMENT



$$Z_{22} = \frac{V_2}{I_2} \Big|_{I_1 = 0}$$

FIGURE 6

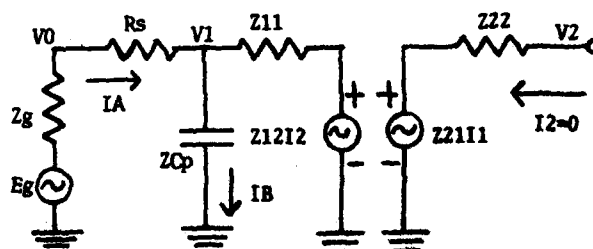
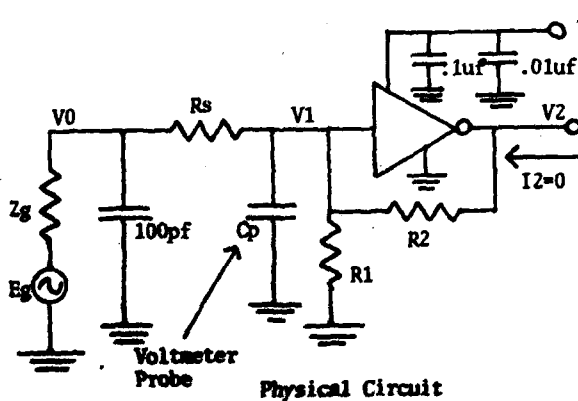
TEST SETUP FOR MEASURING Z_{22}



$$Z_{12} = \frac{V_1 R_s}{V_0 - V_2} \left(1 + \frac{Z_{11}}{Z_{Cp}} \right)$$

FIGURE 7

TEST SETUP FOR MEASURING Z_{12}



$$Z_{21} = \frac{V_2 R_s Z_{Cp}}{V_0 Z_{Cp} - V_1 (Z_{Cp} + R_s)}$$

FIGURE 8

TEST SETUP FOR Z_{21} MEASUREMENT

Measurement Results

Two different biasing resistor values were investigated. Figure 9 shows both types. Both versions put the input and output amplifier voltages in the active region of the inverter's transfer curve. The 1200 ohm bias resistor version offers better phase stability as a function of supply voltage while the unequal bias resistor version provides increased gain for more reliable oscillations. Measurements were made on sample amplifiers at room temperature in accordance with the methods described in the last section. Amplifiers were mounted on a brass plate which was utilized as a ground plane with test connections made with the shortest possible leads. A Hewlett-Packard vector impedance meter model 4815A was used for Z11 and Z22 measurements. This is a low level signal impedance meter and, therefore, consistent with the small signal analysis circuit. In measuring Z12 and Z21, a Hewlett-Packard vector voltmeter model 8405 was used. For the measurement of Z12, R_S was 50 ohms, C_g was .1 microfarads and C_s was 100 picofarads. Voltmeter connections were accomplished via BNC type adapters. This connection generated an approximate 6 picofarad probe capacitance which was used for the C_p value. V1 is the reference channel for the voltmeter and is maintained at 2mV RMS. C_s has the effect of eliminating any change in signal levels when the second voltmeter probe is moved from V0 to V2 during measurement.

Z21 measurements were made per the Z21 test set-up with $R_S=50$ ohms and V1, the phase reference, maintained at 10mV RMS. Z-parameter values were generated for frequencies from 1 to 65 MHz using Texas Instruments SN54S04 dual-in-line hex inverters. Results of the measurements are shown in Table 1 and 2 for both bias resistor values. Observing these results it can be seen that the Z11 term ranges from 68 to 160 ohms for both amplifiers with a decreasing phase angle as frequency increases. In both cases the Z21 term, which is the feed-forward gain of the amplifier, is seen to have an exact 180 degree phase relationship to the input current at low frequencies and steadily decreases as frequency increases. This is due to the propagation delay of the amplifier as previously discussed. Also the Z21 magnitude is seen to be greater for the 2400, 1500 ohm biased amplifier as would be expected. In both cases the Z22 term is seen to be less than twenty ohms at any frequency indicating a very low output impedance.

In order to verify the accuracy of the Z-parameter values, the test circuit of Figure 10 was used. This is an inverter amplifier driven by a generator with 100 ohms in series with its input. A load of R_L and a voltmeter capacitance C_L is present on the output. If a signal applied to the input produces an output of the same magnitude and phase as is predicted by a computer using the Z-parameter model, then the Z-parameter values will be considered to be correct. Three different tests were run; one using the 1200 ohm bias resistors with a 470 ohm R_L , one with the same bias with a 47 ohm R_L and one using the 2400 ohm, 1500 ohm bias and a 470 ohm R_L .

All three tests were run with the V0 voltage at 10mV RMS. Circuit response was converted into gain (V_2/V_0) and compared to the computer predicted performance using the transfer function of Figure 10. Figure 11 is a comparison plot of predicted and experimental results for the 1200 ohm biased amplifier with a 470 ohm load. Figure 12 is the same for a 47 ohm load. In both cases, the correlation of data is seen to be quite good. Phase error is typically

2 to 3 degrees and magnitude error is insignificant. Due to the high correlation of data, these Z-parameter values are considered to be correct and able to represent the amplifier under any source and load impedance conditions.

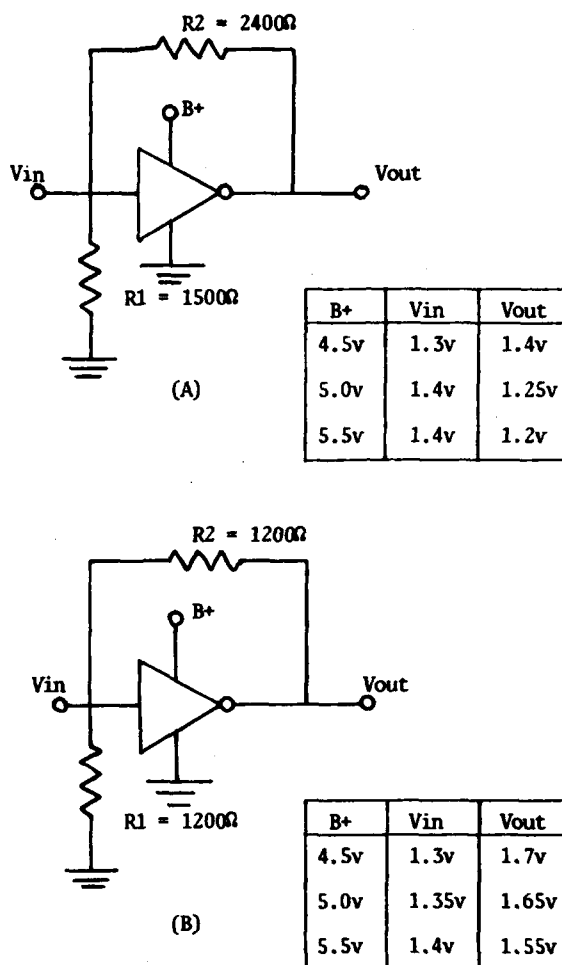
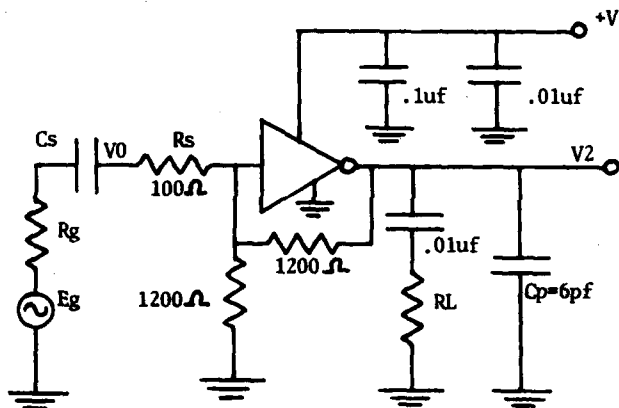
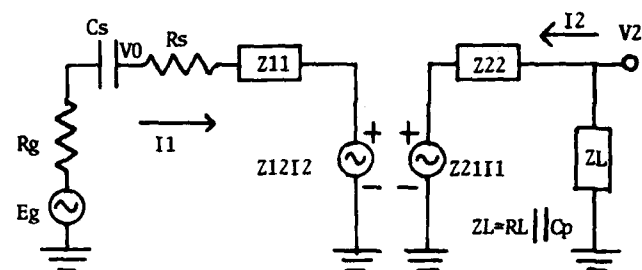


FIGURE 9
TWO PROPOSED METHODS OF AMPLIFIER BIAS

Figure 13 shows the results for the unequal bias resistor amplifier. Here, phase angle prediction is excellent with about a 10% error in magnitude prediction. The general shape of both the magnitude curves is very much the same indicating reasonably good correlation of data. With Z-parameter models verified, computer analysis of the oscillating loop can be done.



Physical Circuit



Equivalent Network

$$\frac{V_2}{V_0} = \frac{Z_{21} \cdot Z_L}{(Z_L + Z_{22})(R_s + Z_{11}) - Z_{12} \cdot Z_{21}}$$

Figure 10
Test Circuit for Z-parameters

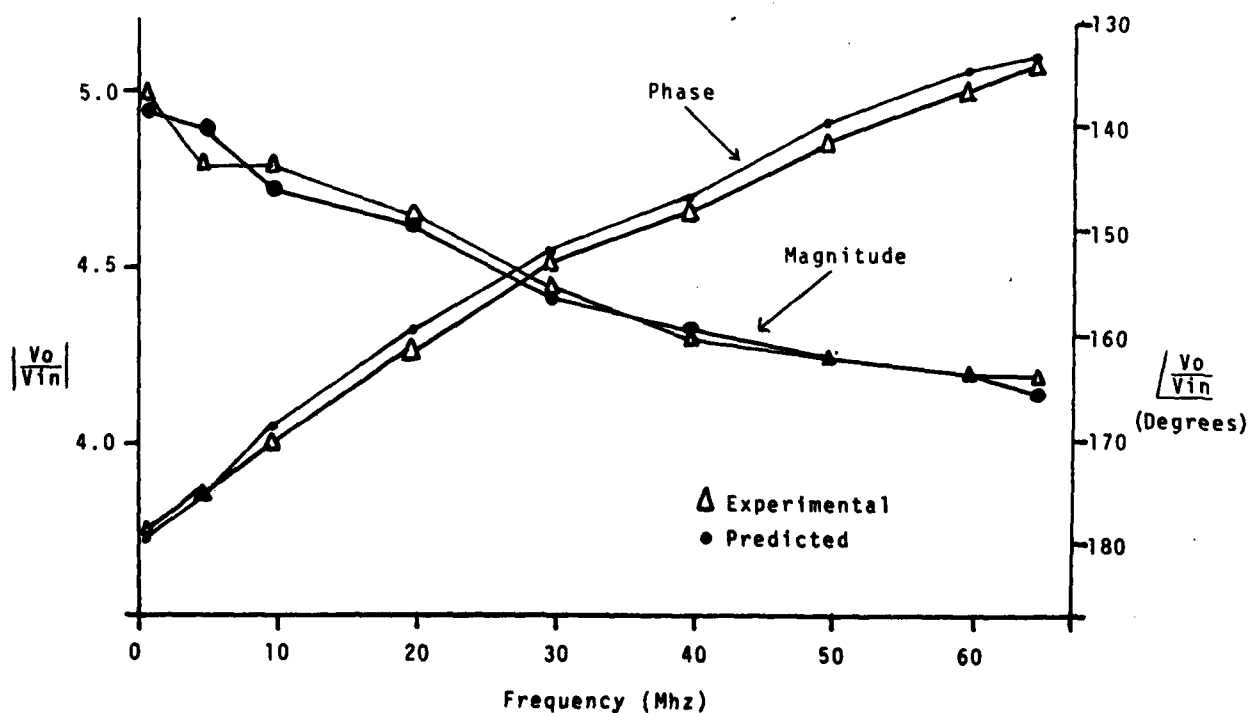


Figure 11
Predicted and Experimental Results of Z-parameter Test
Circuit With 1200 Ω Bias Resistors and 470 Ω Load

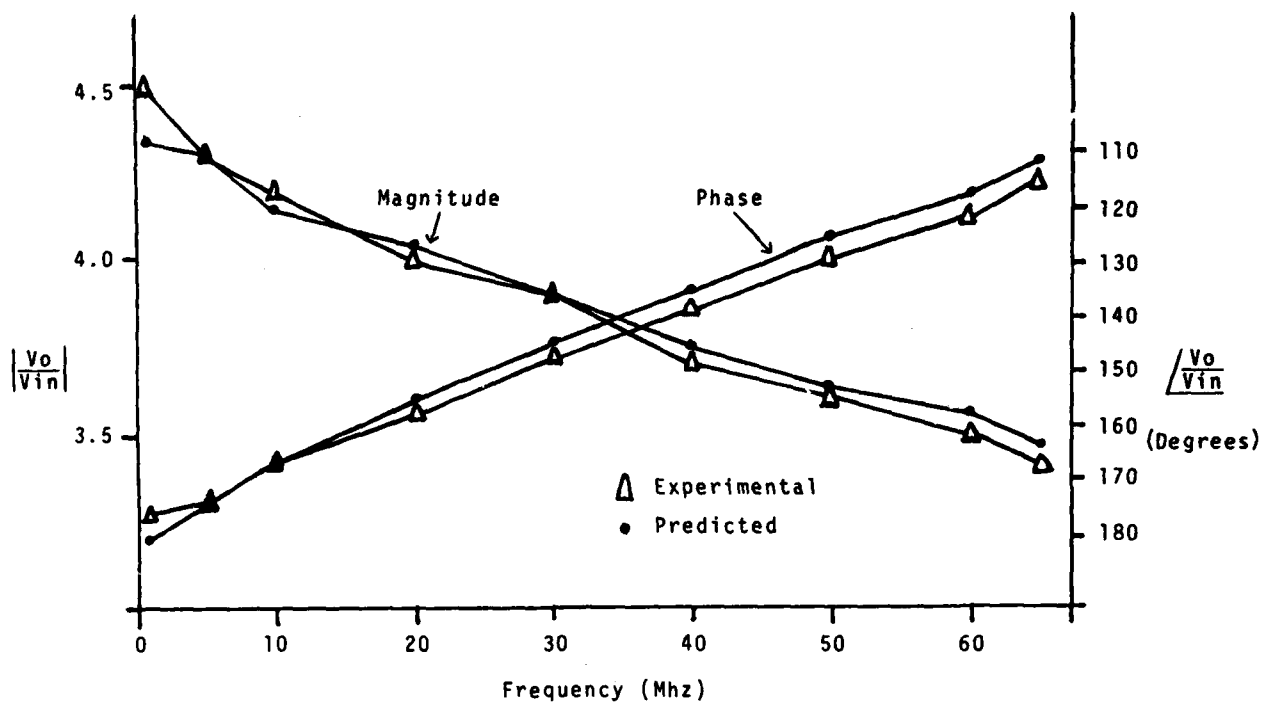


Figure 12

Predicted and Experimental Results of Z-parameter Test
Circuit With 1200 Ω Bias Resistors and 47 Ω Load

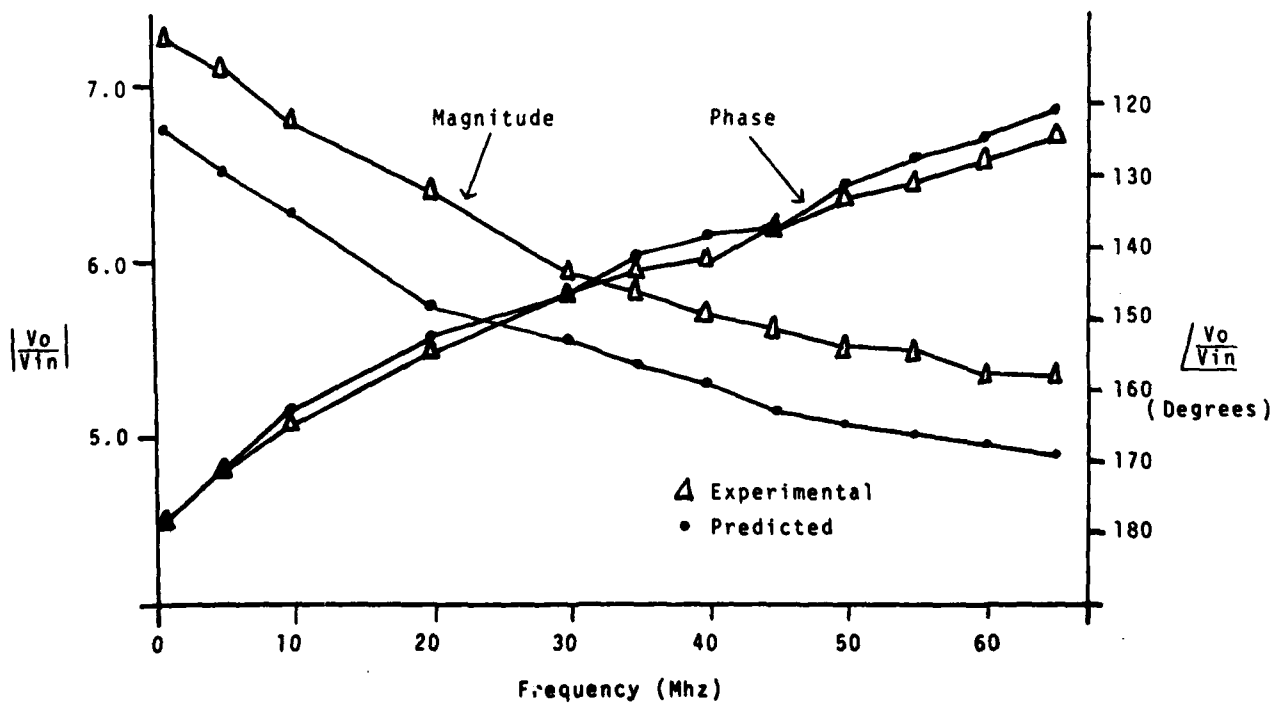


Figure 13

Predicted and Experimental Results of Z-parameter Test
Circuit With Unequal Bias Resistors and 470 Ω Load

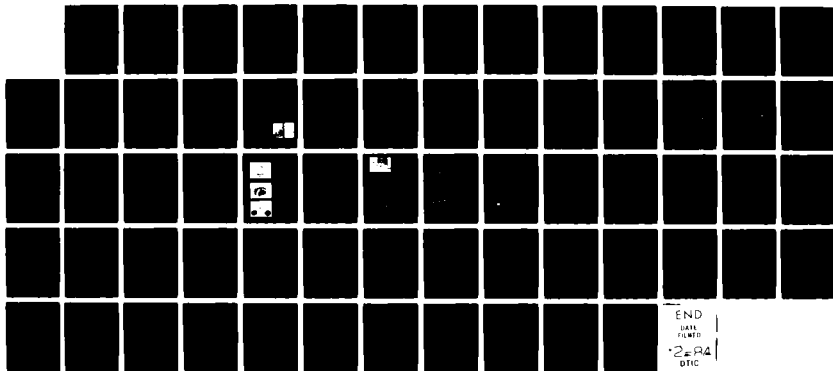
AD-A136 673

PROCEEDINGS OF THE ANNUAL SYMPOSIUM ON FREQUENCY
CONTROL (37TH) 1-3 JUNE 1..(U) ARMY ELECTRONICS
RESEARCH AND DEVELOPMENT COMMAND FORT MONMOUTH 1983

F/G 9/1

NL

UNCLASSIFIED





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Loop Optimization

Figure 14 shows the twin-tee network utilized. A general twin-tee filter has 6 components or 6 variables to be optimized. This many variables is impractical to optimize simultaneously, so the Figure 14 network was devised with three variables, R, C and K. K is an unbalancing factor which allows a variable degree of dissymmetry about a central point in the network. This allows one-half of the network to be at one impedance level while the other half is scaled to a different impedance level. This will allow an optimum network to be solved for that will not be significantly inferior to one which has all six possible variables optimized at once. Furthermore, if the frequency of the notch is fixed at one-third the oscillating frequency, any R value will have only one possible value of C as the following notch frequency equation indicates:⁴

$$\omega_{\text{notch}} = \frac{1}{RC}$$

This reduces the filter variables to two, R and K.

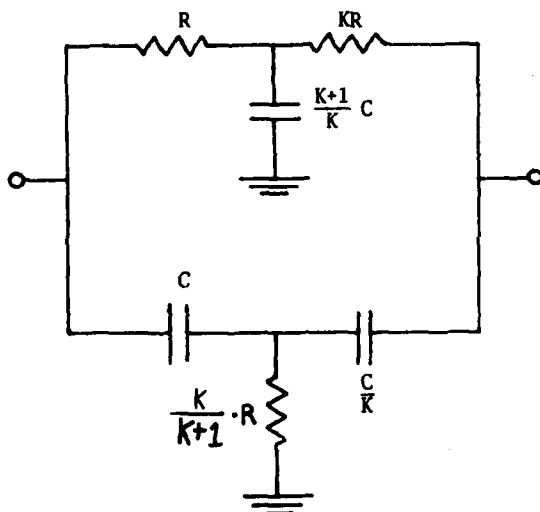


FIGURE 14
TWIN-TEE NETWORK

Figure 15 is an ac equivalent model for the oscillator of Figure 2. The inverter has been replaced with its Z-parameter equivalent model, the monolithic resonator replaced with an ac equivalent circuit and the output driver is represented as a load of R_L and C_L on the inverter's output. In order to find the best values for R and K, an open loop technique was used. The Z2111 generator was replaced with a test generator, Z21IT. This effectively removes the feedback and opens the oscillator loop. Since I_T and I_I are the same current when the loop is closed, os-

cillations can be predicted by searching for the frequency where the phase of I_T and I_I are the same and the magnitude ratio of I_I to I_T is greater than 1. A computer was used to search through many values of R and K and solve for the frequency of oscillation (if one exists) for each set of values. For many values of R in a given range, the computer determined what value of K gave the most excess loop current gain (I_I/I_T) at the zero phase frequency. Then the value of C, K, frequency, I_I/I_T magnitude and I_I/I_T phase (a number very close to zero) is printed out for each R value. For analysis of the Figure 15 circuit, a 40 MHz crystal resonator was used with the following parameters:

F1 = 40.084789 MHz (symmetric response)
F2 = 40.090358 MHz (anti-symmetric response)
RM1 = 89 ohm
RM2 = 84 ohm
CM1 = .000578 pfd
CM2 = .000568 pfd
LM1 = 27.263 MH
LM2 = 27.747 MH
Co1 = 2.5 pf
Co2 = 2.5 pf
K = .000138901

40 MHz Z-parameter values for the 2400 ohm, 1500 ohm biased amplifier were used from Table 2. R_L = 256 ohms and C_L = 25.8 pf. These values were determined from input impedance tests run on inverters biased in the configuration of the output driver of Figure 2. Table 3 is a portion of the output table of this computer run. Observing the table indicates that the largest value of open loop current gain occurs when R = 140 ohms and K = 1.3 at which point the current gain is 1.747.

For increased gain, the circuit of Figure 16 is proposed. Here are two inverter amplifiers cascaded driving the notch filter and monolithic resonator. Now the two amplifiers together provide an approximate zero degree total phase shift and the monolithic resonator provides feedback and frequency control on its third overtone anti-symmetric mode of response. The notch filter is present to stop fundamental oscillations and lead compensate the amplifiers. This circuit was optimized in the same manner as the previous example. This time the 1200 ohm biased amplifier Z-parameters were used along with the same crystal parameters. Table 4 shows the computer output generated for this circuit. Best case current gain is 10.85 or 20.7 db and occurs at two places; K = 1.5, R = 200 ohm and K = 1.3, R = 220 ohm. This circuit has considerably more gain than the single inverter oscillator and should be a much more reliable oscillator.

Fundamental Mode Attenuation

Obviously a most important characteristic of these oscillators is that they never operate on the fundamental mode of the crystal. Analysis of the fundamental mode attenuation of these oscillators is possible by a modification of the computer program used for third overtone loop optimization. For analysis, the two inverter oscillator of Figure 16 was used. In this case the amplifiers were biased with R_2 = 2400 ohms and R_1 = 1500 ohms, this being the higher gain of the two amplifier versions characterized which should be the most susceptible to fundamental oscillations. First the circuit was optimized for third overtone operation using the previously introduced methods. The same 40 MHz crystal parameters were used with the notch filter again tuned to 13.3 MHz. Maximum open loop current gain is 11.11 and occurs when R = 300 ohms and K = 1. By performing an open loop current gain test

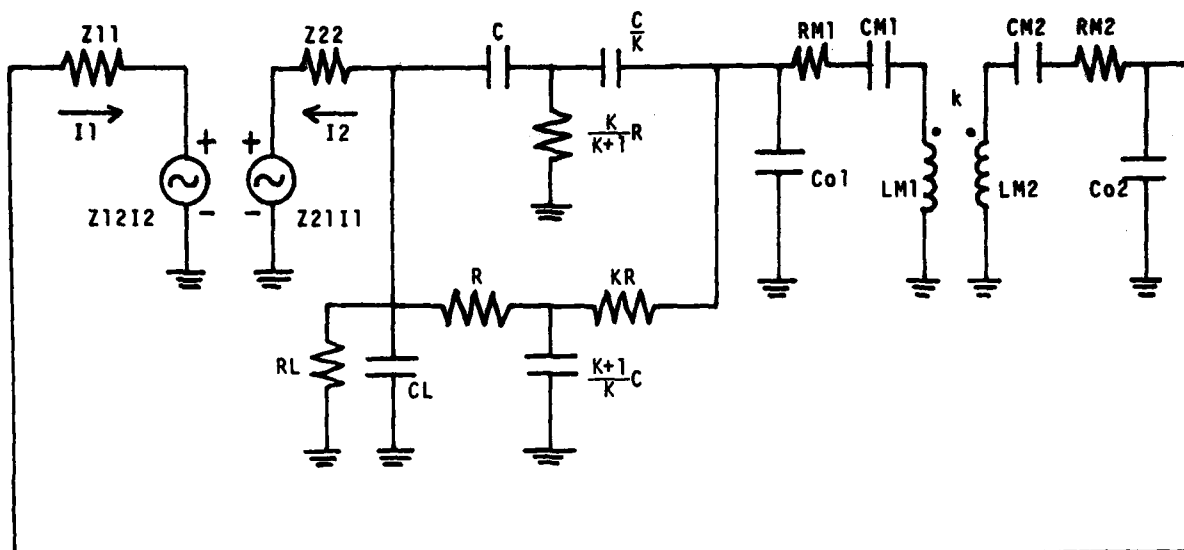


Figure 15

A-C Equivalent of Twin-Tee Inverter Oscillator

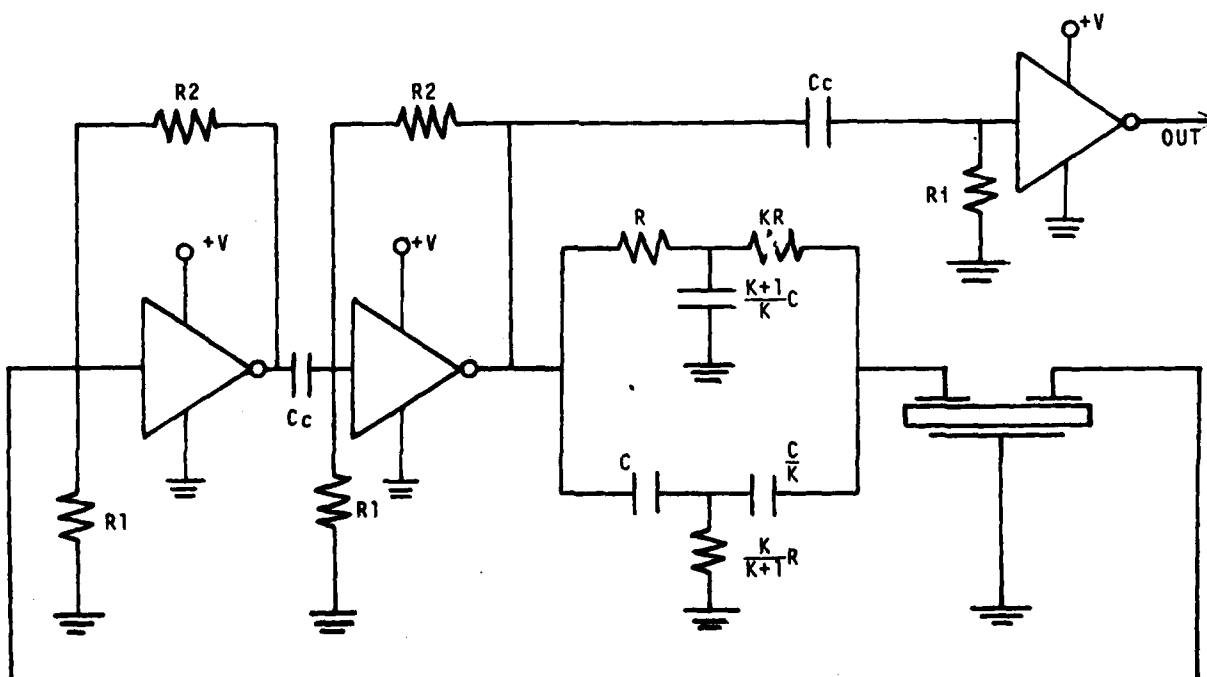


Figure 16

Two Inverter Twin-Tee Oscillator

at the fundamental crystal frequency, the probability of fundamental mode oscillations can be determined. A frequency sweep through the fundamental response of the circuit was done yielding the curves of Figure 17. The fundamental parameters of the same 40 MHz crystal were used and are as follows:

F1 = 13.366090 MHz
 F2 = 13.399720 MHz
 RM1 = 20 ohms
 RM2 = 20 ohms
 CM1 = .005196 pf
 CM2 = .005182 pf
 LM1 = 27.220 mH
 LM2 = 27.291 mH
 Co1 = 2.5 pf
 Co2 = 2.5 pf
 K = .002512886

The 13 MHz Z-parameters were linearly interpolated from Table 2. Figure 17 shows that at the fundamental response peak the current gain is .1124 for a loss of 18.9 db. At this point, the loop phase is 49 degrees off from zero phase, further guaranteeing the absence of oscillations. At the point where the phase angle is zero, the gain is even less, .085, or -21.4 db. Comparing this value of circuit gain to the value obtained for third overtone operation (+20.89 db), indicates that the circuit has 42.29 db more gain at the third overtone zero phase frequency than at the fundamental zero phase frequency. Based on this analysis, fundamental oscillations would be extremely unlikely for the circuits proposed here.

Empirical Tests Of Proposed Oscillators

Both the single inverter oscillator using the unequal bias resistors and the two inverter oscillator using the 1200 ohm bias resistors have been constructed and tested. The same 40 MHz crystal was used whose equivalent circuit parameters were used in the computer optimization. The chosen twin-tee notch filter R, C, and K values were used. Table 5 gives test results for the single inverter oscillator using the 2400 ohm, 1500 ohm biased amplifier. It can be seen that this oscillator provides the appropriate TTL output wave and frequency as required over the 0° to +70°C temperature range. Table 6 is the performance for the two inverter oscillator with 1200 ohm bias resistors. Again, this data demonstrates acceptable performance for a third overtone TTL clock oscillator.

Conclusions

It has been demonstrated that biased TTL inverter amplifiers can be represented by linear models for small signal circuit analysis. Techniques have been developed for the open loop analysis of clock oscillators that utilize these amplifiers. The Schottky inverter has been shown to have sufficient gain at high enough frequencies to be used in overtone crystal applications. The twin-tee network has been shown to be capable of preventing fundamental oscillations and correcting loop phase error caused by amplifier propagation delay.

References

1. Patent No. 4044317, CTS Corporation
2. T. Luxmore and D.E. Newell, "The MCO-Monolithic Crystal Oscillator", Proc. 31st Ann. Symposium on Freq. Control; pp. 396-399 (1977).
3. D. Nehring, "Third Overtone TTL Crystal Oscillator for Hybrid Applications", unpublished Masters Thesis, Northern Illinois University; (1982).
4. John D. Ryder, "Networks, Lines and Fields", 2nd edition, Prentice-Hall, Englewood Cliffs, NJ; pp. 12-14 (1955).

TABLE 1
Z-PARAMETER VALUES FOR AMPLIFIER WITH R1, R2
EQUAL 1200Ω, B+ EQUALS 5.0 VOLTS

FREQ. (MHz)	Z11 (MAGNITUDE) (PHASE)	Z12 (MAGNITUDE) (PHASE)	Z21 (MAGNITUDE) (PHASE)	Z22 (MAGNITUDE) (PHASE)
1.000	0.9100E+02 1.00	0.8951E+00 -8.44	0.9616E+03 179.62	0.3000E+01 0.00
5.000	0.9200E+02 -1.00	0.9045E+00 7.70	0.9529E+03 174.58	0.3000E+01 19.00
10.000	0.9300E+02 -2.00	0.8959E+00 18.46	0.9256E+03 168.30	0.3600E+01 36.00
20.000	0.9300E+02 -6.00	0.1039E+01 37.12	0.9039E+03 157.61	0.5000E+01 59.00
30.000	0.9000E+02 -11.00	0.1150E+01 53.43	0.8425E+03 148.34	0.6800E+01 72.00
40.000	0.8400E+02 -14.00	0.1377E+01 66.72	0.7883E+03 141.40	0.8800E+01 81.00
50.000	0.7700E+02 -16.00	0.1598E+01 76.35	0.7343E+03 134.92	0.1100E+02 85.00
60.000	0.7100E+02 -17.00	0.1841E+01 87.55	0.6897E+03 130.22	0.1300E+02 90.00
65.00	0.6800E+02 -18.00	0.2109E+01 92.50	0.6602E+03 127.79	0.1450E+02 94.00

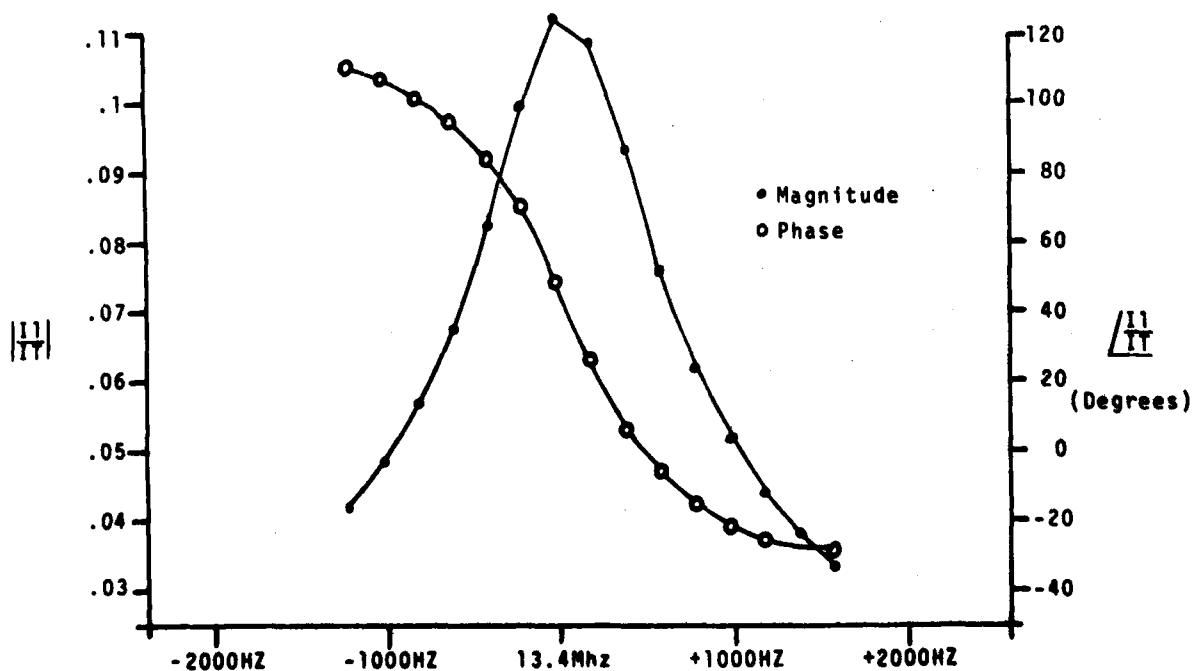


Figure 17 Fundamental Mode Characteristics of Inverter Oscillator With Twin-Tee Network

TABLE 2

Z-PARAMETER VALUES FOR AMPLIFIER WITH
 $R_1 = 1500\Omega$, $R_2 = 2400\Omega$, $B+ = 5.0$ VOLTS

FREQ.	Z11	Z12	Z21	Z22
(MHZ)	(MAGNITUDE) (PHASE)	(MAGNITUDE) (PHASE)	(MAGNITUDE) (PHASE)	(MAGNITUDE) (PHASE)
1.000	0.1600E+03 1.00	0.8253E+00 -10.07	0.1786E+04 178.34	0.3500E+01 7.00
5.000	0.1600E+03 -6.00	0.8066E+00 7.34	0.1719E+04 168.92	0.3800E+01 20.00
10.000	0.1570E+03 -11.00	0.8331E+00 18.12	0.1636E+04 158.01	0.4200E+01 36.00
20.000	0.1420E+03 -22.00	0.9544E+00 36.17	0.1383E+04 142.74	0.6000E+01 57.00
30.000	0.1250E+03 -29.00	0.1197E+01 49.34	0.1219E+04 131.73	0.8000E+01 69.00
35.000	0.1160E+03 -32.00	0.1303E+01 55.03	0.1128E+04 126.65	0.9500E+01 72.00
40.000	0.1060E+03 -34.00	0.1440E+01 59.72	0.1047E+04 122.87	0.1000E+02 75.00
45.000	0.1000E+03 -37.00	0.1568E+01 63.25	0.9769E+03 122.08	0.1200E+02 78.00
50.000	0.8800E+02 -37.00	0.1716E+01 67.75	0.8955E+03 117.35	0.1300E+02 82.00
55.000	0.8200E+02 -38.00	0.1879E+01 71.48	0.8491E+03 114.57	0.1400E+02 84.00
60.000	0.7500E+02 -38.00	0.2069E+01 74.15	0.8046E+03 113.02	0.1500E+02 87.00
65.000	0.6800E+02 -38.00	0.2292E+01 77.90	0.7573E+03 109.36	0.1700E+02 89.00

TABLE 3

OPEN LOOP RESPONSE OF SINGLE INVERTER OSCILLATOR
WITH R1 = 1500Ω AND R2 = 2400Ω

<u>R</u>	<u>C</u>	<u>K</u>	<u>FREQ.</u>	<u>I1/IT(MAG)</u>	<u>I1/IT(PHASE)</u>
0.4000E+02	0.2984E-09	0.3500E+01	0.40084040D+08	0.7361E+00	-0.1631E-01
0.6000E+02	0.1989E-09	0.1800E+01	0.40084500D+08	0.1368E+01	-0.9354E-01
0.8000E+02	0.1492E-09	0.1700E+01	0.40084620D+08	0.1609E+01	-0.9457E-01
0.1000E+03	0.1194E-09	0.1400E+01	0.40084688D+08	0.1705E+01	-0.4440E-02
0.1200E+03	0.9947E-10	0.1600E+01	0.40084724D+08	0.1743E+01	-0.2761E+00
0.1400E+03	0.8526E-10	0.1300E+01	0.40084760D+08	0.1747E+01	-0.2269E+00
0.1600E+03	0.7460E-10	0.1400E+01	0.40084780D+08	0.1738E+01	-0.2832E+00
0.1800E+03	0.6631E-10	0.1300E+01	0.40084800D+08	0.1717E+01	-0.2696E+00
0.2000E+03	0.5968E-10	0.1300E+01	0.40084816D+08	0.1691E+01	-0.1503E+00
0.2200E+03	0.5426E-10	0.1300E+01	0.40084828D+08	0.1662E+01	-0.7224E-01
0.2400E+03	0.4974E-10	0.1200E+01	0.40084844D+08	0.1638E+01	-0.2412E+00
0.2600E+03	0.4591E-10	0.1200E+01	0.40084852D+08	0.1605E+01	0.6935E-01
0.2800E+03	0.4263E-10	0.1200E+01	0.40084864D+08	0.1581E+01	-0.1312E+00
0.3000E+03	0.3979E-10	0.1200E+01	0.40084872D+08	0.1552E+01	0.3037E-01

TABLE 4

OPEN LOOP RESPONSE OF TWO INVERTER OSCILLATOR
WITH R1 = R2 = 1200Ω

<u>R</u>	<u>C</u>	<u>K</u>	<u>FREQ.</u>	<u>I1/IT(MAG)</u>	<u>I1/IT(PHASE)</u>
0.8000E+02	0.1492E-09	0.2100E+01	0.40089868D+08	0.8236E+01	0.1330E-01
0.1000E+03	0.1194E-09	0.1800E+01	0.40089972D+08	0.9293E+01	-0.8726E-01
0.1200E+03	0.9947E-10	0.1800E+01	0.40090036D+08	0.9965E+01	-0.8846E-01
0.1400E+03	0.8526E-10	0.1500E+01	0.40090092D+08	0.1040E+02	-0.1121E+00
0.1600E+03	0.7460E-10	0.1500E+01	0.40090128D+08	0.1065E+02	-0.1157E+00
0.1800E+03	0.6631E-10	0.1400E+01	0.40090160D+08	0.1078E+02	-0.2399E-01
0.2000E+03	0.5968E-10	0.1500E+01	0.40090184D+08	0.1085E+02	-0.2057E+00
0.2200E+03	0.5426E-10	0.1300E+01	0.40090208D+08	0.1085E+02	-0.6669E-01
0.2400E+03	0.4974E-10	0.1200E+01	0.40090228D+08	0.1081E+02	-0.1059E+00
0.2600E+03	0.4591E-10	0.1300E+01	0.40090244D+08	0.1074E+02	-0.1467E-01
0.2800E+03	0.4263E-10	0.1200E+01	0.40090260D+08	0.1067E+02	-0.1643E+00
0.3000E+03	0.3979E-10	0.1400E+01	0.40090276D+08	0.1054E+02	-0.2054E+00

Table 5
Performance of Single Inverter Oscillator
With Monolithic Crystal

Results at 0°C			
Input Voltage	4.5 volts	5.0 volts	5.5 volts
Frequency (Mhz)	40.084539	40.084417	40.084363
1 Level (v)	3.0	3.7	4.7
0 Level (v)	.3	.3	.3
Symmetry	60/40	55/45	56/47
Rise Time (ns)	3.6	2.5	2.2
Fall Time (ns)	3.8	2.2	1.8
Results at +25°C			
Input Voltage	4.5 volts	5.0 volts	5.5 volts
Frequency (Mhz)	40.084399	40.084288	40.084344
Current (ma)	36.5	43.3	48.4
1 Level (v)	3.3	4.3	5.0
0 Level (v)	.2	.2	.2
Symmetry	54/46	53/47	54/46
Rise Time (ns)	3	2.4	2.1
Fall Time (ns)	3	1.9	1.7
Results at +70°C			
Input Voltage	4.5 volts	5.0 volts	5.5 volts
Frequency (Mhz)	40.084375	40.084454	40.084492
1 Level (v)	3.6	4.4	5.0
0 Level (v)	.2	.3	.3
Symmetry	52/48	54/46	56/44
Rise Time (ns)	3.8	3.0	2.5
Fall Time (ns)	2.6	2.0	1.7

Table 6

Performance of Two Inverter Oscillator With
Monolithic Crystal and $R_2, R_1 = 1200 \Omega$

Results at 0°C			
Input Voltage	4.5 volts	5.0 volts	5.5 volts
Frequency (Mhz)	40.090181	40.090208	40.090215
1 Level (v)	3.1	3.6	4.0
0 Level (v)	.2	.2	.2
Symmetry	44/56	54/46	53/47
Rise Time (ns)	3.4	2.6	2.4
Fall Time (ns)	2.5	2.0	1.7
Results at +25°C			
Input Voltage	4.5 volts	5.0 volts	5.5 volts
Frequency (Mhz)	40.090073	40.090096	40.090135
Current (ma)	33.6	41.7	48.1
1 Level (v)	3.3	3.9	4.3
0 Level (v)	.2	.2	.2
Symmetry	53/47	51/49	50/50
Rise Time (ns)	3.4	2.6	2.2
Fall Time (ns)	2.4	1.9	1.5
Results at +70°C			
Input Voltage	4.5 volts	5.0 volts	5.5 volts
Frequency (Mhz)	40.090188	40.090202	40.090254
1 Level (v)	3.2	3.7	4.2
0 Level (v)	.2	.2	.2
Symmetry	51/49	50/50	52/48
Rise Time (ns)	4.0	2.4	1.9
Fall Time (ns)	2.4	1.6	1.3

MSSW DELAY LINE BASED OSCILLATORS

V. Lander* and J.P. Parekh**

*Airtron Corporation, Division of Litton, 200 East Hanover Avenue,
Morris Plains, N.J. 07950

**Department of Electrical Engineering, State University of New York
at Stony Brook, Stony Brook, N.Y. 11794

ABSTRACT

The present paper reports an experimental study of magnetostatic-wave (MSW) delay-line based oscillators utilizing magnetostatic surface waves (MSSW's) in epitaxial YIG films. A novel approach at narrowbanding the delay line was realized through the use of lifted transducers. It is found that the tuning range for single-mode operation is limited primarily by the electrical length of the circuit external to the delay line with frequency jumping occurring everytime this electrical length changes by 360° . A maximum tuning range of about 600 MHz was obtained, with this range limited primarily by the AVANTEK GaAs FET amplifier that was used whose electrical length was equivalent to about 45cm of air transmission line. It is concluded that, if the electrical length of the external circuit is sufficiently reduced (ideally made zero), single-mode operation over an octave bandwidth or more is possible.

1. INTRODUCTION

An important application area for the emerging technology based on magnetostatic waves (MSW's) in epitaxial YIG films is that of tunable microwave oscillators operating directly within the band 2 GHz to 20 GHz. Since its inception, the area of tunable oscillators has been dominated by YIG-sphere resonators through their deployment as the frequency-control element within the oscillator feedback loop. The potential for MSW oscillators to compete with the well-entrenched YIG-sphere oscillators, or to carve out a market niche of their own, has been well recognized. An attractive property of MSW oscillators is their superior phase-noise characteristics to the YIG-sphere oscillator. MSW oscillators are thus a potential candidate for applications in communication and electronic warfare where further improvement in the spectral purity of microwave sources is an important requirement. Another attractive feature of MSW oscillators is the relative simplicity of the planar technology that is the basis of fabrication of the constituent MSW frequency-control feedback element, i.e., a delay line or a resonator. This advantageous construction feature of MSW oscillators contrasts with the complex alignment problem present in the manufacture of YIG-sphere oscillators.

In the original studies of MSW oscillators,¹⁻³ the MSW device was a wideband delay line which was chosen over a resonator for the obvious reason of simplicity in its construction. The experiments, however, revealed that, as the oscillator was tuned via change of bias field, the undesirable phenomena of multimoding and mode hopping occurred which limited the range of tunability to the frequency range between any two mode hops, typically in the range of few tens of MHz or thereabouts. While the frequency spacing between mode hops, in agreement with expectation, could be increased by reducing the delay time of the delay line, for a wideband delay line the problem of mode hopping was observed to be a serious one.

The use of delay lines employing multiple-element microstrip transducers^{4,5} as a means of suppressing the mode hopping problem has also been reported. With this narrowbanding approach, single-mode oscillator operation was extended to a frequency band of about 100 MHz or more but the problem of mode hopping was not removed.

The difficulties associated with delay-line based MSW oscillators led generally to an abandonment of this approach in favour of a resonator-based approach,^{1,6} with ion-milled groove reflectors being the favoured means of implementing high-Q resonators. The recent realization of a magnetostatic-forward-volume-wave (MSFVW) resonator-based oscillator with single-mode operation over the 3.5 GHz to 7 GHz octave range represents an important achievement in establishing the viability of tunable MSW oscillators.

The present paper reports significantly improved and encouraging single-mode MSW delay-line based oscillator

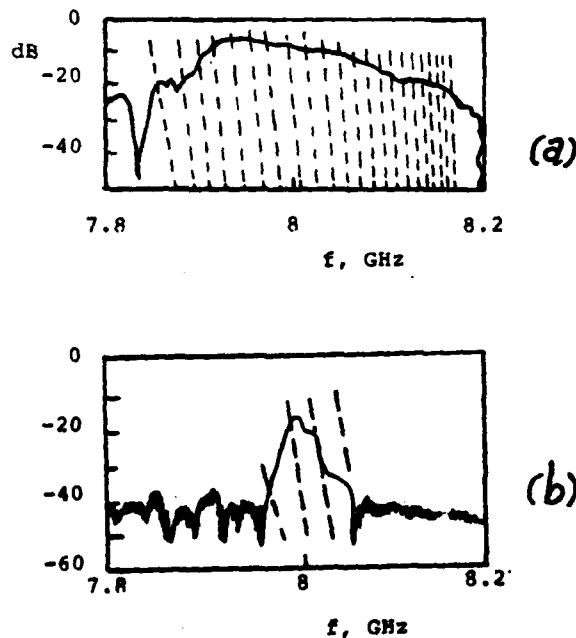


Fig. 1. Insertion loss and phase characteristics of an MSSW delay line: (a) with microstrip transducers directly on the YIG film; (b) with microstrip transducers lifted 500 μ m above the YIG film. The solid curves represent insertion loss in dB (with scale at the left) and the dashed curves represent phase.

results obtained with magnetostatic surface waves (MSSW's). A re-examination of delay-line based oscillators was undertaken for two reasons: (i) delay lines are much easier to fabricate than resonators; and (ii) a detailed understanding of the cause of mode hopping and the resulting limitation on tuning range did not appear to be available in the literature. In the present experiment, single microstrip transducers were used and a novel approach at narrowbanding the delay line was realized through the use of lifted transducers. The experiment was performed over the frequency range of 5 GHz to 10 GHz.

The narrowbanding of the delay line arrested the multimoding and mode hopping problems but a new problem arose instead. This was that the oscillator frequency exhibited a "jump" with tuning, with the jump appearing with a more or less constant period in frequency. This jump phenomenon has been determined to be related

to the electrical length of the circuit (including the amplifier) that is external to the delay line, with the frequency interval between jumps becoming shorter if this electrical length is increased. It has been found that the frequency jump occurs every time the electrical length of the external circuit changes by 360° . In the experiment, single-mode operation was realized over a 600 MHz range, with this range limited primarily by the Avantek AMT-12017 GaAs FET amplifier that was used which, with an electrical length of about 45cm of air transmission line, was the longest element in the external circuit. An important conclusion drawn in the present work is that, if the electrical length of the external circuit is sufficiently reduced (ideally made zero), single-mode oscillator operation over an octave bandwidth or more is possible.

II. MSSW DELAY-LINE NARROWBANDING VIA LIFTED TRANSDUCERS

A characteristic feature of MSSW's is that their fields decay exponentially away from the YIG film. This feature implies a decoupling in the electromagnetic excitation of MSSW's with a lifting of the exciting microstrip transducer above the YIG-film surface. The level of decoupling is determined by the electrical spacing kh of the transducer from the YIG-film surface rather than the mechanical spacing h . Now the MSSW wavenumber k increases monotonically with frequency, ranging from a value of zero at the low-frequency bound of the MSSW spectrum to a value of infinity at the high-frequency bound of the spectrum. It follows that, although the transducer may be mechanically spaced from the YIG-film surface, electrically it is as good as on the YIG-film surface at the low-frequency bound of the MSSW spectrum, with the electrical spacing increasing monotonically and at an increasing pace with increase in frequency. Correspondingly, the MSSW excitation efficiency of a transducer is independent of h at the low-frequency bound of the MSSW spectrum, with a decoupling in excitation becoming increasingly more pronounced as the frequency is raised. In the context of an MSSW delay line with transducers conveniently lifted by the same amount h , one expects the delay-line passband to become narrower with increase in h and to be situated at the low-frequency bound (or the low- k regime) of the MSSW spectrum, with the delay-line passband insertion loss to be unaffected by change in h .

The foregoing qualitative prediction of narrowbanding via the use of lifted transducers is to a large measure borne out experimentally. In Fig. 1, the insertion loss and phase characteristics of an MSSW delay line with 2-mil wide shorted single-microstrip transducers defined on a 10-mil thick Cu-clad Duroid substrate are illustrated for two positions of transducer placement, i.e., transducers placed directly on YIG film yielding a wideband delay line (Fig. 1a) and transducers lifted 500 μ m above the YIG film yielding a narrowband delay line (Fig. 1b). In Fig. 1b, the passband was centered at about 8 GHz through a change in bias field from that for Fig. 1a. The crowding of the phase slopes with increasing frequency in Fig. 1a is indicative of the dispersive character of MSSW's. A few phase slopes are present within the passband of Fig. 1b but these can be reduced to a single one by further lifting the transducers as was done in the remainder of the experiment.

In Fig. 2, the insertion loss and phase characteristics of a narrowband MSSW delay line (narrowbanded via the use of lifted transducers) are shown as the device is tuned over an octave bandwidth (5 GHz to 10 GHz). A conspicuous feature of this figure is that the amplitude and phase responses of the single mode existing within the passband of the delay line tune in synchronism with each other as the bias field is changed. The locking of amplitude and phase responses with tun-

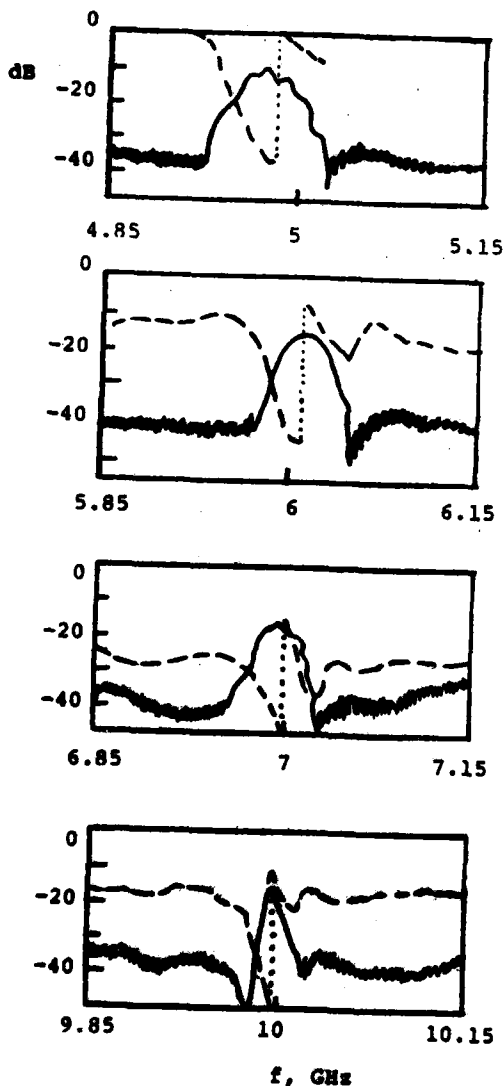


Fig. 2. Illustrating the synchronization with tuning of the amplitude and phase responses of an MSSW delay line with lifted transducers. The solid and dashed curves have the same meaning as in Fig. 1. Notice that the narrowbanding becomes more pronounced as the device is tuned to higher frequencies.

ing is an important observation that provides a clear indication that single-mode operation without mode hopping or frequency jumping is in principle possible over an octave bandwidth or more. The maximum frequency range over which synchronization of amplitude and phase responses of a delay line alone is obtained with tuning has not been fully explored at this time.

III. OSCILLATOR CHARACTERISTICS

The major limitation to the tuning range of the oscillator for single-mode operation arises from the fact that the circuit external to the delay line has a finite length. In order to minimize this length, the directional coupler was removed from the oscillator loop by fabricating it directly on the Duroid substrate on which the delay-line microstrip transducers were defined. Thus, the overall device was a three-port with two ports constituting the input and output ports of the delay line and the third port constituting the output port of the directional coupler. The circuit external to the delay line was, therefore, comprised of three elements: (i) the delay line with the YIG

film removed from it; (ii) the external cable; and (iii) the amplifier.

Although the amplitude and phase responses of the delay line by itself are locked into synchronism as the device is tuned (see Fig. 2), the open-loop response of the oscillator exhibits desynchronization of the amplitude and phase responses with device tuning. A systematic evaluation of this desynchronization process was undertaken and is illustrated in Figs. 3 and 4.

In Fig. 3, the oscillator open-loop response with the YIG film physically removed from the delay line is given for three different lengths of the external cable, i.e., 2 inches (Fig. 3a), 4 inches (Fig. 3b) and 2 feet 6 inches (Fig. 3c). It is seen that the frequency range over which a phase change of 360° takes place is slightly under 500 MHz for the case of an external cable of length 2 inches, and that this frequency range goes down with increase in cable length. It has been found that the frequency range over which this open-loop phase changes by 360° coincides with the frequency interval between frequency jumps in the oscillator.

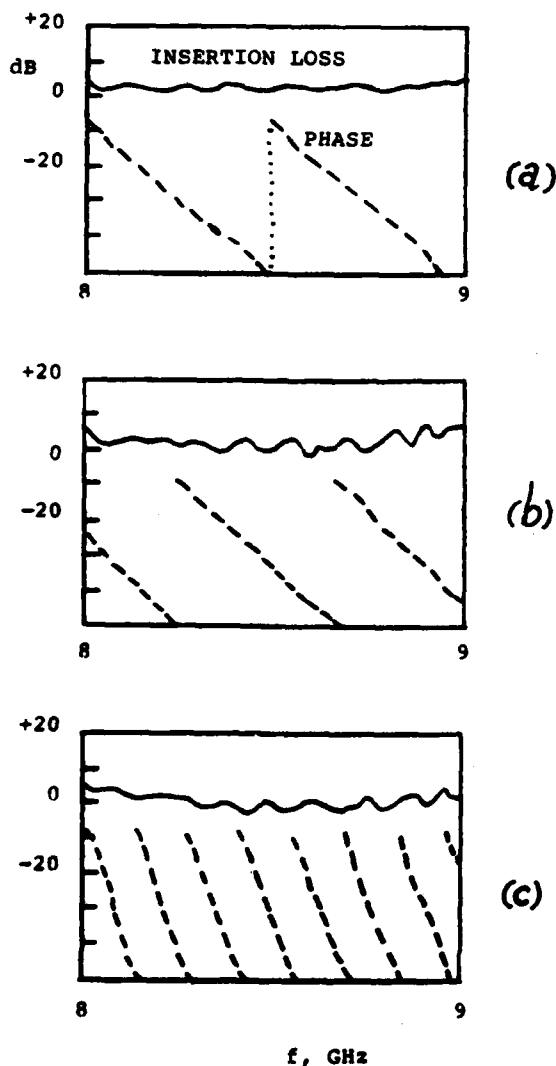


Fig. 3. Oscillator open loop response with the YIG film removed from the delay line. The measurements are for an external cable length of: (a) 2"; (b) 4"; and (c) 2'6". The solid and dashed curves have the same meaning as in Fig. 1.

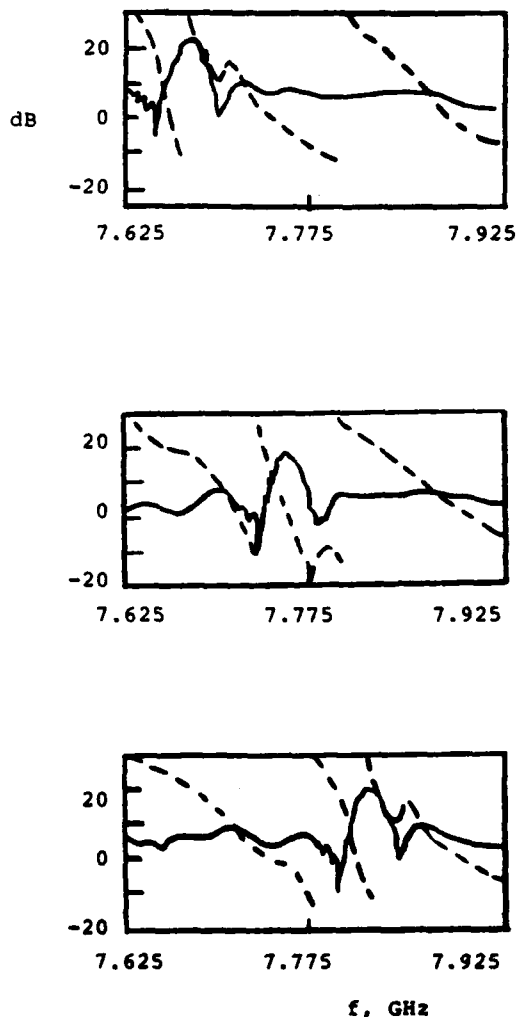


Fig. 4. Illustrating the desynchronization with tuning of the oscillator open-loop amplitude and phase responses with the YIG film in place in the delay line. The measurements are for a fixed external cable length of 2'6". The solid and dashed curves have the same meaning as in Fig. 1.

The amplifier used in the experiment was an Avantek AMT-1207 GaAs FET amplifier (50 dB gain over the 7 GHz to 12 GHz range). The amplifier phase changes by 360° approximately every 800 MHz so that, if the length of the external circuit other than the amplifier were made zero, frequency jumping would occur every 800 MHz. The amplifier delay line is about 1.5ns which is approximately equal to the delay time of an air transmission line of length 45cm. The amplifier constitutes the longest element in the circuit external to the delay line.

In Fig. 4, the measured open-loop response obtained with the YIG film put back into the delay line is given for the case of a 2 feet 6 inch long external cable. As the device is tuned, the phase response lags on the amplitude response, with the phase slope of a given mode at some point moving out of the amplitude response and a new phase slope coming within the amplitude response. It is at the latter point that frequency jumping occurs.

With minimization of the external cable length, the largest value of the frequency interval between frequency jumps was obtained to be about 600 MHz. The tuning response of the oscillator representing the variation of oscillator frequency with bias field is essentially linear as shown in Fig. 5. The measurements in this figure are for an external cable of

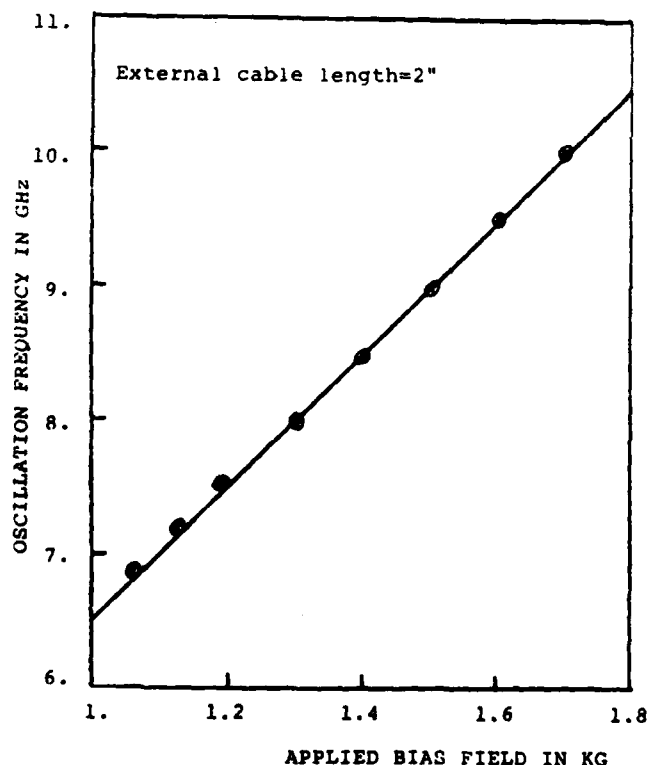


Fig. 5. MSSW oscillator tuning curve for fixed external cable length of 2".

length 2 inches. Frequency jumping which occurs at intervals of about 500 MHz for this cable length is not shown in Fig. 5.

IV. CONCLUSION

The present paper has reported an experimental study of an MSSW delay-line based oscillator. A novel approach in the present work has been the use of lifted transducers for narrowbanding the delay line. It has been found that, if the delay time associated with the circuit external to the delay line can be reduced to zero, or as a minimum reduced to a length much shorter than the shortest wavelength in the tuning range, then single-mode oscillator operation over an octave bandwidth or more is possible. An important observation critical to this optimistic conclusion was that the amplitude and phase responses of the delay line are locked in synchronism as the device is tuned. Interestingly, the foregoing need to minimize the length of the circuit external to the delay line is also a condition that applies to MSW resonator-based oscillators as well as YIG-sphere oscillators.

REFERENCES

1. R.L. Carter, J.M. Owens, W.R. Brinlee, Y.W. Sam and C.V. Smith, Jr., in Proceedings of the 1981 MTT-S International Microwave Symposium.
2. J. Haworth, in Proceedings of the 1975 IEEE MTT-S International Microwave Symposium.
3. N.D.J. Miller and D. Brown, Electronics Letters 12, 209 (1976).
4. J.P. Casters, IEEE Trans. Magnetics, MAG-14, 826 (1978).
5. J.C. Sethares and M.R. Stiglitz, J. Appl. Phys. 52, 2273 (1981).
6. J.P. Casters, P. Hartemann, J.M. Dupont, Y. LeTron, A. Bert and A. Trillaud, in Proceedings of the 1983 IEEE MTT-S International Microwave Symposium.
7. V. Lander and J.P. Parekh, to be published.

THE STATUS OF MAGNETOSTATIC
WAVE OSCILLATORS

R. L. CARTER & J. M. OWENS
Electrical Engineering Dept.
University of Texas at Arlington
Arlington, Texas USA 76019

Abstract

The continual demand for increased performance in modern communication and radar systems in terms of increased bandwidths and higher operating frequencies has led to investigation of novel techniques and technologies for analog signal processing. In particular, surface acoustic waves (SAW) have been extensively exploited with great success to this end, but systems requiring center frequencies greater than 1 GHz have pushed SAW devices nearly the practical physical limit of the technology. A novel technology promising higher frequencies is based on magnetostatic waves (MSW) propagating in epitaxial ferrite films such as Yttrium Iron Garnet (YIG). These waves can be exploited in devices offering instantaneous bandwidths up to 2.2 GHz at microwave center frequencies from 0.5 to 30 GHz.

This MSW signal processing technology, based on transversal filtering concepts used in SAW, has been under extensive investigation for the past 10 years. This paper will first discuss physical properties and limits of magnetostatic waves, and consider similarities and differences with SAW. Next MSW resonators will be discussed, followed by the theory and status of MSW oscillators.

Introduction

Magnetostatic waves are slow, dispersive, magnetically dominated electromagnetic waves which propagate in magnetically biased ferrite materials at microwave frequencies (1-30 GHz in Yttrium Iron Garnet). A key to the recent interest in MSW devices has been the development of high quality low line-width (low loss), large area, Yttrium Iron Garnet (YIG, $Y_3Fe_5O_{12}$) grown on Gadolinium Gallium Garnet (GGG, $Gd_3Ga_5O_{12}$) substrates as epitaxial films. As an example, magnetostatic waves in EPI-YIG offer lower propagation loss than SAW on lithium niobate at 3 GHz and losses are less than 13 dB/msec at 10GHz.

Magnetostatic wave propagation in thin films has been extensively considered, and three major propagating modes with the propagation direction in the film plane have been used in device applications. These three modes are determined by the relative orientation of bias field, propagation direction and the crystalline orientation of the ferrite slab normal. The modes are dispersive and characterized by a limited propagation passband width and magnetic bias field tunability. This tunability principle can permit propagation of a particular wave length at any particular frequency in a range of frequencies. First consider the Magnetostatic Surface Wave (MSSW) mode in which the bias field, H , is perpendicular to the direction of the wave propagation and in the plane of the film. This mode has highly anisotropic propagation in the plane of the film. This mode has highly anisotropic propagation in the film plane and the mode energy is

confined to the "top" surface for forward propagation and to the "bottom" surface for reverse propagation as indicated in Figure 1. A second

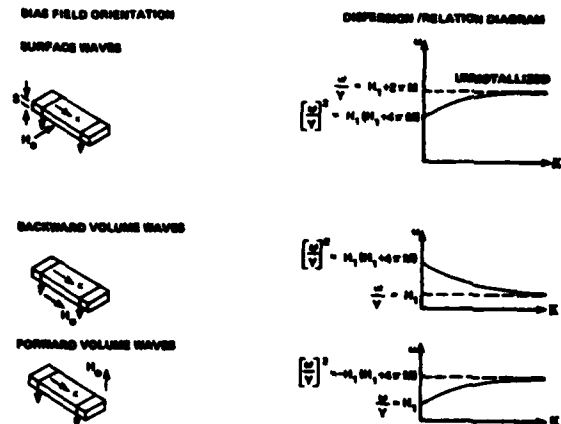


Figure 1 Principal MSW Propagation Modes

mode is the Magnetostatic Forward Volume Wave (MSFVW) in which the bias field, H , is perpendicular to the film plane and is characterized by nearly isotropic propagation in the plane of the slab with mode energy distributions resembling those found in rectangular metal pipe wave-guides. The lowest order mode is usually most easily excited. The third mode, the Magnetostatic Backward Volume Wave (MSBVW) exists when the direction of the bias field and propagation are the same in the film plane. This mode has opposite phase and group velocity directions, is highly anisotropic and is multimoded as in the MSFVW case.

All modes typically have velocities in the 3-300 km/sec range and wavelengths from 100 to 1 mm for 10 μ m thick epitaxial films; thus, transducer and periodic structure dimensions may be chosen to permit easy fabrication by standard photolithographic techniques. Since practically any wavelength can be achieved at any frequency, devices designed to work at a specific wavelength will also work at any frequency provided a suitable bias field can be obtained and the ferrite is uniformly saturated. Propagation delays of ten to several hundred nanoseconds per CM are typical. High Q tunable resonator structures can be fabricated in MSW technology in a similar manner to that used in SAW technology. Either Fabry-Perot or delay line structures can yield high quality tunable oscillators when operated in the feedback loop of an amplifier. Oscillator work that has centered on the Fabry-Perot structures due to problems associated with achieving stable broad band single mode tunable delay line oscillators. The attraction of MSW resonators lies in their high loaded Q 's (Q 's > 1500 have been observed). At present, the dominant microwave broadband tunable oscillator technology is

based on Yig spheres operated at ferromagnetic resonance. These devices have loaded Q's of less than 200 and require tedious alignment to achieve high performance. Since phase noise is inversely proportional to Q^2 the attraction of MSW devices are obvious.(2)

MSW Resonators

As in the SAW case the simplest resonator consists of a pair of wavelength selective periodic reflecting arrays separated by a delay section. These elements are arranged relative to the propagation path to form a resonator cavity and a means for transduction is established within the cavity. The single port MSSW reflecting array resonator of this type has been reported by Collins et.al.(3) This paper treats the two-port MSSW and MSFVW reflecting array resonator with an input-output transducer pair located in the resonant cavity; the geometry of this device is shown in Figure 2. This structure is of particular interest in the MSSW case due to its inherent isolation between input and output ports and ease of application in filter and oscillator systems. The MSFVW case is of interest due to its higher saturation power and easier magnet design. These resonators retain the desirable features of tunability and micro-wave frequency operation associated with previous YIG technology. Further, submicron wavelengths are not required so conventional microelectronics fabrication techniques are used for device fabrication.

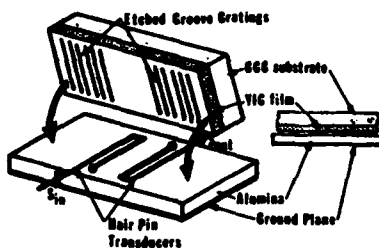


Figure 2 Schematic of a "Two Port" Magnetostatic Wave Resonator

Modeling and Theory and MSFVW'S

Since the MSSW and MSFVW'S correspond to limiting cases of electromagnetic modes, gratings in the reflecting array resonator can be modeled in terms of cascaded equivalent transmission line sections. Such modeling as applied to SAW filters and acoustic dispersive delay lines has been presented by Sittig and Coquin.(4) For MSW, however, dispersion propagation loss and nonreciprocity of propagation due to the proximity of ground planes must be included. In general, each array within the resonator can be considered to be a cascade of grating elements and the resonator cavity treated as two such arrays spaced by an appropriate delay section. The inverse transmission matrix for a line of length L has been presented by Collins et.al.(5) along with a discussion of Storch's(6) method for calculation of overall transfer matrices for individual arrays. For nonreciprocal propagation, as in the MSSW case,(for the volume wave cases propagation is reciprocal) this method is used to calculate the overall forward and reverse transmission matrices $(T)_F$ $(T)_R$ for an array. Denoting these matrices for an N element grating as

$$(T)_F = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^N \quad (1)$$

and

$$(T)_R = \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}^N \quad (2)$$

and a ... h are the transmission matrix coefficients for the forward and reverse line section. The various reflection and transmission coefficients are then evaluated as:

$$Z_F = \frac{Z_F - Z_R}{Z_F A + B - Z_R (Z_F C + D)} \quad (3)$$

$$P_F = \frac{Z_R (D - A + Z_F C - D / Z_F)}{Z_F A + B - Z_R (Z_F C + D)} \quad (4)$$

$$Z_R = \frac{Z_R - Z_F}{Z_R E + F - Z_F (Z_R G + H)} \quad (5)$$

and

$$P_R = \frac{Z_F (H - E + Z_R G - H / Z_R)}{Z_R E + F - Z_F (Z_R G + H)} \quad (6)$$

The corresponding insertion and reflection losses can be obtained from finding $-20 \log_{10}$ of the magnitudes of these factors. These quantities do not account for transducer loading effects.

The resonator cavity can be considered equivalent to a microwave Fabry-Perot resonator(7) if the concept of an effective mirror plane is introduced.(8) The distance, d , the effective mirror plane is located into an array can be determined using the reactance slope, the rate of phase change of the reflection factor, at the insertion loss notch;

$$d = \frac{1}{2vg} \frac{d\theta}{dw} \quad (7)$$

Where vg is the group speed, θ , the phase and w the radian frequency. For a given resonator design the effective resonator cavity spacing is just $L = d_F + d_R + l$ where l is the reflecting array spacing and the subscripts denote the forward and reverse wave factor; the mode spacing for the resonator is

$$\Delta f_m = \frac{1}{L} (T_F + T_R) \quad (8)$$

where T_F and T_R are respectively the group delay times per unit length. The remaining parameter of interest is the resonator quality factor or Q , which is equivalent to the reciprocal fractional bandwidth, $W_0/\Delta W$ where ΔW is the half-power bandwidth for the resonance at W_0 . This

identification permits determination of Q's directly from insertion loss versus frequency characteristics for both theory and experiment. Grouping individual resonator power losses into array, propagation and transducer losses give a loaded quality factor Q < as

$$1/Q_L = \frac{1}{Q_u} + \frac{1}{Q_{ext}} \quad (9)$$

where the external quality factor Q_{ext} is associated with transducer loading and the unloaded quality factor Q_u is the parallel combination of the array quality factor Q_R , and the material or propagation quality factor Q_m . Q_R can be obtained in terms of the effective cavity length, d, and the geometric mean of the forward and reverse reflection factors p as (9)

$$Q_R = \frac{2\pi d}{\lambda_0(1-(p/2))} \quad (10)$$

where λ is the resonant wave length. The material quality factor Q_m is related to the wave attenuations and group speeds as

$$Q_m = W_0/2\alpha v_g$$

where α and v_g are respectively the associated geometric mean values for forward and reverse propagation.

Experimental Resonator Results

A range of MSSW and MSFVW resonators have been fabricated and evaluated experimentally. Table I details the results of these studies. Both MSSW and MSFVW resonators have been studied. The best of the MSFVW devices use multiple cavities, and/or reflectors within the cavity to achieve higher spurious rejection and thus stable single mode oscillators.

MSW OSCILLATOR THEORY

Figure 2 shows a block diagram of an MSW oscillator. For this configuration the requirement for oscillation is that the loop gain be 1. The phase noise for a resonator stabilized oscillator is given by Lewis (10),

$$L(F) = 10 \log \frac{GkTf_p^2}{Q^2 P_o(f-f_0)^2} \text{ dbc}/H_z \quad (11)$$

Where G is the amplifier gain, k is Boltzmann's constant, T ambient temperature ($^{\circ}K$) F the noise figure of the amplifier f_0 the carrier frequency, Q the quality factor of the resonator, and P_o the loop power.

In general even for MSFVW resonators the resonator is the saturated element, and the amplifier is operating in the linear range. Thus the noise floor is dominated by the amplifier noise floor. Thus

$$\frac{\text{Noise Power}}{\text{Carrier Power}} = 10 \log \frac{kTFG}{P_s} \quad (12)$$

Where G = Gain of the amplifier, F = noise figure, P_s = saturation power, L_s = insertion loss of the resonator = G = gain of the amplifier. Using this result a noise floor of -150 dBc is possible with an amplifier with a 4 db N.F. A saturation power of -7dbm and an insertion loss of 13db.

d (um)	10	20.5	14.6	8	23
h/d (%)	2.5	4.4	2.7	3.5	1.5
D (mm)	3.6	3.6	3	3	--
N	40	40	60	56	--
t (um)	0	0	0	0	0
F (GHz)	3	3.2	5	2.2	3-7GHz
P (dB)	16	12	16	13	22
R (dB)	10	12	12	6	.5
Q	830	500	1500	550	500
WAVE/ RESONATOR	MSSW/ SIMPLE	MSSW/ SIMPLE	MSSW/ SIMPLE	MSFVW/ SIMPLE	MSFVW/ OBLIQ.
Lab.	Unv. Tx.	Unv. Tx.	Unv. Tx.	Thomson CSF	Thom- CSF

Table 1 - Characteristics and experimental performances obtained with MSSW and MSFVW two-port resonators. (9) d is the thickness of the YIG films, h/d the groove depth to film thickness ratio, D the distance between the gratings, N the number of grooves per array, t the spacing between the YIG and the transducers, F the resonance frequency, P the insertion loss at resonance, R the rejection of the off-resonance transmission and Q the loaded quality factor.

EXPERIMENTAL OSCILLATOR RESULTS

Experimental investigation of MSW resonator based oscillators have been carried out by two groups one at the University of Texas At Arlington on MSSW resonators, and on at The Thomson-CSF Research Center, Orsay, France, utilizing oblique reflection MSFVW resonators. The results of these studies are shown in Table II. The oscillators both have performance specs superior to YIG sphere oscillator performance with the potential for simple less expensive assembly. Figure 3 shows the phase noise performance of the University of Texas oscillator compared with an HP8620C YIG sphere based sweep oscillator. Also shown is the theoretical performance of an optimized MSW resonator oscillator with A Q= 800 and an amplifier noise figure of 3db.

TABLE 11 MSSW OSCILLATOR STUDIES

STUDY	MODE	TYPE OF RESONATOR	AMPLIFIER	Q	TUNING	PHASE NOISE @10KHZ	NOISE FLOOR
Univ. Tx.	MSSW	Simple Cavity	2 Stage Hybrid GaAs FET	270	2.0-4.5GHz	-80dbc	-135dbc
Thomson CSF	MSFW	90° Oblique Reflector 2 Cavity	4 Stage Hybrid GaAs FET	500	3-7GHz	-90dbc	-150dbc

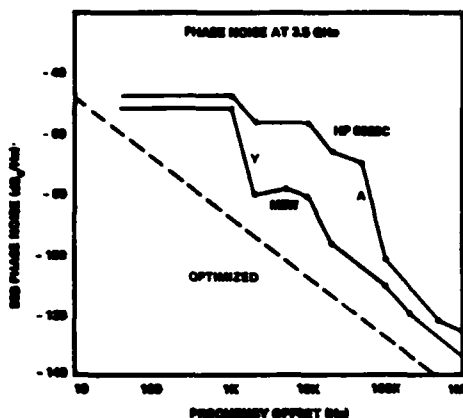


Figure 3. MSSW Oscillator Phase Noise

CONCLUSIONS

MSSW Resonator oscillators have shown superior phase noise performance to their YIG sphere counterparts. They offer the potential for simpler fabrication and potential exists for significant performance improvement.

ACKNOWLEDGEMENT

The authors would like to acknowledge the support of the Air Force Office of Scientific Research through Grant AFOSR 80-0264.

REFERENCES

1. J. M. Owens, C. V. Smith, Jr. and R. L. Carter Proc 35th Freq Control Symp (1981) U.S.A. Eradcom 077030, PP 358-363, 1981.
2. J. M. Owens, and R. L. Carter MSN V.13 No. 3, PP.103-111, 1983.
3. J. H. Collins et. al., Proc IEEE Lett., 65 1090-1092, 1977.
4. E. K. Gittig, G. A. Coquin, IEEE Trans. SU, SU-15 111-119, 1968.
5. J. H. Collins et al. 1977 Ultra Solids Symp Proc. IEEE Cat. No. 77 CH. 1264 - ISU, 1977.
6. L. Storch, AIEE Trans., 73, 616-618, 1955.
7. R.E. Collin, Foundations For Microwave Engineering, McGraw-Hill Book Co., Inc. N. Y. 1966.
8. G. L. Matthaei, et. al. IEEE Trans. SU, SU 23, 99-107, 1976
9. J. P. Castera, P. Hartmann Proceeding of 1981 RADC Microwave Magnetics Technology Workshop, RADC TR 83-15, 218-229, 1983.
10. M.F. Lewis, Proc 28th Annual Fx Control Symp, 34-37 1974.
11. J. M. Owens, R. L. Carter, Y.W. Sam to be published Proc 1983 MTTs, June 1983.
12. J. P. Castera, P. Hartman, J. M. DuPont to be published Proc 1983 MTTs, June 1983.

AN X-BAND GaAs FET OSCILLATOR USING A DIELECTRIC RESONATOR

K. Wakino, T. Nishikawa, S. Tamura and H. Tamura

Murata Manufacturing Co., LTD
26-10, 2-Chome Tenjin
Nagaokakyo-shi, Kyoto, Japan

Summary

An X-band GaAs FET oscillator highly stabilized by a dielectric resonator was developed. The oscillator's frequency stability is less than ± 1 MHz (-40°C to 85°C), and its size has been greatly miniaturized (20.3mm x 12.6mm x 8.8mm).

For the dielectric resonator, a new material $\text{Ba}(\text{NiTa})\text{O}_3$ - $\text{Ba}(\text{ZrZnTa})\text{O}_3$ has been developed which has an extremely high Q value and high temperature stability. The characteristics of the resonator are $K=29$, $Q=10,000$ at 10 GHz and $\tau_f=0$ ppm/ $^{\circ}\text{C}$.

have larger lattice constants than that of $\text{Ba}(\text{ZnTa})\text{O}_3$, that is 4.10 Å, the Q value of $\text{Ba}(\text{ZnTa})\text{O}_3$ is improved. And in the reverse case, the Q value is degraded.

When the perovskite materials are added to $\text{Ba}(\text{ZnTa})\text{O}_3$ ceramics, the ceramics

Table 1. Dielectric characteristics of $\text{Ba}(\text{ZnTa})\text{O}_3$ ceramics to which 5 mol% perovskite materials are added.

additives	(A) lattice parameter	K	(at 7GHz) Q	(ppm/ $^{\circ}\text{C}$) τ_f
none	---	29.5	10,000	1.0
BaZrO_3	4.18	30.4	14,500	7.5
BaSnO_3	4.12	28.7	10,500	3.1
SrZrO_3	4.10	30.9	13,400	18.0
SrSnO_3	4.03	29.1	5,000	11.4
BaTiO_3	3.99	31.6	6,800	1.1
SrTiO_3	3.91	32.3	300	20.9

Introduction

Many kinds of dielectric resonators have been used for microwave oscillators.^{1,2} The dielectric losses of known ceramics, however, are not sufficiently small at the X-band or at higher frequencies. In order to realize an X-band oscillator, a new material had to be developed.

In this paper, we report an X-band GaAs FET oscillator using a dielectric resonator, which has high temperature stability (± 1 MHz from -40°C to 85°C at 11.8GHz) and small size (20.3mm x 12.6mm x 8.8mm).

We also propose a $\text{Ba}(\text{NiTa})\text{O}_3$ - $\text{Ba}(\text{ZrZnTa})\text{O}_3$ material for the dielectric resonator, which has characteristics of $K=29$, $Q=10,000$ at 10GHz and temperature coefficient of resonant frequency $\tau_f=0$ ppm/ $^{\circ}\text{C}$.

Dielectric Resonator

At the beginning of the 1960's, F. Galasso and J. Pyle investigated the crystal structures of $\text{Ba}(\text{B}^{0.33}\text{Ta}^{0.67})\text{O}_3$ type materials, and they reported that these materials had perovskite pseudocells and formed the hexagonal superstructures due to the ordering of $\text{B}^{0.33}$ -Ta ions.³ In 1977, it was reported that $\text{Ba}(\text{Zn}^{1/3}\text{Nb}^{2/3})\text{O}_3$ - $\text{Ba}(\text{Zn}^{1/3}\text{Ta}^{2/3})\text{O}_3$ ceramics had high Q values at microwave frequencies.⁴

We studied the $\text{Ba}(\text{ZnTa})\text{O}_3$ with additives of well known perovskite materials and found an interesting result. Table 1 shows the dielectric characteristics of $\text{Ba}(\text{Zn}^{1/3}\text{Ta}^{2/3})\text{O}_3$ ceramics to which 5 mol% perovskite materials are added. When the added perovskite materials



$\text{Ba}(\text{ZnTa})\text{O}_3$



$\text{Ba}(\text{ZnTa})\text{O}_3 + 4\% \text{ BaSrO}_3$

Fig.1 Scanning electron micrographs of polished and etched surface of dielectric resonator.

do not come to form the superstructure. Then the grain growth and crystallization of $\text{Ba}(\text{ZnTa})\text{O}_3$ ceramics are accelerated. This acceleration of the crystallization improves the Q value. In order to get higher Q value, however, the added perovskite materials must have larger lattice constants than the $\text{Ba}(\text{ZnTa})\text{O}_3$.

Fig. 1 shows the grain growth of $\text{Ba}(\text{ZnTa})\text{O}_3$ ceramic to which 5 mol% BaZrO_3 is added. Both samples were sintered at

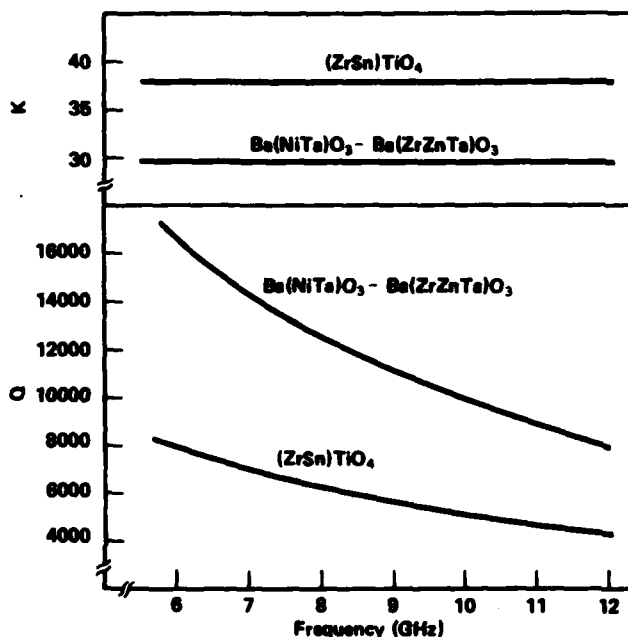


Fig. 2 Frequency vs K and Q value of $\text{Ba}(\text{NiTa})\text{O}_3$ - $\text{Ba}(\text{ZrZnTa})\text{O}_3$ ceramic.

1,500 °C for 4 hours.

As $\text{Ba}(\text{ZrZnTa})\text{O}_3$ material have a positive temperature coefficient, we added $\text{Ba}(\text{NiTa})\text{O}_3$ material which have a negative temperature coefficient. Thus, the characteristics of $K=29$, $Q=10,000$ at 10 GHz and $\tau_f=0$ ppm/°C are obtained at the composition of $0.2\text{Ba}(\text{Ni}/3\text{Ta}2/3)\text{O}_3$ - $0.8\text{Ba}(\text{Zr}0.04\text{Zn}0.32\text{Ta}0.64)\text{O}_3$.

Fig. 2 shows the frequency dependence of K and Q value, and Fig. 3 shows frequency temperature dependence of resonant frequency and Q value.

K and Q values are measured by Hakki and Coleman's dielectric resonator method.⁵

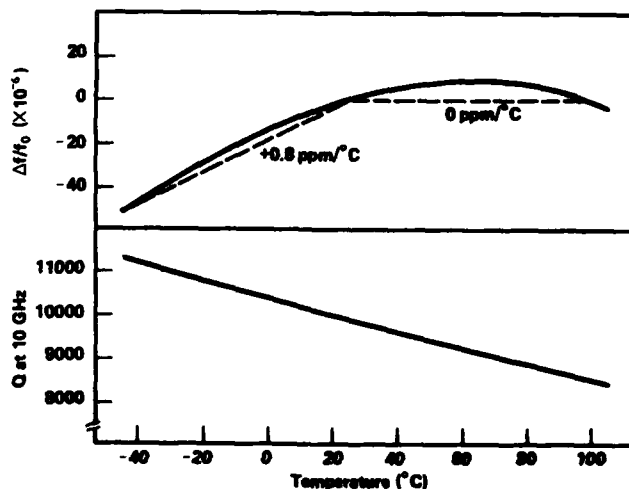


Fig. 3 Temperature vs resonant frequency and Q value of $\text{Ba}(\text{NiTa})\text{O}_3$ - $\text{Ba}(\text{ZrZnTa})\text{O}_3$ ceramic.

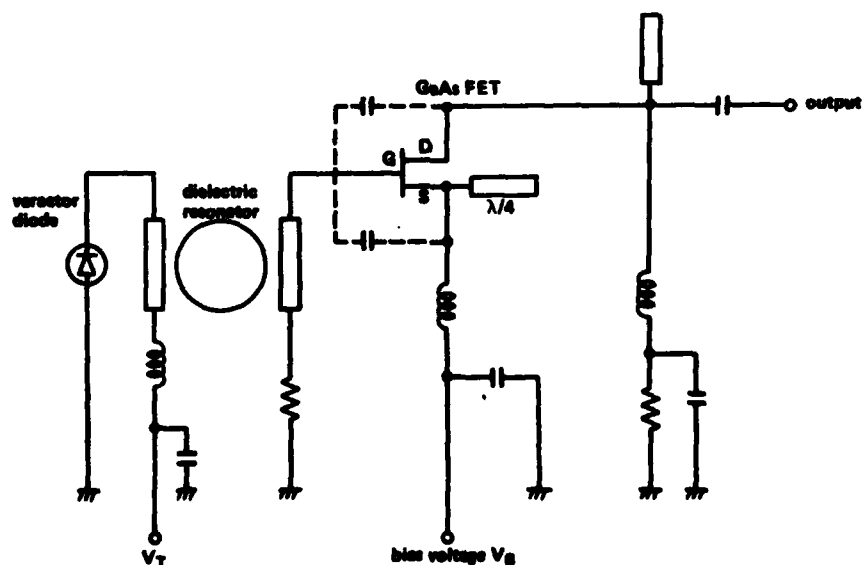


Fig. 4 equivalent circuit of GaAs FET oscillator

Oscillator Design

Fig. 4 shows the equivalent circuit of the oscillator.

An oscillator which consists of a GaAs FET, a dielectric resonator and a varactor diode is mounted on MIC. The GaAs FET, FSX51X of Fujitsu Limited was used in this work. The dielectric resonator is mounted on an Alumina substrate with a spacer made of forsterite.

The resonator is magnetically coupled to gate line, and its resonant frequency is pretuned mechanically by the dimensions $5.5\text{mm} \times 2.4\text{mm}$ for the frequency of 11,800 MHz. Slight deviation from the specified oscillation frequency, after the resonator is housed in the case, is controlled by a bias voltage (V_T) which controls the capacitance of the varactor diode. Fig. 5 shows the tuning characteristic of V_T versus oscillation frequency. The oscillation frequency changes linearly about 2MHz, when V_T changes from 0 to 5 volts.

The temperature coefficient of a dielectric material can be controlled around zero ppm/ $^{\circ}\text{C}$ with high precision by changing the ratio of $\text{Ba}(\text{NiTa})\text{O}_3$ and $\text{Ba}(\text{ZrZnTa})\text{O}_3$. The temperature coefficient is chosen so that it compensates the thermal linear expansion of the case and the impedance deviation of the FET amplifier. In this oscillator, 3 ppm/ $^{\circ}\text{C}$ was chosen.

The oscillator is hermetically sealed in the SPCC case to keep stable operation even in severe environmental conditions. The RF output and DC bias terminals are made of pin type leads.

Oscillator Performance

The electrical characteristics of the oscillator are summarized in Table 2. The measured operation is $I_D=38\text{mA}$, $f_o=11,800\text{MHz}$ and $P_o=12\text{dBm}$ at bias voltage of 5 Volts at room temperature.

Fig. 6 shows the temperature stability of the oscillation frequency. For comparison, two kinds of dielectric resonators are used. A $\text{Ba}(\text{NiTa})\text{O}_3\text{-Ba}(\text{ZrZnTa})\text{O}_3$ resonator has the Q value of 8,500 and a $(\text{ZrZn})\text{TiO}_4$ resonator has the Q value of 4,400, at 11.8GHz. The position of these resonators were adjusted so that the same output power can be obtained at the same DC power supply. Both of these resonators have almost the same linearities of resonant frequency. But the linearity of the oscillation frequency using the resonator of higher Q value was better than the other. The frequency stability of less than $\pm 1\text{MHz}$ from -40°C to 85°C was obtained when the $\text{Ba}(\text{NiTa})\text{O}_3\text{-Ba}(\text{ZrZnTa})\text{O}_3$ resonator is used.

Fig. 7 shows the sensitivity of the frequency (f_o) and output power (P_o) to variations in bias voltage. The pushing figure is 400KHz/V, and the efficiency is 8.3% at 5 volts of DC power supply.

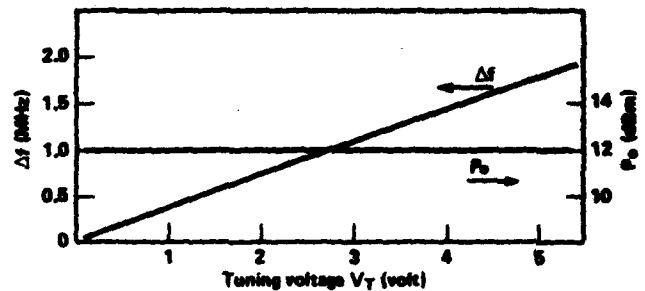


Fig.5 Tuning bias voltage vs oscillation frequency.

Table 2 Oscillation performance

Frequency	11,800 MHz
Output power	12.0 dBm
Efficiency	8.3%
Frequency stability	$\pm 1\text{ MHz } (-40^{\circ}\text{C} \sim 85^{\circ}\text{C})$
Pushing figure	400 KHz/V
Pulling figure	900 KHz (VSWR = 1.5)
Tuning range	1.8 MHz
DC power supply voltage	5 Volts
Current	38 mA
Dimension	$20.3 \times 12.6 \times 8.8\text{ mm}$
Weight	4.8 g

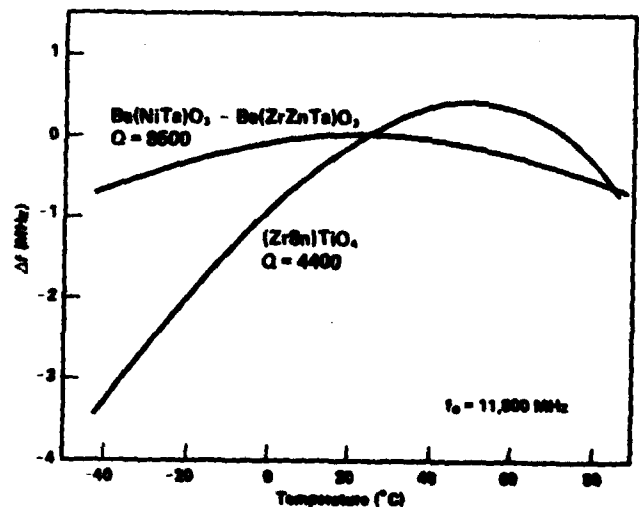


Fig.6 Temperature vs oscillation frequency using two kinds of dielectric resonators.

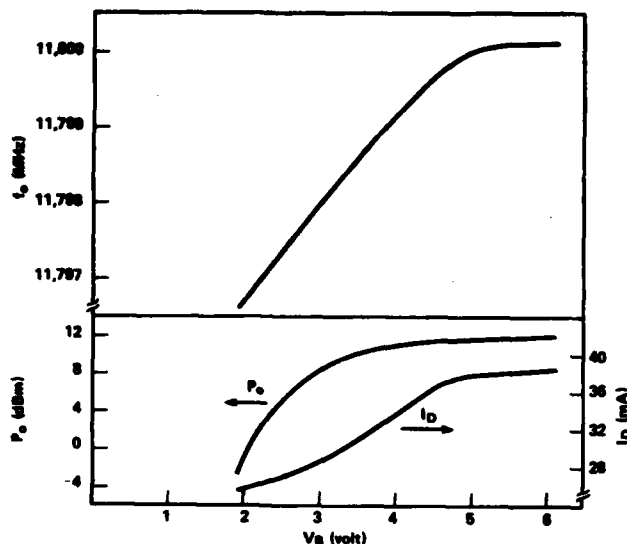


Fig.7 DC bias voltage vs resonant frequency and out put power.

Conclusion

A dielectric resonator of Ba(NiTa)O₃-Ba(Zr₂NbTa)O₃ material has been developed, which has very high Q value (Q=10,000 at 10GHz) with K=29 and $\tau_f=0$ ppm/°C.

Using the resonator, an X-band miniaturized GaAs FET oscillator was developed which provided frequency stability of +1 ppm/°C (-40°C~+85°C), output power of 12dBm and power efficiency of 8.3%.

As the oscillator has a simple construction, it is suited to mass production and can be produced at low cost. This X-band oscillator is suitable for use within an SHF-UHF converter for DBS.

Acknowledgements

The authors wish to thank Mr. M. Itoh of Fujitsu for his helpful suggestions.

The authors also would like to thank Mr. M. Sakurai of Toshiba for his helpful discussions.

References

- 1) J. K. Plourde, D. F. Linn, I. Tatsuguchi, C. B. Swan, "A Dielectric Resonator Oscillator with 5 PPM Long Term Frequency Stability at 4 GHz", IEEE-MTT-8, IEEE Cat. No.77CH1219-5 MTT, p.273 (1977).
- 2) T. Saito, Y. Arai, H. Komizo, Y. Itoh, T. Nishikawa, "A 6GHz Highly Stabilized GaAs FET Oscillator Using a Dielectric Resonator", IEEE-MTT-8, IEEE Cat. No.79CH1439-9 MTT-8, P.197 (1979).
- 3) F. Galasso and J. Pyle, "Ordering in Compounds of the A(B^{0.33}Ta^{0.67})O₃ Type", Inorganic Chemistry, 2, p482 (1963).
- 4) S. Kawashima, M. Nishida, I. Ueda, H. Ouchi and S. Hayakawa, "Dielectric Properties of Ba(Zn_{1/3}Nb_{2/3})O₃-Ba(Zn_{1/3}Ta_{2/3})O₃ Ceramics at Microwave Frequency", proceedings of the 1st Meeting on Ferroelectric Materials and their Applications, 1,p293 (1977).
- 5) B. W. Hakki and P.D. Coleman, "A Dielectric Resonator Method of Measuring Inductive Capacitance in the Millimeter Range", IRE Trans. on MTT MTT-8, p.402 (1960).

AN ANALYSIS OF A MICROPROCESSOR CONTROLLED
DISCIPLINED FREQUENCY STANDARD

Bill Bourke & Bruce Penrod
Austron, Inc. P.O. Box 14766
Austin, Texas 78761

Summary

This paper summarizes the ideas, experiments and conclusions from an effort to develop a new, producible disciplined quartz frequency standard. Some of the original thinking comes from a paper published in the late sixties which described the use of a third order phase locked loop (PLL) system to fully correct a drifting frequency source. Further, the paper described the behavior of the corrected frequency source when the PLL was without reference input, or coasting, and showed that much of the drift had been effectively removed by the third order control. These results lead to the development at Austron of the model 2010B Disciplined Frequency Standard. The objective here is to implement this same idea at the current state-of-the-art, taking advantage of the power and versatility of the microprocessor.

Introduction

Briefly, our problem is to phase lock a high quality quartz oscillator to a reference having superior long term stability, such as cesium or Loran-C. The phase lock technique must be able to correct the oscillator's phase, frequency and drift in frequency (ageing) in such a manner that the phase error approaches zero when the PLL achieves steady state. This requirement allows performance of the second, and more important, function of a disciplined frequency standard, i.e., backup of a primary standard in the event of its failure. If the above described PLL has operated long enough to achieve zero steady state phase error then it has "learned" the ageing rate of the oscillator under control and, in the event of reference failure, the PLL will continue to correct the oscillator in such a manner that only a fraction of its inherent ageing will be observed over the duration of the reference outage. In this way, a relatively inexpensive oscillator may offer atomic standard performance for intervals of 24 hours or less.

Definitions and Assumptions

In order to more specifically define this problem and to prevent any confusion in terms, the following basic formulae are presented.

First the fractional frequency offset:

$$y(t) = \frac{1}{2\pi\nu_0} \frac{d\phi}{dt} \quad (1)$$

where ν_0 is the nominal center frequency of oscillation and ϕ is the instantaneous phase of the oscillator versus some ideal source of ν_0 .

Now the phase time

$$x(t) = \int_0^t y(t') dt' \quad (2)$$

which yields a quantity whose units are time and allows direct comparison of clocks.

A model of the quartz oscillator to be disciplined, in these terms might be

$$y(t, T) = y(t, T_0) + y_r(t) + Dt + (T - T_0) \frac{\partial y(t, T_0)}{\partial T} \quad (3)$$

where T is the temperature of the oscillator, T_0 is the average temperature of the oscillator, $y_r(t)$ is the random contribution of frequency, and D is the coefficient of drift. The addition of the temperature dependence to the familiar oscillator model comes from observations by the authors that indicate that its magnitude is far from negligible.

The model for the phase time is

$$x(t, T) = y(t, T_0)t + x_r(t) + Dt^2/2 + (T - T_0) \frac{\partial y(t, T_0)}{\partial T} t + x(0) \quad (4)$$

Now the problem may be quite clearly defined. The initial phase offset $x(0)$, average over temperature frequency offset $y(t, T_0)$, and the drift rate D must be determined and corrected in the presence of the random contributions and temperature effects.

Proposed Solution

Fig. 1 illustrates a classic type III PLL in the Laplace transform notation. The type III servo employs three pure integrations in its feedback. Servo theory shows that in order to track with zero steady state error a quadratic phase input, as included in eq. 4 and arising from the drift in frequency, a type III PLL is necessary and sufficient. Fig. 1 also shows that a technique of forcing the loop input phase error, on detection of reference failure, to zero is employed and it also indicates which points in the loop contain the values pertaining to the ageing and frequency corrections, at steady state.

It should be noted that three parameters, K , τ_1 , and τ_2 may be adjusted to set such loop characteristics as ageing determination averaging time, and phase determination averaging time. These parameters will determine the effectiveness of the PLL in determining the crucial oscillator parameters in the presence of random and temperature effects.

Implementation

A digital technique for implementing Fig. 1 is shown in Fig. 2. The output of the phase detector is digitized and the control integrations are performed in microprocessor software using a 48 bit floating point arithmetic. After scaling, the

CLASSIC PLL APPROACH TO DISCIPLINING

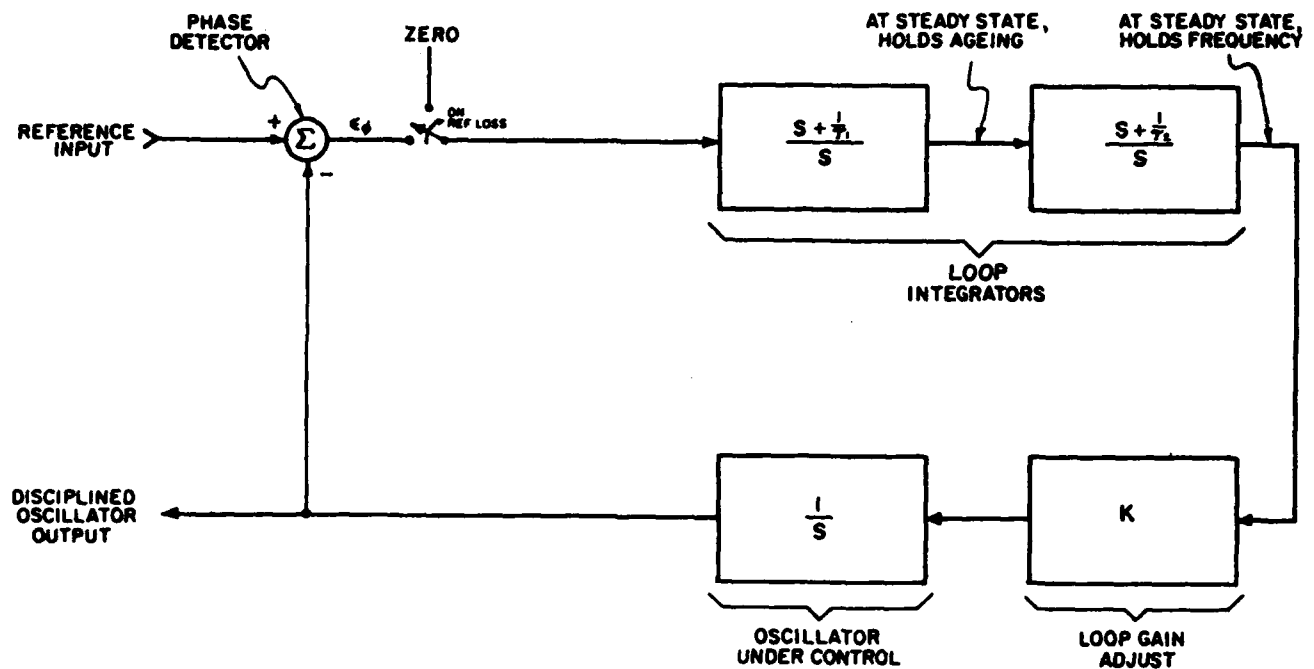


FIG. 1

MICROPROCESSOR IMPLEMENTATION

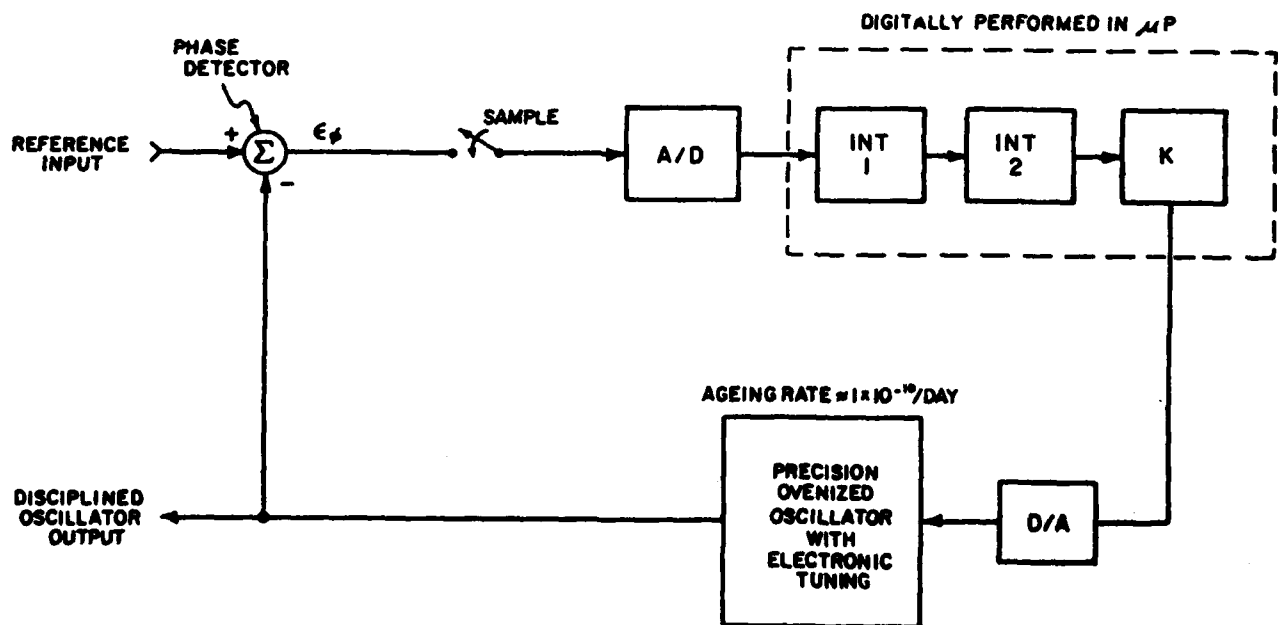


FIG. 2

control word is sent to a high resolution (16 bits), low temperature coefficient digital to analog converter whose output voltage drives the varactor tuning on a model 1150, 5th overtone AT-cut precision ovenized oscillator. This particular oscillator exhibited approximately a 1×10^{-10} /day ageing rate.

Experiment

In order to verify that the proposed solution could work, an experiment was performed and is schematically represented by Fig. 3. The model 2100 Loran-C receiver obtains its reference from the disciplined frequency standard, applies phase corrections necessary to synthesize this reference input with Loran-C and outputs this phase corrected signal to the reference input of the disciplined frequency standard. The phase detector output voltage, prior to digitizing, is output to an analog chart recorder to allow observation both during the initial acquisition time and after the loop has been intentionally opened to simulate reference failure.

Since the PLL parameters K , τ_1 , τ_2 are readily variable, the software in this disciplined frequency standard included an algorithm to effectively increase the various averaging times during PLL acquisition. In this way the initial turn on transient is minimized and a useable output from the disciplined frequency standard is available in several hours rather than days. The ultimate averaging times reached after approximately 48 hours of PLL acquisition were:

Ageing Determination - 8 days
Frequency Determination - 5000 seconds
Phase Determination - 4000 seconds

In this experiment, the PLL was allowed to acquire for 7 days. At that time, the loop was opened in software and the input to the loop integrators was set to zero. Twelve days of operation were recorded in this mode and the data which follow are from that period.

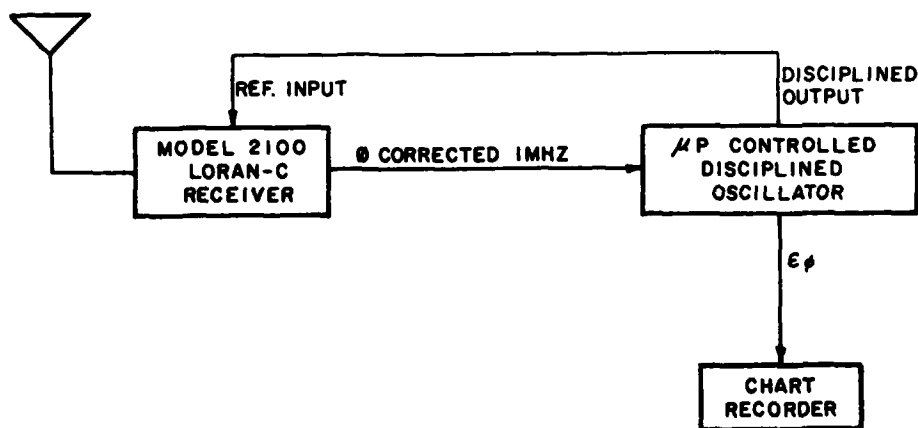
Results

Fig. 4 is a compressed version of the phase record from the chart recorder of Fig. 3 obtained by sampling at 1800 second intervals. The 575 points were then plotted by an HP-85 computer. The succeeding plots arise from various manipulations of these original data.

The phase time accumulation in Fig. 4 shows that the ideal correction of the oscillator did not occur. In fact $21 \mu\text{s}$ of phase error resulted over the 12 day opened loop interval. However, a free running oscillator with an ageing rate of 1×10^{-10} /day would have accumulated $622 \mu\text{s}$ just due to ageing.

Something which should be noticed in Fig. 4 is the presence of a diurnal shift in phase. These shifts have been found to be due to the variations in building temperature from night to day causing a corresponding frequency shift in the oscillator. Unfortunately a temperature record was not kept

EXPERIMENT VS LORAN-C



- VARIABLE TIME CONSTANTS EMPLOYED
- FINAL EFFECTIVE MEASUREMENT TIMES
FREQUENCY - ≈ 5000 SECONDS
AGEING - ≈ 8 DAYS
- 7 DAYS ACQUISITION
- 12 DAYS OPENED LOOP

FIG. 3

OPEN LOOP RESULTS - RAW DATA

FIG. 4
PHASE TIME

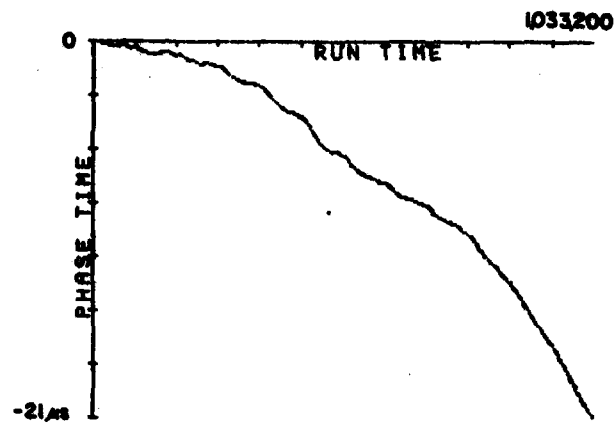
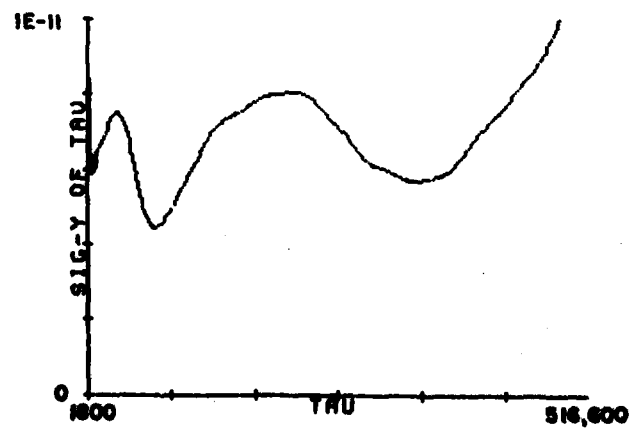


FIG. 5
FIRST DIFFERENCES (FREQ)

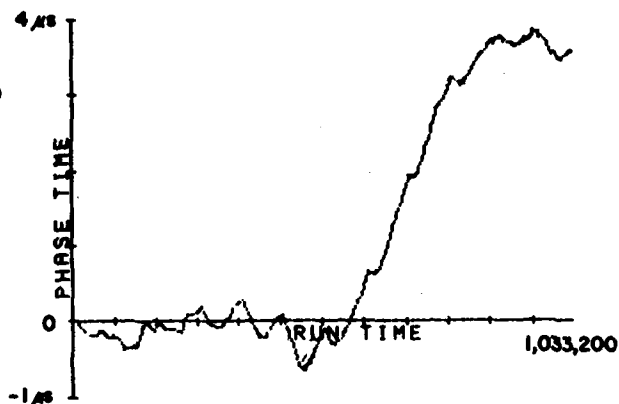


FIG. 6
ALLAN VARIANCE
FROM PHASE TIME



OPEN LOOP RESULTS - PROCESSED DATA

FIG. 7
PHASE DATA - DRIFT REMOVED

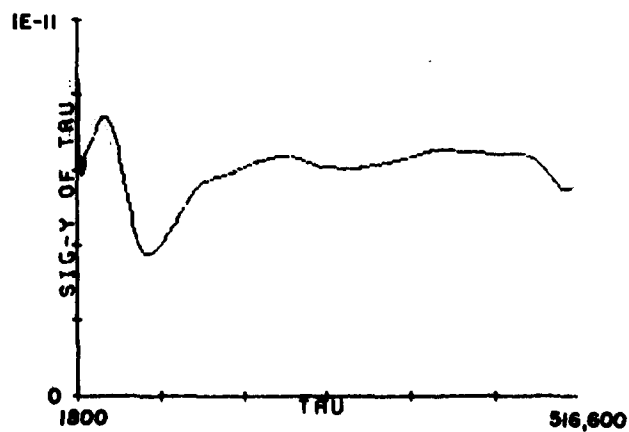


- COMPUTED AGEING FROM AVERAGE OF SECOND DIFFERENCES = -3.96×10^{-12} /DAY
- RESIDUAL FREQUENCY OFFSET = 3.49×10^{-12}

FIG. 8
1st DIFFERENCES - DRIFT REMOVED



FIG. 9
ALLAN VARIANCE FROM
PHASE DATA - DRIFT REMOVED



over the duration of the test and specific temperature versus frequency data is not available. It is worth mentioning that at the time of the experiment the authors noted the greatest changes occurred on days experiencing large night-to-day temperature differentials as obtained from normal weather data.

Fig. 5 is a plot of the first differences of the phase time data yielding fractional frequency offset data. Here can be seen an apparent residual drift in frequency over the duration of the test as well as an approximately 3×10^{-11} peak-to-peak diurnal shift which occurs at the fifth day. Smaller shifts are also present for other days.

Fig. 6 is Allan Variance data for all τ 's from 1800 seconds to approximately six days. An obvious periodic disturbance occurs for $\tau = 1/2$ day and $\tau = 1$ day which results from the presence of the diurnal temperature effect. For $\tau \geq 4$ days an apparent increase in $\sigma_y(\tau)$ proportional to τ occurs indicating the presence of residual drift in frequency.

From these plots it can be seen that the PLL has failed to perform perfectly and that there are temperature effects. It would now be interesting to attempt to determine what the errors in estimating $Y(t, T_0)$ and D of eq. 3 were and, by comparison with the Allan Variance information, see if these errors are noise induced or due to some conceptual or practical error.

Due to the apparent flicker FM background of Fig. 6, fitting a quadratic function to the raw phase data of Fig. 4 will yield a less than optimal

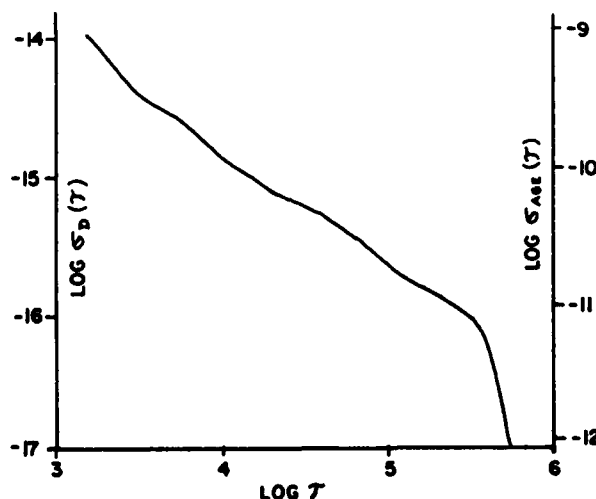
result since the residuals of phase will be decidedly nonwhite. A better technique would fit a linear function to the frequency data of Fig. 5. The residuals are still not white but they are closer to being so. However the technique employed here involved fitting a constant to the second differences of phase, the drift data. These residuals are not white either and have opposite and equal spectral density slope compared to the second technique. The resulting drift rate was $-3.96 \times 10^{-12}/\text{day}$.

Often it is useful to remove the effect of frequency drift from experimental data in order to obtain a better Allan Variance estimate and in order to determine the frequency offset present in the data. Fig. 7 shows the data of Fig. 4 after the phase accumulation due to $-3.96 \times 10^{-12}/\text{day}$ ageing has been removed. It is obvious from this plot that frequency drift was responsible for the largest part of the phase accumulation since a total accumulation of less than $4 \mu\text{s}$ remains to be attributed to an offset in frequency present at the time that the PLL was opened. The apparent magnitude of this offset is 3.49×10^{-12} and it is positive.

The first differences of the driftless data of Fig. 7 are shown in Fig. 8 as a visual check on the effects of the previous operation. It appears that the drift has largely been removed. The Allan Variance of this driftless data is now shown in Fig. 9. It should be noted here that aside from the effect of the diurnal frequency shift, the residual noise background is constant over averaging time in the manner associated with flicker FM. The level of this noise is very nearly 6×10^{-12} .

ALLAN VARIANCE APPLIED TO FREQUENCY DATA

THE ALLAN VARIANCE OF FREQUENCY DATA WILL GIVE INFORMATION CONCERNING THE VARIANCE OF THE AGEING RATE VERSUS AVERAGING TIME.



AN AGEING AVERAGING TIME OF 10^6 SHOULD YIELD AN ESTIMATE WITH $\sigma_{\text{age}} = 3 \times 10^{-12}/\text{DAY}$.

FIG. 10

Impact of Results on PLL Parameters

From the preceding results and analyses there should be some indication concerning the proper selection of the PLL parameters. First, it appears that the presence of a diurnal frequency shift in the oscillator will force the use of a quite long frequency determination averaging time in order to accurately determine $y(t, T_0)$. However, in so doing, the oscillator will not be held in nearly as tight a phase lock; a drawback which could be undesirable in some applications (the authors believe it would be better to reduce the thermal sensitivity of oscillators used in this instrument). The second observation which may be made concerning PLL parameter selection is that, due to the predominance of flicker FM background noise, an inordinately long frequency determination averaging time is of no benefit since the ambiguity of the measurement remains constant versus averaging time. On the positive side, the third observation, also due to flicker FM, is that a long ageing determination averaging time is highly advantageous. This may be shown quite convincingly by applying the Allan Variance to the frequency data of Fig. 5 and in so doing obtaining information concerning the Allan Variance of the drift versus averaging time.

Figure 10 illustrates this relationship and shows a $1/\tau$ relationship between the Allan Variance of the drift coefficient $\sigma_D(\tau)$, versus τ . Specifically, for an averaging time of 10^6 seconds the drift should be measurable with an ambiguity of approximately 3×10^{-12} /day. The drift measurement error realized in the experiment, with an averaging time of 6×10^5 seconds, and being -3.96×10^{-12} lends some credence to this idea.

Conclusions

The authors realize that a single 12 day run of data is not sufficient to make concrete statements concerning the performance of this instrument, however the following general remarks are felt to be in order:

1. The performance realized appears to be noise limited.
2. The temperature sensitivity of the oscillator under control is of prime importance.
3. To compensate ageing, as long an averaging time as practical is to be attained.

Acknowledgements

The authors wish to acknowledge the several encouraging as well as enlightening discussions with Dr. James Barnes during the preparation of this document.

References

1. D.W. Allan, L. Frey, H.E. Machlan, J.A. Barnes, "An Ultra-Precise Time Synchronization System Designed by Computer Simulation," Frequency, January 1968.

AD P 002520

BVA Quartz Crystal Resonator and Oscillator Production A statistical review

E.P. Graf and U.R. Peier
OSCILLOQUARTZ SA SWITZERLAND

In earlier publications^{1, 2, 3} the design of the BVA quartz resonator has been presented. All these papers dealt with preproduction models of the BVA. Since that time it has been necessary to find a design that at the same time allows easy manufacturing and good production repeatability.

Actual design of the BVA Resonator

The BVA consists of an electrodeless resonator which is decoupled from its mounting structure by four bridges (Fig. 1). These bridges have a width of only .4 mm to get the mounting stresses away from the active centerpart as much as possible. The electrodes are plated on two more pieces, called condensers, which are made also of AT Quartz blanks with the same cutangle as the resonator blank. Since the BVA is designed as an oven quartz this is the only possibility to avoid stresses due to thermal expansion.

The three parts, we call it 'sandwich', are rigidly hold together with clips made from stainless steel (Fig. 2). To protect the fragile bridges against shocks the whole sandwich is mounted with springs into a rigid cage consisting of the base plate and a cover plate which are fixed to 4 columns (Fig. 3).

This fairly complicated mounting structure was not used initially. Earlier mounting structures resulted in long term stability measurements such as shown in Fig. 4. The solid line is the frequency deviation $\Delta f/f$ vs. time and the dashed line is the barometric pressure. There is no need to calculate a correlation factor. After that unhappy experience there was no doubt that better decoupling of the sandwich from the enclosure was needed. With todays suspension we cannot

measure this effect.

Another typical characteristic of the BVA is its special cold weld enclosure. Because we wanted to be able to bake out the finished resonator at 250°C while pumping it to 10^{-8} mbar with a cryo pumping system, we designed the cap with a pinch-off tube. In this way we can avoid expensive manipulations in the vacuum.

The entire mounting is performed in a glove box filled with dry nitrogen (< 40 ppm H₂O). Experience has shown that humidity and presence of oxygen are the bigger enemies than dust, since the presence of the latter can be detected almost instantly whereas oxidation problems are not revealed until the ageing measurement.

Production Statistic of BVA resonator

We have produced about 200 resonators so far, thus a quantity that provides us with production statistics. Since we were trying different raw material some of the histograms show a secondary peak.

Load resonance frequency deviation at turnover is shown in Fig. 5. Because with an electrodeless resonator it is not possible to measure the frequency at the same time you adjust it, it was clear that this was one of the main problems of the BVA manufacturing process. In this production step we had to go the whole learning curve, but as can be seen from the attained mean value of -.066 ppm ($\sigma = .26$ ppm) this problem has been overcome.

Turnover temperature. A histogram is shown in Fig. 6. We got a mean of 85.5 C and a $\sigma = 6.4$ C which seems large for an angle tolerance of ± 15 sec of arc but which can be

explained by the different raw material used.

Series resistance R_1 , histogram Fig. 7, shows a secondary peak near 66Ω , the mean value is 72.6Ω ($\sigma = 4.7\Omega$). Of course these resistances are related to:

quality factor, Fig. 8. Here a secondary peak appears due to the use of russian quartz (near $2.9 \cdot 10^6$).

Motional Inductance, Fig. 9, shows a very large dispersion. Beside the secondary peak, the L_1 values scatter because in the case of an electrodeless quartz the gaps (nominal 10μ resp. 6μ) intervene.

Motional Capacitance, Fig. 10, of course shows the same distribution as L_1 .

Initial ageing. Typical ageing curves can be found in Fig. 11 and 12. $2 \cdot 10^{-11}$ /day is guaranteed.

2 g Sensitivity Vector

For the 2 g sensitivity measurements we use a turntable with 50 positions that is shown in Fig. 13. We measure every 7:2 the frequency by a beat method against a synthesizer. Each frequency measurement is repeated 20 times and the average as well as the sigma is calculated and dumped into the memory of a HP 9835 desktop calculator. The sigmas we get in this way are in the low 10^{-11} which is sufficient to measure frequency variations of 10^{-10} due to the acceleration vector.

In the electrodeless resonator one of the main stress sources, the electrode stress, is avoided, so we felt that the direction of 2 g sensitivity vector \vec{r} could tell us something about its origin. Since in the case of the BVA the position of the resonator blank within the enclosure is known (and always the same, which isn't the case for third overtone units in a HC-36/U can) it was easily possible to trace the \vec{r} vector in blank coordinates.

In the following we will use an axonometric

plot as in Fig. 14. We have a sketch of the blank (with the bridges) and the axis X, n_{AT} and Z' with n_{AT} being the normal to the AT blank.

The measurements give directly the $\Delta f/f$ vs. ψ plots for the rotations 1, 2 and 3 that are marked on the left side of Fig. 15. A least squares fit is then passed through the data to obtain the peak deviation and the offset angles. The solid line are the least squares fit values. After some vectorial calculation we get the direction of the \vec{r} vector in blank coordinates (shown on the right side).

Fig. 16 shows the same thing for another quartz. You see $r \approx 7 \cdot 10^{-10}/g$ what is about 3 times smaller than for a classical 5th OT AT resonator. Since \vec{r} wasn't parallel to n_{AT} as we first expected, we concluded that there must be a stress configuration that is caused by the mounting structure in general and especially in the mounting clips. Improving the precision in manufacturing the clips resulted in \vec{r} measurements of Fig. 17 with $r = 3 \cdot 10^{-10}/g$.

For comparison, Fig. 18 shows the same plot for a 5 MHz 3rd OT SC BVA resonator. Here we get $r = 1.7 \cdot 10^{-10}/g$ which is about 10 times smaller than the classic AT resonator.

We feel that the reason for these small values must lie in the bridge locations since the shape of the active part of the resonator is exactly the same as for a classical 5th OT resonator. Fig. 19 is a plot provided by A. Ballato⁴. The location of the bridges are exactly the points A through D. These are the angles where the force-frequency coefficient $K_f(\psi)$ becomes zero.

Fig. 20 shows a summary of about 50 measurements in a general position. By rotation of the axonometric axis we can get easily the in-plane components or in another case the out-of-plane components of the \vec{r} vector. We hope that these plots will reveal the origin

of the stress configuration, but results are not yet available.

Performance of oscillators with BVA resonators

To take full advantage of the BVA resonator, it has been necessary to develop a new oscillator. That development was reported by E.P. Graf³ in more detail. This paper will be restricted to measurements only.

The measurements are made on Oscilloquartz Model 8600 production oscillator units. Frequency measurements are made against our house reference consisting of 3 cesium beam standards. Sometimes it was necessary to use of a separate cesium beam standard to avoid noise pick up in the long distribution lines; in this case the used standard is traceable to the reference.

Frequency vs. ambient temperature. Fig. 21 gives the distribution for the relative frequency deviation vs. ambient temperature (-30 to +70°C). In the operating temperature range we get a mean deviation $y = 3.7 \cdot 10^{-10}$

Frequency vs. change of power supply voltage. For a $\pm 10\%$ change in power supply voltage we get a mean relative frequency change of $y = 3.8 \cdot 10^{-11}$. (Fig. 22)

Short term stability. The two sample Allan Variance for $\tau = 10$ sec gives a mean of $3.5 \cdot 10^{-13}$ (Fig. 23). This measurement is made against a second Model 8600 oscillator. Measurements were made against an Oscilloquartz EFOS-1 Hydrogen maser; these results are published elsewhere⁵.

Phase noise data. On the right side of Fig. 24 we find the characteristic slopes f^{-3} , f^{-1} and f^0 . On the left side we give histograms for the intersection of these straight lines with the 1 Hz Ordinate. By this way we get a mean value of -122.2 db for f^{-3} -135.3 db for f^{-1} and a white phase noise of -152.7 db.

Conclusions

1. Although of complicated design, the BVA resonator can be manufactured with repetitive performances. Frequency adjustment can be done with excellent precision even though in situ plating cannot be applied.

2. Systematic 2 g sensitivity measurements turn out to be a very efficient tool for stress studies, especially in the case of an electrodeless resonator where mounting stresses are not masked by electrode stress. It's not easy to measure the f vector, but once the data is available for statistical purposes they can be used as 'three dimensional needles' pointing straight forward to points where something is wrong.

3. The Model 8600 oscillator exhibits excellent performances in regard of stability (long term and short term) with a very small phase noise near the carrier.

Acknowledgement

The authors take this opportunity to thank Mr. Arthur Ballato (ERADCOM) for providing Fig. 19, Mr. Lüdi and Mr. J.-D. Droz (Oscilloquartz SA) for the oscillator measurements and Miss Rölli for her expert typing.

References

1. R. Besson, "A new piezoelectric resonator design", Proc. 30th AFCS (1976), p. 78 - 83
2. R. Besson and U. Peier, "Further advances on BVA quartz resonators", Proc. 34th AFCS (1980), P. 175 - 182
3. E.P. Graf, D.A. Emmons et al., "Performance of new oscillators designed for 'electrodeless' crystals", Proc. 34th AFCS (1980), p. 457 - 462
4. T.J. Lukszek, A. Ballato, "Resonators for severe environment", Proc. 33rd AFCS (1979), p. 311 - 321

5. 6. Busca, private communication, to be published

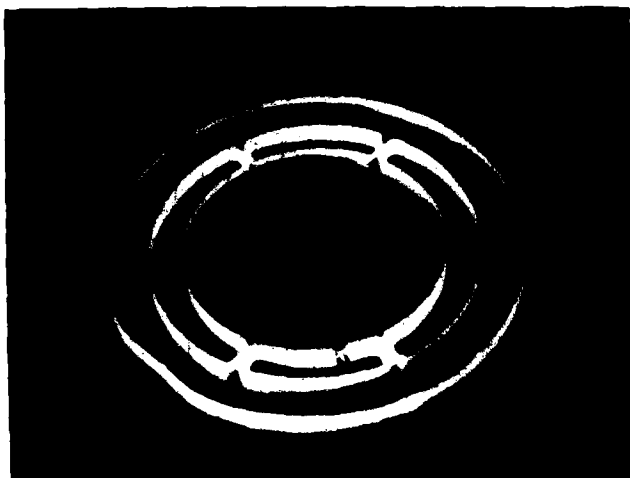


Fig. 1. BVA resonator



Fig. 2. BVA 'sandwich' held together with clips

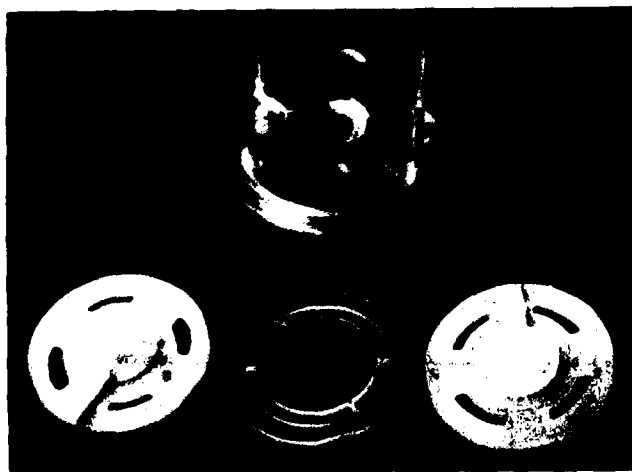


Fig. 3. Components of the cage (clips not shown)

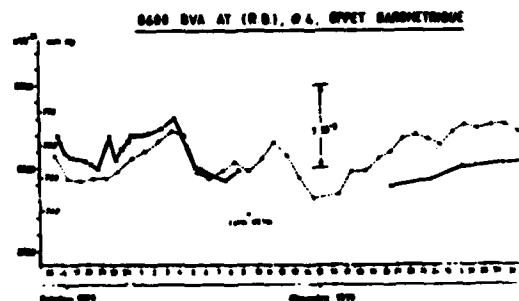


Fig. 4. Frequency dependence of barometric pressure shown in early prototypes

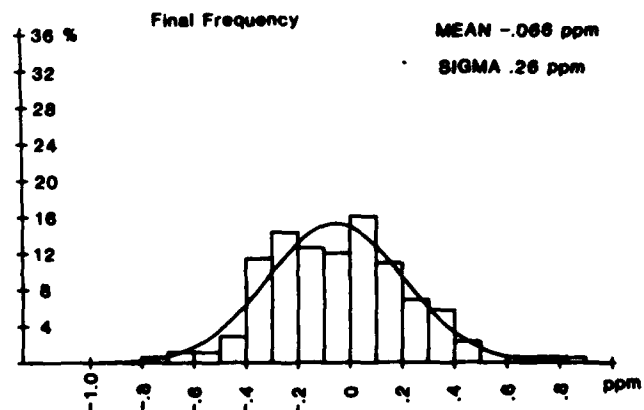


Fig. 5. Histogram of final frequency adjustment

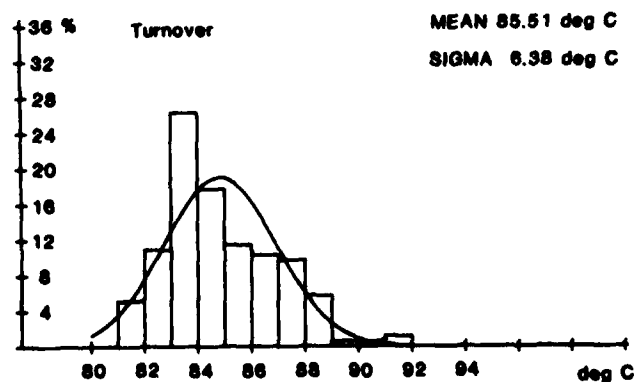


Fig. 6. Histogram of turnover temperature

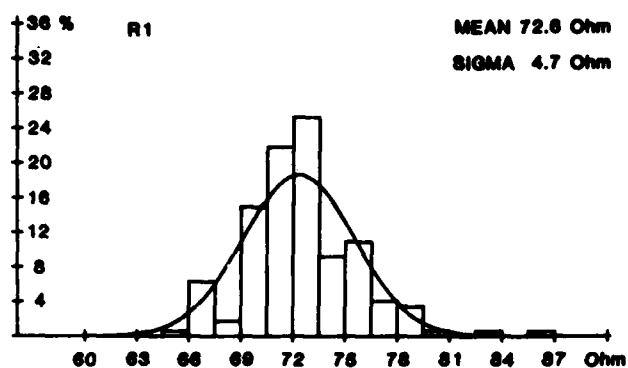


Fig. 7. Histogram of series resistance R_1

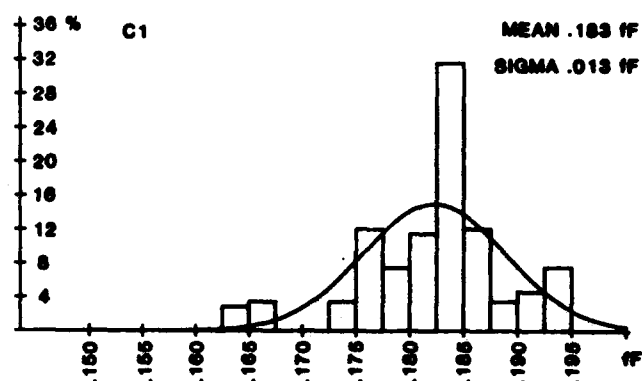


Fig. 10. Histogram of motional capacitance C_1

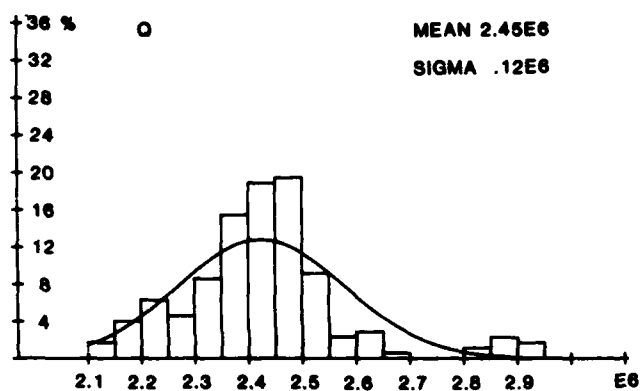


Fig. 8. Histogram of quality factor Q

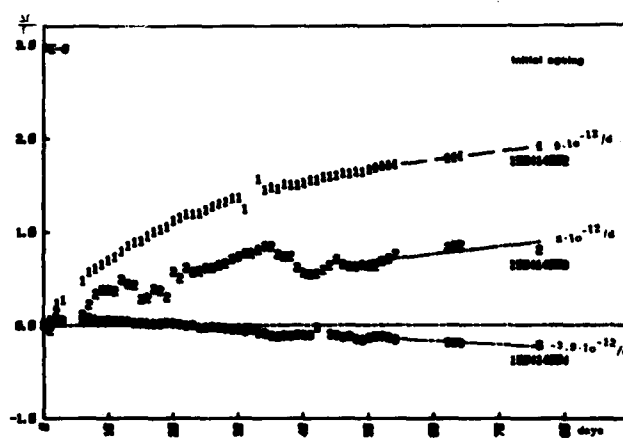


Fig. 11. Initial ageing

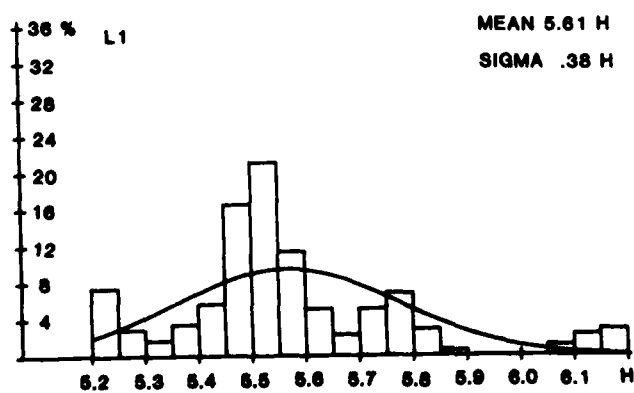


Fig. 9. Histogram of motional inductance L_1

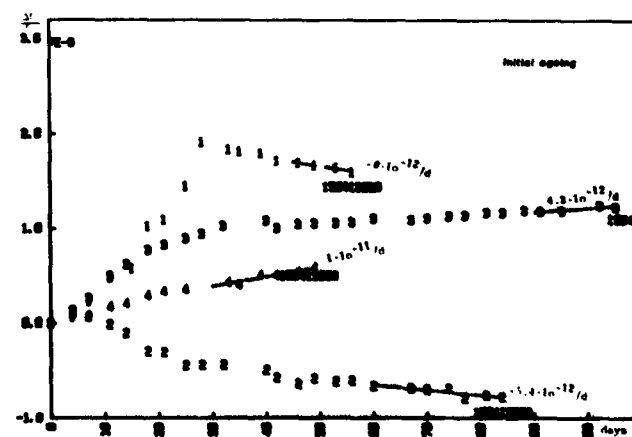


Fig. 12. Initial ageing

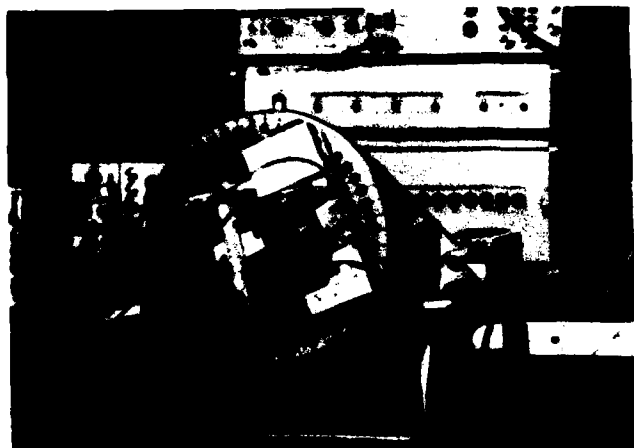


Fig. 13. Turntable for 2 g sensitivity measurements

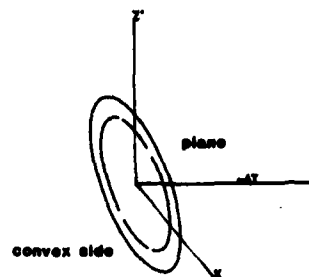


Fig. 14. Axonometric plot of BVA blank in general position (the diameter corresponds to $5 \cdot 10^{-10}/g$).

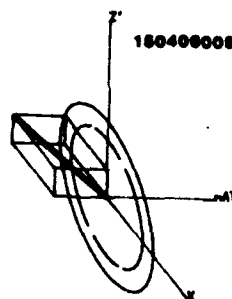
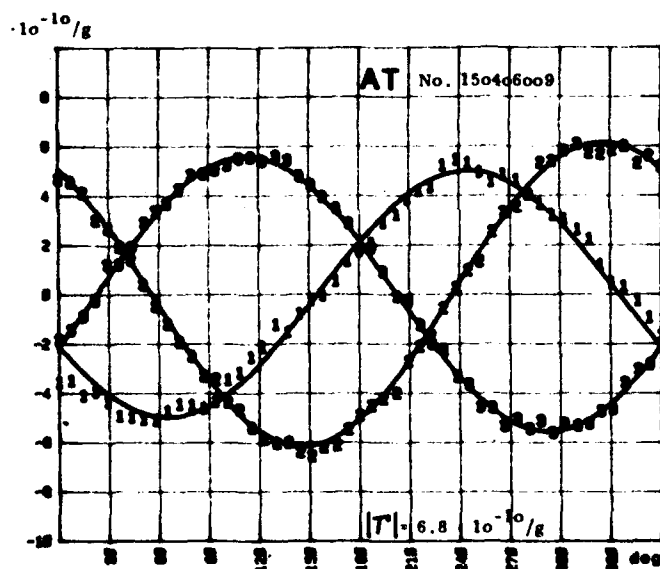


Fig. 15. r vs ψ and \hat{r} in blank coordinates for an early unit

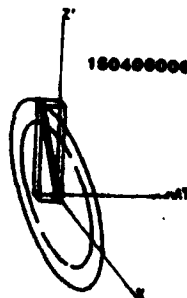
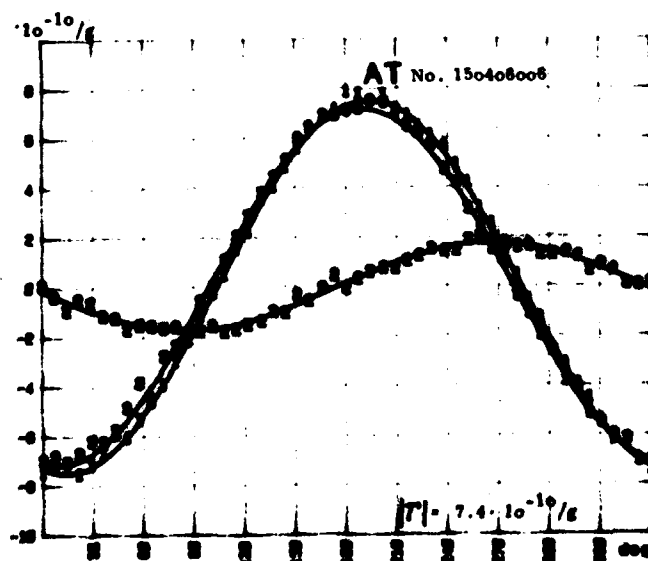


Fig. 16. r vs ψ and \hat{r} in blank coordinates showing extreme clip influence

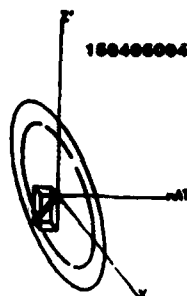
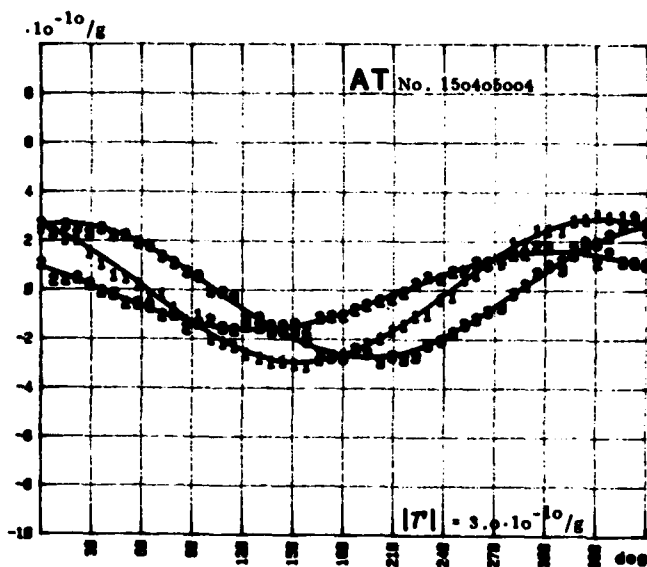


Fig. 17. r vs ψ , improved clips

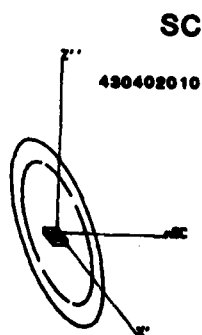
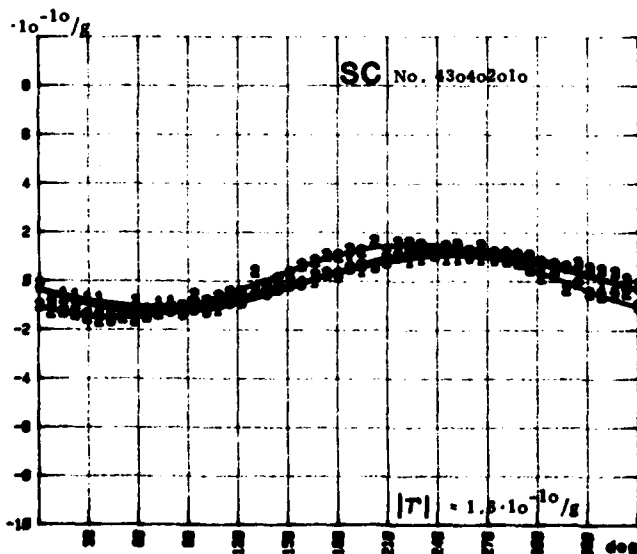


Fig. 18. r vs ψ for a 3rd OT SC BVA resonator (typical value)

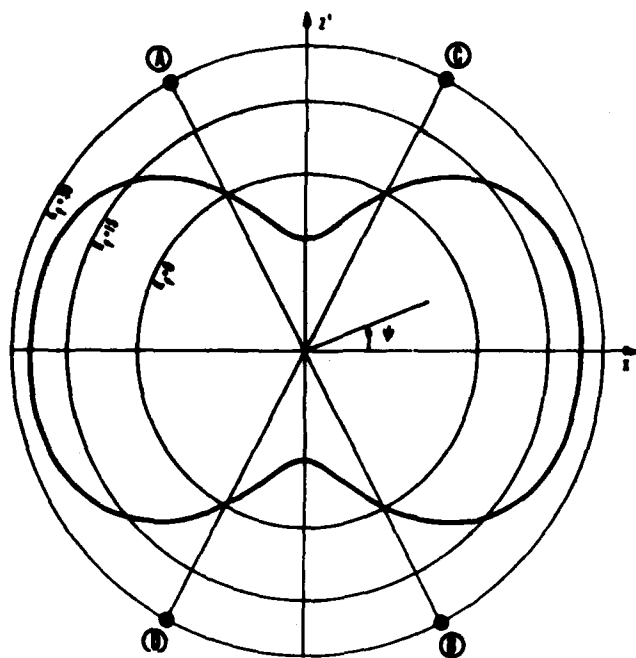


Fig. 19. Bridge location at the zero crossing of $K_F(\psi)$ to minimize r

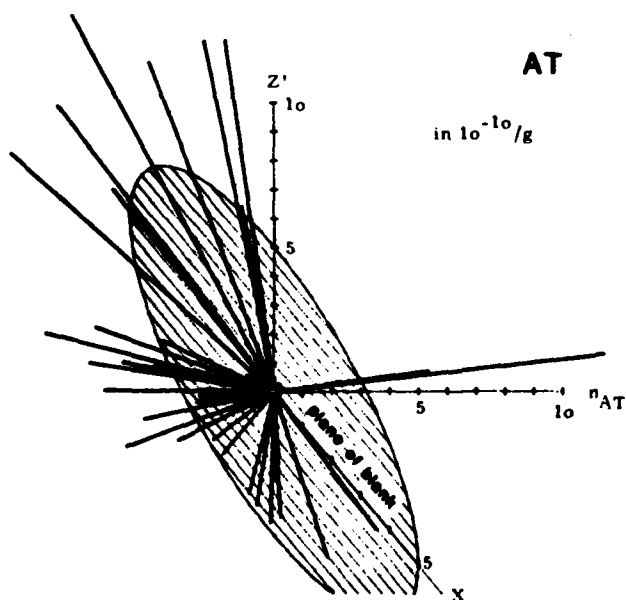


Fig. 20. Summary of 50 r measurements shown in a general position

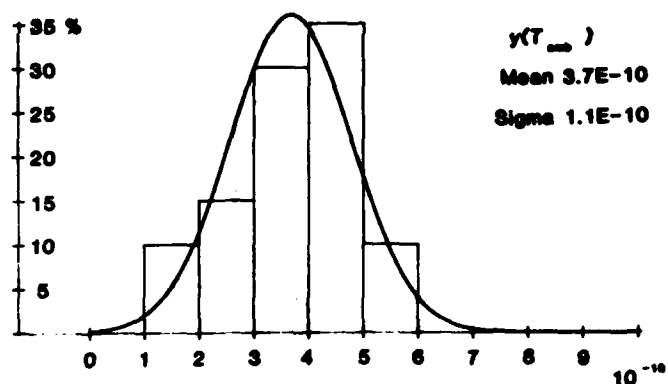


Fig. 21. Histogram of relative frequency deviation vs. ambient temperature in the range from -30° to $+70^\circ\text{C}$

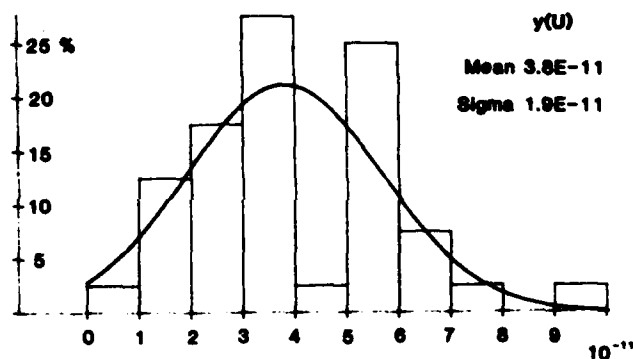


Fig. 22. Histogram of frequency vs. supply voltage change of $\pm 10\%$

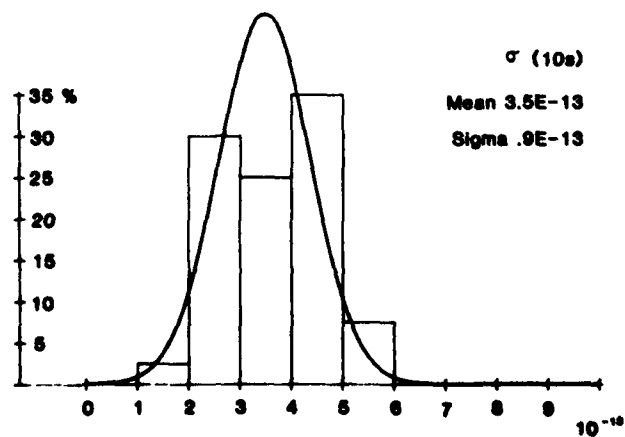


Fig. 23. Histogram of two sample Allan variance for $\tau = 10$ sec

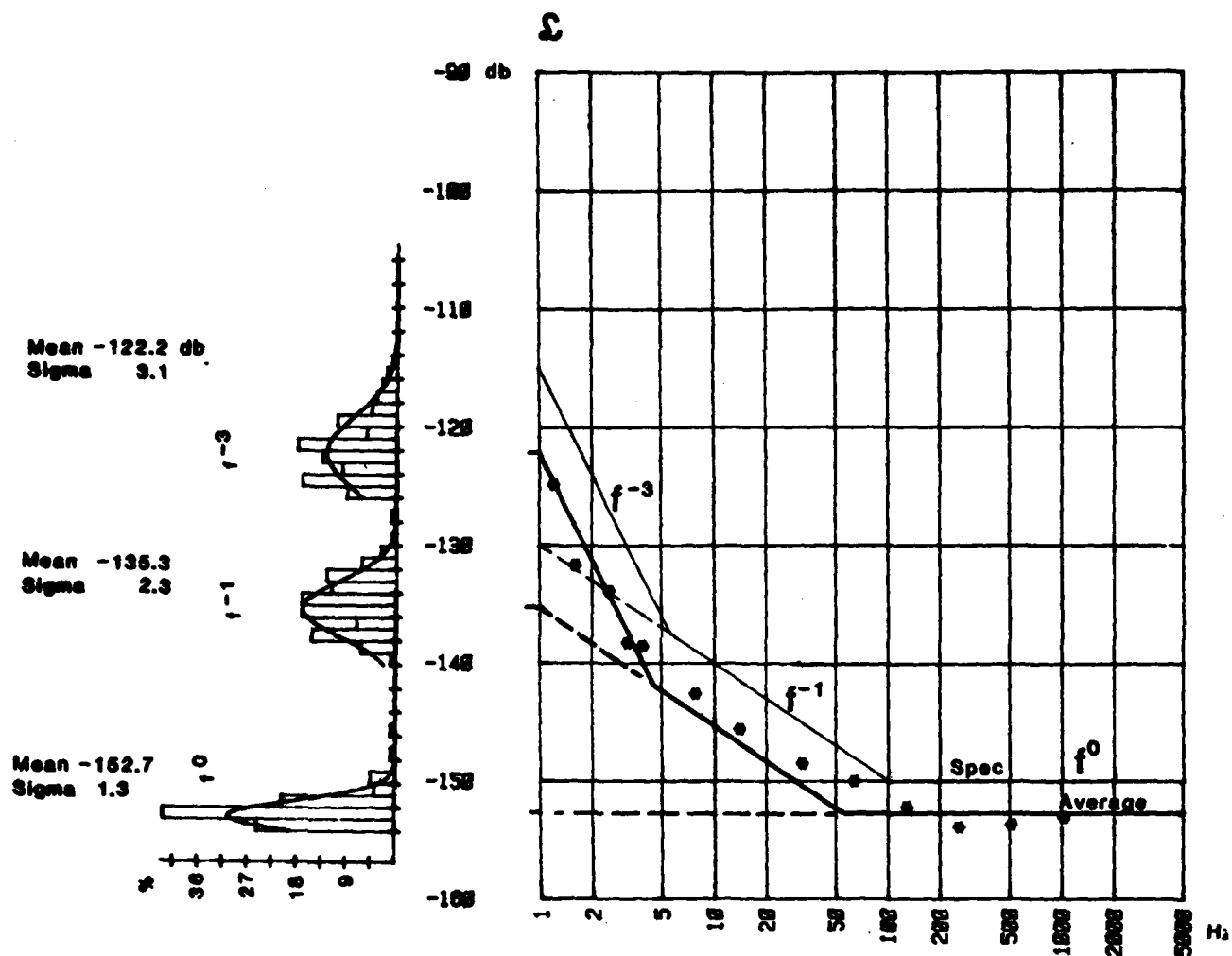


Fig. 24. Σ data with histograms of the intersection of the straight lines f^{-3} , f^{-1} and f^0 with the 1 Hz ordinate

AD P 002521

TEMPERATURE COMPENSATED CRYSTAL OSCILLATOR SURVEY AND TEST RESULTS

V. Rosati, S. Schodowski and R. Filler

Electronics Technology & Devices Laboratory (ERADCOM)
Fort Monmouth, New Jersey 07703Abstract

→ To assess the availability and quality of moderate and high precision TCXO, a variety of oscillators was purchased and tested. Tests included frequency-temperature stability and aging. Oscillators in the 1 to 5 ppm stability range were found to conform to specifications, except that aging characteristics were widely variable. Oscillators between 0.5 ppm and 1 ppm did not fare as well. Significant percentages were out of specification. Aging for these oscillators was better, but still disappointing in many cases. No correlation between price and performance was found. Users are cautioned to carefully specify and test candidate oscillators before designing them into systems. ←

Introduction

In response to requests by various military systems managers and designers, a survey of crystal oscillator manufacturers was conducted. Sample lots were then ordered and tested.

Survey

All known oscillator manufacturers (107 companies) were solicited by letter to provide price and delivery information on two classes of oscillators. First, an oscillator was sought to fulfill a specific system requirement. The abbreviated specifications were: 21.9 MHz; ± 5 ppm F-T stability over the temperature range -45 to +75°C; 1 ppm first month aging; specified case size and pinout; and the usual Frequency-Voltage, Frequency-Load, etc., stability. Second, each vendor was asked for price and delivery information on their "best available" TCXO. Operating characteristics (frequency, stability, waveform, load, etc.) were all vendor options.

There were 14 positive replies, i.e., 14 vendors indicated that either or both classes could be supplied.

Sample quantities of 5 each were ordered from 13 of the responding vendors. The basis for selection was no exception to the specifications of the 21.9 MHz oscillator or exceptional performance claimed for the vendor-specified oscillator. Available funds limited the quantities purchased to 5 each of each kind from each vendor.

Specifically, 30 of the 21.9 MHz oscillators were ordered. Delivery time ranged from 24 to 32 weeks.

Fifty vendor-specified oscillators were also

ordered. It is interesting that at the time of this writing, one year after the initial order, two vendors have not yet shipped their "best available" oscillators. A third vendor had shipped 10 months after order but those oscillators arrived too late to be included in this report.

In addition to the 80 oscillators purchased outright, 25 oscillator boards were removed from new Army radios which require 1.6 ppm overall frequency accuracy (0.5 ppm allotted for temperature effects). These 25 oscillators were tested and results are included in this report.

Therefore, the data to follow is based on a cross section of moderate (5.0 ppm) to high (0.5 ppm) stability TCXO.

The price range for all oscillators was between \$48 and \$900 each.

Testing

All 105 oscillators were subjected to the following tests: Frequency-Voltage; Frequency-Load; Frequency-Temperature; and Aging.

The Frequency-Voltage and Frequency-Load tests were straightforward and followed the procedures recommended in the current Specification for Quartz Crystal Oscillators (MIL-O-55310).

Frequency vs. Temperature (F-T) Stability

The F-T tests were conducted as follows: The oscillators under test were placed in a temperature chamber at room ambient and quickly cooled (within 10 minutes) to -45°C. The chamber was held at the low temperature for two hours after which time the "temperature run" commenced. Temperature was increased, stepwise, at a 1 degree per four minute rate until the maximum operating temperature was reached. Temperature was then decreased at the same rate to -45°C. Frequency readings were recorded at one degree intervals.

Upon completion of the run, the F-T behavior of each oscillator was plotted. Also presented on each graph were calculations to determine a set point (at 30°C) that would center the frequency excursions about the oscillator's nominal frequency.

Stability was calculated from the formula:

$$\text{Stability} = \frac{F_{\text{max}} - F_{\text{min}}}{2} \quad (1)$$

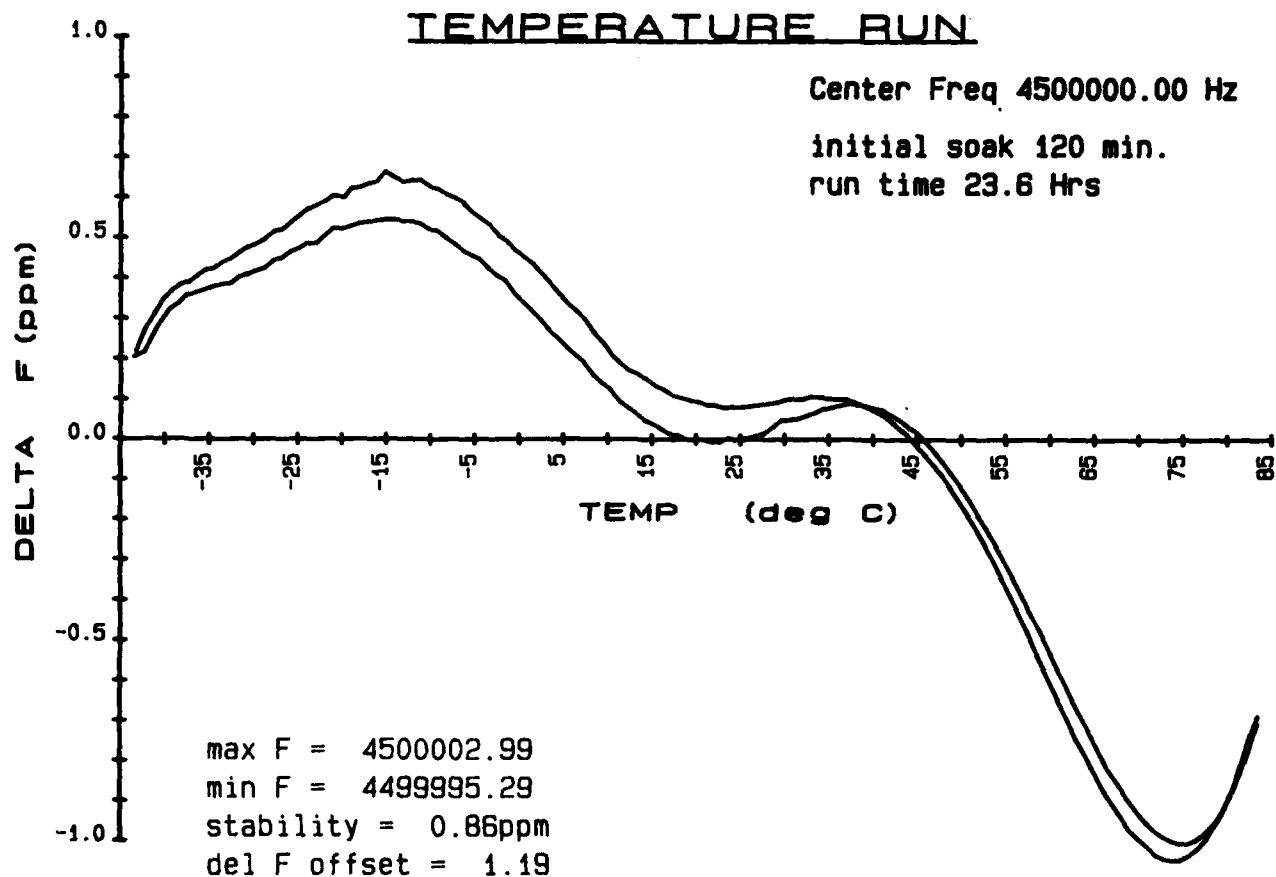


FIGURE 1. REPRESENTATIVE F-T CURVE

Since temperature was changed slowly, approximating steady-state conditions, any difference between the two F-T curves (temperature increasing vs temperature decreasing) was interpreted as thermal hysteresis.

A representative F-T curve is shown as Fig. 1.

Aging

Oscillators were aged while being held at $60^{\circ} \pm 0.05^{\circ}\text{C}$ for a minimum of thirty days. The oscillators were allowed a 24 hour stabilization period before the first data point was recorded. Subsequently, frequency measurements were made daily. At the conclusion of the aging measurement cycle, frequency vs. time graphs were plotted for each oscillator. Figures 2 and 3, respectively, represent "well behaved" and "erratic" performance. About 60% of the oscillators were well behaved.

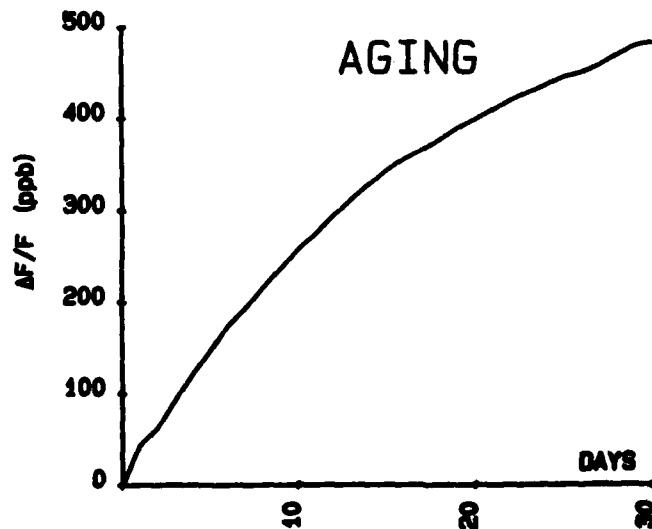


FIGURE 2. TYPICAL SHAPE OF THE AGING CURVE OF A "WELL BEHAVED" OSCILLATOR.

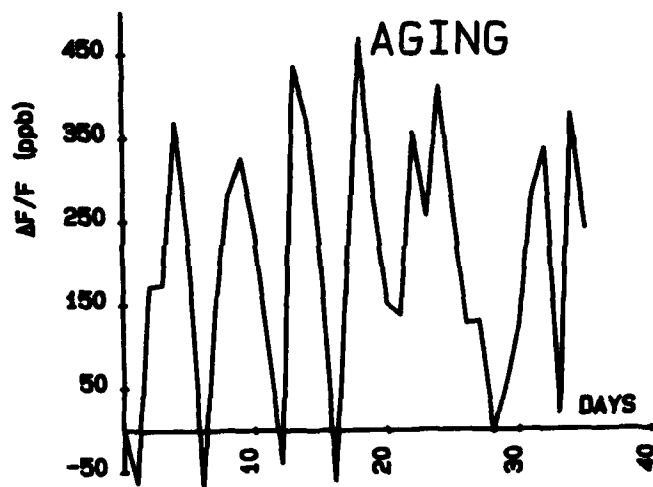


FIGURE 3. TYPICAL BEHAVIOR OF AN OSCILLATOR WHICH AGES ERRATICALLY.

Test Results

All 105 oscillators passed the Frequency-Voltage and Frequency-Load tests.

Table I is a summary of the performance of the 21.9 MHz, ± 5 ppm F-T stability, oscillators. (It should be pointed out that the oscillators were tested "as received". No frequency adjustments were made, nor were tests yet performed to determine the magnitude of the "trim-effect".)

Table II is a summary of the performance of the vendor-specified "best available" oscillators.

It should be pointed out that here, also, the oscillators were tested "as received".

The data from Tables I and II were further reduced to a pass/fail matrix which constitutes Table 3.

TABLE I

RESULTS: 21.9 MHz OSCILLATORS

NOMINAL SPECIFICATIONS:

21.9 MHz: 5 PPM (-45 TO 75°C); 1 PPM/MO AGING

GROUP	PRICE \$	F-T STABILITY (PPM)		HYSTERESIS (PPM)		AGING (PPM/MO)
		AVG	WORST	AVG	WORST	
1	513	.85	3.4	.61	1.08	.246*
2	480	1.1	1.8	.21	.37	.036
3	600	1.7	3.6	.45	.78	.360
4	225	3.0	5.5	2.0	4.7	.288
5	400	3.1	3.6	1.1	1.9	> 6
6	239	3.2	4.2	1.3	3.1	2.73*

* ONE UNIT CEASED OPERATING DURING AGING TEST.

TABLE II

RESULTS: VENDOR SPECIFIED OSCILLATORS

GROUP	PRICE (\$)	F-T STABILITY (PPM)			TEMP RANGE (°C)	FREQUENCY (MHz)	HYSTERESIS (PPM)	AGING (PPM/MO)
		CLAIMED	AVG	WORST				
1	266	0.5	.62	1.07	-45 TO +85	4.5	.15	.36
2	812	0.5	.47	.58	-55 TO +85	9.9	.20	.30
3	315	0.6	1.4	1.6	-45 TO +75	1.0	.48	.5 (E)
4	668	0.6	.71	.98	-40 TO +75	3.3	.18	.34
5	385	1.0	.91	1.15	-40 TO +75	3.5	.18	.18
6	66	2.0	1.2	1.82	-40 TO +95	14.4	.14	.57
7	48	3.0	1.9	2.5	-30 TO +70	3.6	1.44	.25 (E)
8	209	0.5	NOT DELIVERED		-45 TO +75	5.0	----	----
9	488	0.5	NOT DELIVERED		-55 TO +85	3.2	----	----
10	900	0.2	NOT DELIVERED		-45 TO +85	5.0	----	----

(E) = ERRATIC

TABLE III

SUMMARY OF ALL OSCILLATORS

STABILITY CLAIMED (PPM)	F-T STABILITY			AGING/MO (PPM)
	#TESTED	#PASSED	#FAILED	
5	30	29	1	0.2 (Avg)
3	5	5	0	.25 (E)
2	5	5	0	.57
1	5	4	1	.18
0.6	5	0	5	.34
0.6	5	4	1	.5 (E)
0.5	5	2	3	.30
0.5	25	9	16	.36

(E)= ERRATIC

Note that the likelihood of finding good units diminishes as the specified stability gets better. (Fifteen of the oscillators that were ordered, 10 at 0.5 ppm and 5 at 0.2 ppm, have not yet been delivered. This puts the vendors approximately five months behind their own delivery schedule.)

Interesting plots of performance vs price are shown in Figs. 4 and 5.

Fig. 4 is a plot of F-T stability vs price. Note the lack of correlation.

Fig. 5 is a plot of 30 day aging vs price. Again, there is a striking lack of correlation. For these oscillators good F-T performers are not necessarily good agers. That is, there is also no good correlation between Figs. 4 and 5.

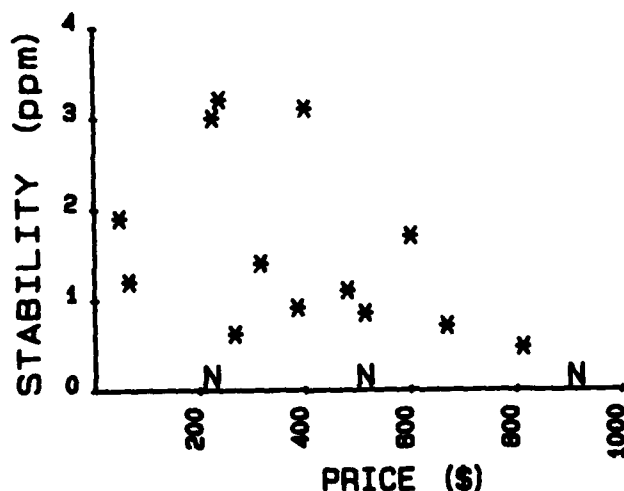


FIGURE 4. FREQUENCY STABILITY VS. PRICE. "N" SIGNIFIES OSCILLATORS NOT DELIVERED.

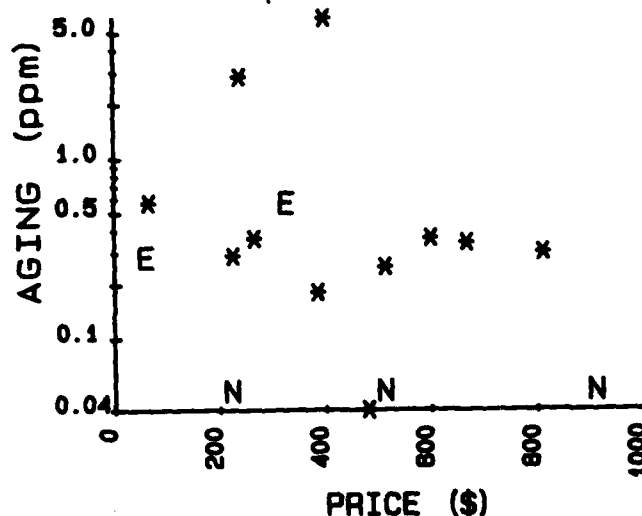


FIGURE 5. AVERAGE AGING VS. PRICE. "N" SIGNIFIES OSCILLATORS NOT DELIVERED. "E" SIGNIFIES ERRATIC AGING BEHAVIOR.

Conclusions

Oscillators in the 5 ppm to 1 ppm category, with average first-month aging of 0.5 ppm, appear to be readily available. However, even in this moderate stability category, systems designers and manufacturers are cautioned to provide careful quality assurance inspections.

Oscillators in the 0.5 ppm class appear to be at the very edge of the state-of-the-art. Great care should be exercised in the specification and testing of these oscillators and it is suggested that supporting data be requested for each oscillator of this class. A user, therefore, ought to expect to pay a few dollars for the extra testing, but it appears that the payoff in troubleshooting time saved and life cycle cost

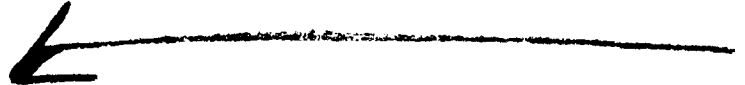
reduction could be well worth the investment. On the other hand, price alone does not seem to be a realistic guide to expected performance.

It is probable that many of the higher stability units were selected for best performance. It is also probable that those vendors who have not yet delivered are still in the selection process. Moreover, our solicitation letter clearly identified the buying organization and did not conceal the fact that the purchased oscillators would be tested. In view of this, one should be doubly cautious when contemplating the use of high stability TCXO.

This testing program will continue. "New" and "better" oscillators will continue to be sought and evaluated.

Acknowledgements

The authors thank Mr. Peter Thompson of the Frequency Control Branch at Fort Monmouth for his diligence and care in setting up the oscillator tests.



A NEW FREQUENCY CALIBRATION SERVICE OFFERED BY
THE NATIONAL BUREAU OF STANDARDS

G. Kamas and J. L. Jespersen
Time and Frequency Division, National Bureau of Standards
Boulder, Colorado 80303

Summary

AD P002522 After many years of helping calibration laboratories solve frequency and timing problems on an ad hoc basis, the Time and Frequency Division of the National Bureau of Standards is offering a new frequency calibration service that is specifically matched to the user's needs. This new service is intended to take advantage of the latest in measurement instrumentation, while at the same time keeping the user costs low.

The new service consists of identifying the calibration problem, obtaining and installing the necessary measurement equipment, and finally, a training seminar for the user. The user then obtains corrections for the accumulated calibration data via a telephone modem to the NBS frequency measurement computer. The service can use low frequency radio signals, or Loran-C signals, for NBS traceability.

Introduction

A frequency calibration means that a frequency source, a crystal, rubidium, or cesium oscillator, is measured using a standard reference frequency. Data are recorded, but, for several reasons, precision oscillators are normally not adjusted during each calibration. A measurement is made and a record is kept, but adjustments are made only infrequently. As a result, considerable record keeping is involved. The new NBS Frequency Calibration system automatically measures and records these calibration data.

Historically, the user of a high quality frequency standard has purchased equipment, read the manuals, and then proceeded with the calibrations. Although this approach to frequency calibration has changed little in several decades, we at the NBS have noticed a change in the needs of frequency calibrations users. Our telephone calls, letters, visitors, and attendees at NBS seminars have shown us that frequency calibration users need a different kind of service. That's why this service will emphasize a different approach to frequency calibrations. This approach is based on NBS experiences in recent years, dealing with a wide variety of users at many levels of accuracy.

Background

To start the discussion of the new NBS Frequency Calibration service, consider the history of frequency calibration. Most laboratories simply tune WWV on short wave, and zero-

beat a local oscillator against the WWV signal. This is the most widely used frequency calibration technique and is quite adequate for many users. The next step toward higher accuracies (better than a part in 10^7) involves using a VLF or a Loran receiver (Figure 1.).

The new frequency calibration system is not so much a change in what is being used, but how it is being used. A calibration laboratory must decide what equipment to buy, what it will cost, how it can be made traceable, how much manpower it will involve, and so on. Once these decisions have been made, the equipment is purchased and installed without much trouble. However, the new NBS Frequency Calibration service does not simply limit itself to equipment purchase and operation. Training for standards laboratory personnel is provided, as is a direct telephone data link to the NBS. After the equipment is installed and running, the user can be assured that the data are correct and that a traceable calibration has been performed.

As mentioned previously, the need for both training and follow-up has revealed itself over many years of customer contact. NBS has addressed this need in several ways. The two most important steps have been the offering of frequency calibration training seminars at Boulder and the publication of user-oriented technical manuals. Technical Note 559 is in a second printing and more than 8000 copies have been distributed to aid users in making choices among several different frequency calibration techniques.

Two separate frequency calibration seminars have been offered. One deals with frequency standards and clocks. It discusses time scales, international agencies for time and frequency, short and long term stability, and noise in oscillators; and reviews commercial standards, timekeeping, and clock modeling. A second seminar covers the calibration of oscillators, the use of frequency counters, the day to day measurement and adjustment of frequency standards and the use of Loran-C and VLF signals (Figure 2.)

Even after distributing several thousand copies of the NBS User's Manual and giving many seminars, there was still a missing element. The attendees many times went away with the idea that many frequency calibration techniques could be used, but they were not always quite sure how to use them. This same problem was revealed by our close contact with the manufacturers of WWVB, WWV and Loran-C receivers.

Oscillator manufacturers and others who saw this problem would often tell their customers to call NBS regarding frequency traceability.

As the equipment purchased during the 50's and 60's gradually became obsolete, users indicated a desire to upgrade their laboratory equipment. NBS is an active member in the National Conference of Standards Laboratories. The membership of that group was most vocal in wanting higher quality calibrations. As new equipment is specified and purchased, the level of calibration accuracy moves up from parts in 10^7 or 10^8 , towards parts in 10^{11} . The new NBS Frequency Calibration service can accommodate all of these requirements. When companies invest in new lab equipment, they want it to meet their needs as far into the future as possible. But as they increase calibration accuracy, operator skill becomes more of a factor. These are the two main areas the new NBS Frequency Calibration system addresses: what to buy, how to use it, and more important how to insure good traceability at high accuracy levels without becoming overly labor intensive.

This is an important feature. Users want their frequency calibrations to be precise and trouble free. Many laboratories cannot afford to spend large amounts of time on frequency calibrations since they usually have other demands for available resources. The need is for an automatic system where the operator isn't required to learn a complex procedure. A suggestion was once made to call this kind of calibration service a utility, like a telephone. One can use it without much training, and if it stops working, someone else is concerned about it. The NBS goal was to deliver state-of-the-art calibrations of frequency in a way that did not unduly burden the user. It is worth noting that frequency calibrations are relatively easy to automate, even at distances of thousands of miles. This cannot be said of all the other basic standards.

In the course of a year, the NBS Time and Frequency Division handles thousands of inquiries relating to frequency calibrations. These range from the most casual use of WWV to the most sophisticated frequency users wanting to increase their calibration accuracies. Many of these users don't necessarily want superior accuracy but rather want advice on installing and operating a simple system that will not take much time to understand. Often the NBS response to these inquiries is to refer them to manufacturers of equipment that will meet their needs. NBS maintains a list of manufacturers just for that purpose. Listed are all those manufacturers of Loran-C receivers, WWV receivers, GOES satellite equipment etc., known to NBS, but without any NBS endorsement or certification of performance. This is done as a service to the manufacturers and the frequency calibration users.

For many years NBS has broadcast frequency calibration signals from several radio stations. This has been a very popular service. NBS has a close working relationship with most of the

manufacturers in the frequency calibration industry and is active in standards labs organizations. Publishing calibration manuals and giving seminars have been useful activities and they have been well received. This new NBS Frequency Calibration service is also a natural step made possible by the low cost computers and telephone modems coupled with the fact that high accuracy radio signals are available nearly everywhere.

A New Frequency Calibration Service

Here is a description of the steps involved in this new NBS Frequency Calibration service. First, there is initial contact with a user. This is followed by attendance at an NBS seminar to help the user decide on choices for signals and equipment. Next, the user obtains equipment that is compatible with the NBS recommended data gathering format, and finally, the user becomes part of an NBS frequency calibration information network. This network has two parts. The first part is the receipt (by mail or computer terminal) of the current frequency data required for laboratory traceability. The second part is an on-line telephone modem link to the NBS for two-way data exchange to verify that the calibrations are correct.

If the initial contact goes beyond a simple suggestion or advice on how to better utilize existing equipment, NBS offers the option of having the user attend one of the NBS Time and Frequency Seminars. The initial contact will usually suggest the level of the difficulty and a possible solution. From experience, and especially dealings with receiver manufacturers, and the NBS staff can determine what to expect in terms of signal reception problems. This information is exchanged with users and manufacturers so that an informed solution can be sought. In addition to suggesting contacting the manufacturer or his representative for help, NBS also suggests that equipment, receivers especially, be tried at the proposed site to detect interference and signal level problems. With many years of experience in these areas, NBS can usually locate problems and suggest alternatives.

An NBS frequency seminar is the second step of the process leading to an improved calibration service. The seminar serves several purposes. It allows the user to actually see the kind of equipment that is available. In the seminar meetings, many equipment options are discussed. Also, traceability to NBS is described, and the user learns what his choices will be in terms of operator involvement. Emphasis is placed on the data and its interpretation. By the time the seminar attendee leaves, his level of familiarity is such that he can make a more intelligent choice. The seminar also lets him see a variety of equipment used to record frequency calibration data. He sees a number of options in terms of signal sources, crystal, rubidium, or cesium. He sees Loran-C, VLF, or LF reception. He is also introduced to time-of-day systems and television methods, and

gets to see satellite systems at work. From this experience, the user is better able to decide what is best for his laboratory. He may well choose to stay with what he has, thinking that the possible added cost or complexity is not worth it.

If a decision is made to get a new system, the next step is obtaining the necessary equipment. This will be similar to the equipment used in the NBS seminars, and the software will match. After getting the equipment installed and working, the user can utilize the telephone data link to NBS and verify correct equipment operation. In addition, to simply exchanging recorded data, the telephone modem allows for system troubleshooting via the phone line. This has proved to be very useful for getting systems started.

To better explain the details of how this service operates, consider the following example. Imagine a user in Salt Lake City who has a frequency calibration problem. He wants to upgrade his frequency calibration service so that he can calibrate newer counters with high accuracy time bases. His actual measurement requirement is one part in 10^{10} , but he decides to exceed this, to allow for higher quality frequency calibration needs in the future. He talks to his manufacturers' representative and views several equipment demonstrations. However, he is unsure about how to achieve traceability, and at the representative's suggestion, calls NBS. From this initial contact, he learns that his options are to use WWVB at 60 kHz or Loran-C from the West Coast. To further increase his understanding, to train a member of his staff, and to see actual equipment in use, he registers to attend a NBS frequency seminar.

This user is looking for cost trade-offs and is concerned about how complex the system will be after it is installed and running. He wants to see real data being recorded and handled on a daily basis and wants to talk about his problem. So the NBS seminar is a good choice. At NBS, we have found that precision frequency calibration services have not always been well understood. Direct user training has been one result of attempts to solve this problem. Even though we now have technical literature that is fairly easy to read, hand-on experience is invaluable.

At the NBS seminar, the user sees Loran-C and WWVB equipment being used and the data from those receivers being plotted and analyzed (Figure 3.). He gets a feeling for the kind of data he will have to handle to achieve his target objective of parts in 10^{11} calibration accuracy. He sees equipment from a number of manufacturers in operation. During the seminar, he talks with other users and with them makes comparisons of the options available to him.

Finally, he decides that even though WWVB is receivable in Salt Lake City, he will select Loran-C as a calibration source and will use an automatic data logger (Figure 4.). This choice

gives him state-of-the-art frequency calibrations. At the NBS seminar, the user sees Loran-C tracking receivers running with little or no attention and also notes that the data logging system takes care of power outages and is designed to minimize operator interaction. At the other end of the telephone line, NBS can provide assurance (automatically) that the data being recorded is valid. The user then gathers up his seminar material and returns to his laboratory ready for the next phase, which is getting the equipment together. Depending on his choices, the equipment costs should remain modest (Figures 5 and 6). This reflects a primary design goal by NBS to make the system affordable, automatic, and easy to operate.

As his equipment is delivered, the user will work with his manufacturers' representatives to get it installed. Using NBS-compatible software, he can go at once into a debugging phase where his data can be compared with data recorded at the NBS. This provides the assurance that the calibrations are valid. The seminar training material will also contain examples of data plots and oscillator performance evaluation.

As questions arise, the user can "see" the performance of his Loran based system by plotting data obtained at his site. Conversely, NBS can also look at his data via telephone modem. This two-way data exchange is a most valuable characteristic since it can resolve many of the problems associated with precision frequency calibrations. This user will also have a Frequency Calibration Bulletin available to him. This is a published list of the performance of some accurate signal sources, such as WWVB, Loran-C, and several VLF stations (Figure 7). He can obtain these data by telephone and will also get a copy by mail at monthly intervals. Since the software is compatible, his plots will overlay the NBS plots and he can identify system troubles.

The user now has a high quality system that he understands and that fulfills his requirements. The system runs with the least trouble and inconvenience to him and can be upgraded as new developments emerge. If the user has staff changes, he can send additional persons to an NBS seminar for training. He will feel that he is a part of larger system and that help is available. This is in contrast to the situation in recent years where each user was essentially isolated.

Another benefit is built into the system. Suppose in our example that an air base near Salt Lake City has a contractor who wants to check some timing equipment. This new user can benefit from the data base aspects of the new NBS Frequency Calibration system. Each new user in an area can compare his data, via the NBS data base, with others. In this example, if two users are tracking Loran-C, any suspicious data can be tracked down and the cause located. In an area that has many users, the overall quality of calibrations will likely improve, especially

if some users maintain very high quality frequency standards.

Conclusion

The NBS feels that this new service plan will meet a critical need for frequency calibration users. It has been planned to provide all the necessary training, data logging, feedback, and follow-up. As improvements in technique are developed at NBS, they can be incorporated by substitution or modification of system software. If lower cost options or peripherals are developed, they too can be easily incorporated.

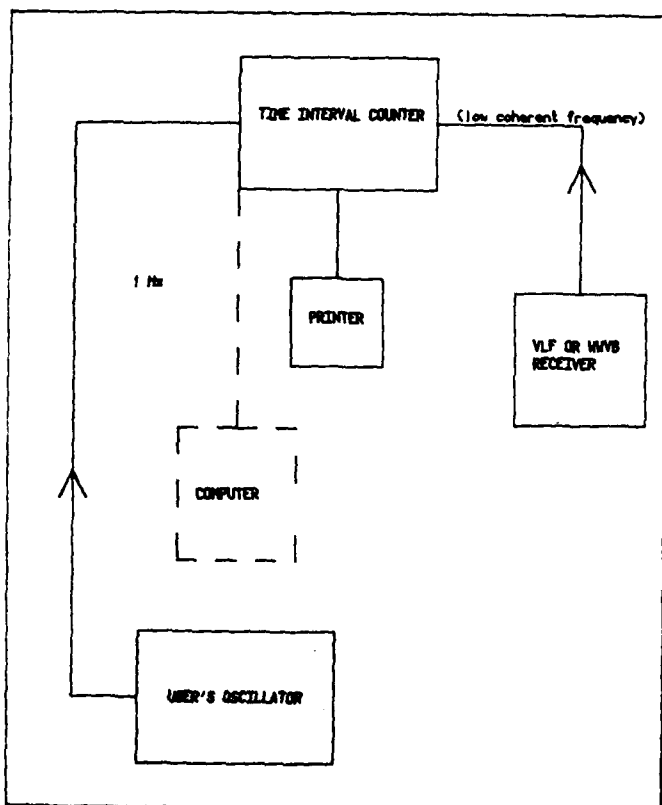


FIGURE 1. SIMPLE FREQUENCY CALIBRATION SYSTEM (COMPUTER OPTIONAL)

SEMINAR ON FREQUENCY MEASUREMENTS AND CALIBRATIONS

This is a new seminar that incorporates new material into the Time & Frequency User's Seminar given in previous years. It is intended for engineers and standards lab technicians involved in making frequency calibrations. The course will be taught at a practical level to satisfy those new in the field as well as more experienced users. Methods taught will use commercially-available equipment.

DATES: October 25, 26, and 27.

LOCATION: National Bureau of Standards
325 Broadway
Boulder, CO 80303

TOPICS COVERED: CRYSTAL OSCILLATOR CALIBRATION
APPLICATIONS OF FREQUENCY COUNTERS
HOW TO CHOOSE A FREQUENCY CALIBRATION SOURCE
CARE AND USE OF FREQUENCY SOURCES
USING LORAN-C AND MWB FOR FREQUENCY CALIBRATIONS
TIME AND FREQUENCY MEASUREMENT ASSURANCE SERVICES AT NBS
ORGANIZATION OF TIME AND FREQUENCY IN THE U.S.
NBS, USNO, AND OTHER PUBLICATIONS

Although all of the above subjects will be covered, emphasis will be on making practical frequency measurements and calibrations.

Registration fee includes teaching materials, supplies, coffee and refreshments, and certificate of completion.

REGISTRATION DEADLINE: October 11.

CANCELLATION: Registration fee will be refunded in full only if notice of cancellation is received at NBS prior to October 11.

MAKE CHECKS PAYABLE TO: SEMINAR ON FREQUENCY MEASUREMENTS AND CALIBRATIONS.

HOTEL INFORMATION: See page 2.

INSTRUCTIONS FOR ATTENDANCE: Complete and return attached application.

FOR FURTHER INFORMATION:

GENERAL INFORMATION: (303) 497-3212
TECHNICAL INFORMATION: (303) 497-3378

FIGURE 2. EXAMPLE OF SEMINAR ANNOUNCEMENT

FIGURE 3. EXAMPLE OF LORAN DATA AS RECEIVED AT THE NBS

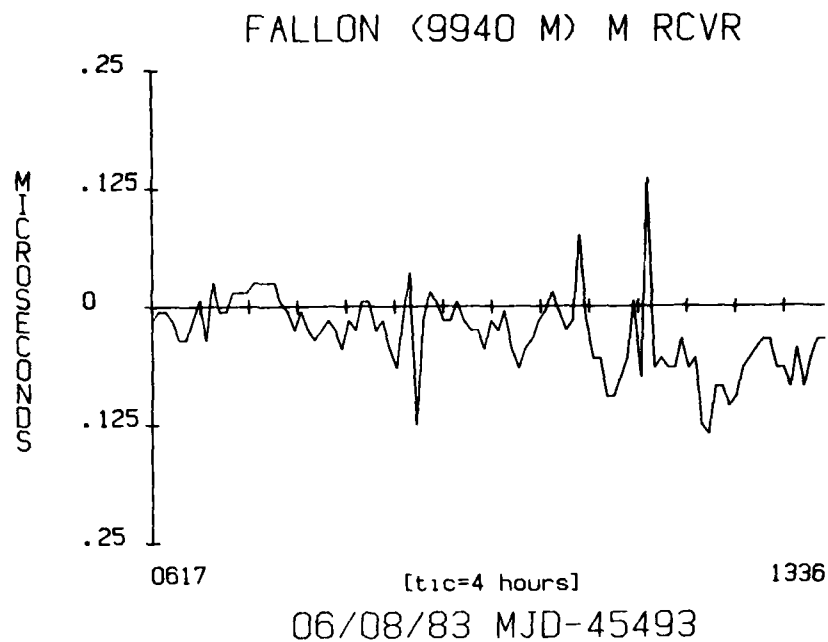


FIGURE 4. LORAN-C STATIONS IN RELATION TO A TYPICAL FIELD SITE

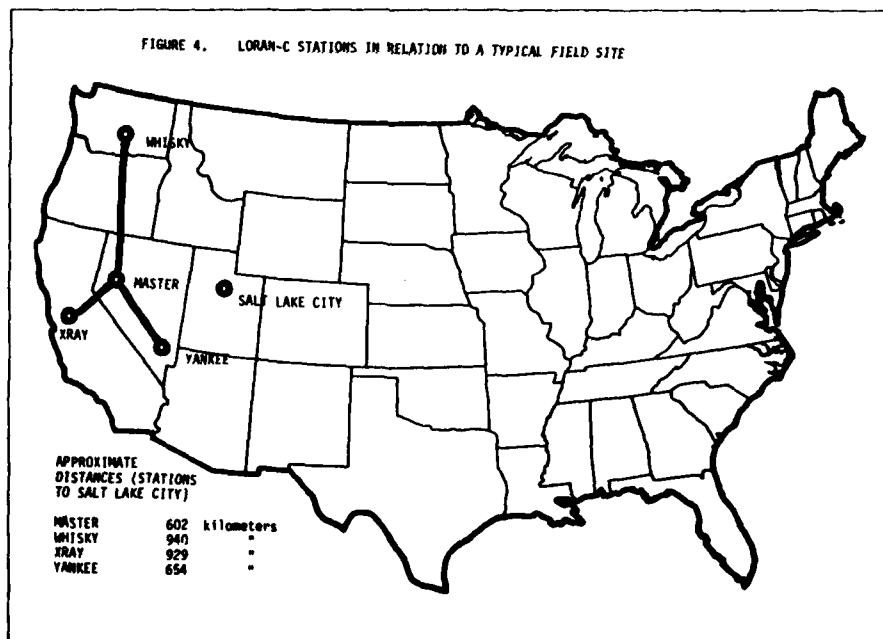


FIGURE 5. EXAMPLE OF LOW COST NBS FREQUENCY CALIBRATION SYSTEM

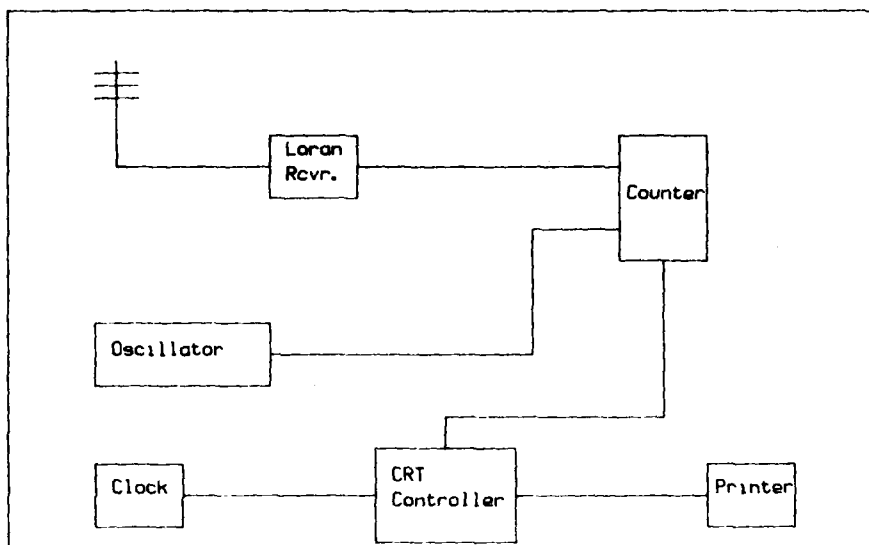
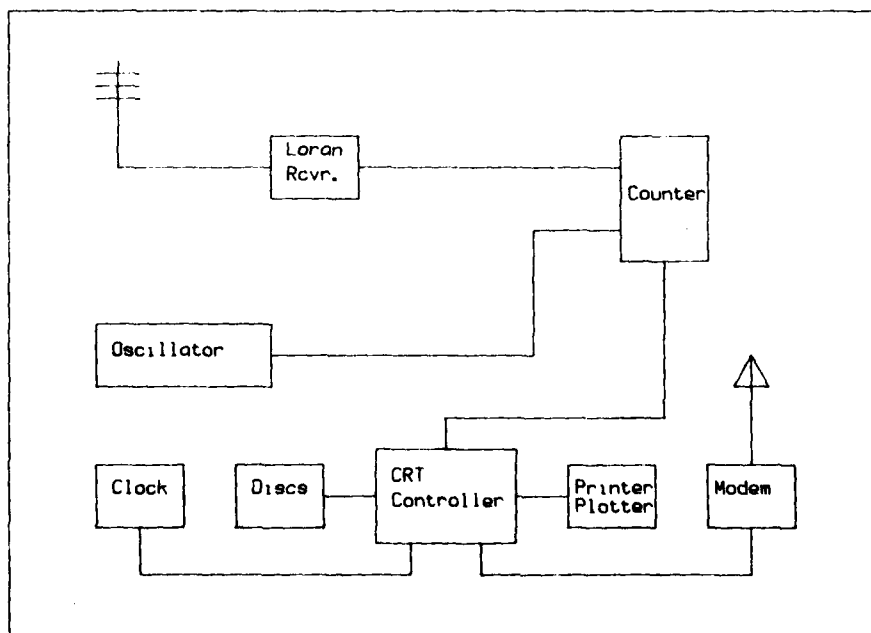


FIGURE 6. EXAMPLE OF NBS FREQUENCY CALIBRATION SYSTEM WITH MODEM AND PLOTTING CAPABILITIES



3. PHASE DEVIATIONS FOR WWVB AND OTHER NBS-MONITORED BROADCASTS

WWVB (60 kHz)

Values given for WWVB are the time difference between the time markers of the UTC(NBS) time scale and the first positive-going zero voltage crossover measured at the transmitting antenna. The uncertainty of individual measurements is plus or minus 0.5 microseconds. Values listed are for 1500 UTC.

Omega, North Dakota (13.1 kHz) and Omega, Hawaii (11.8 kHz)

Relative phase values are given for VLF stations and only the change from the previous available day's reading is published. Days when the data were satisfactory but readings were not taken (for example, on weekends or station maintenance days) are marked (-). If data were lost, continuity is also lost and the indication is (*), which means that reading cannot be compared to the previous day.

LORAN-C Dana, Indiana and LORAN-C Fallon, Nevada, (100 kHz)

Values for Loran-C (Dana) and Loran-C (Fallon) are the time difference between the UTC(NBS) time pulses and the 1 pps output of the Loran-C receiver.

APR	MJD	UTC(NBS) - WWVB (60 kHz) ANTENNA PHASE (in Microseconds)	UTC(NBS) - RECEIVED PHASE (in Microseconds)			
			OMEGA (11.8 kHz)	OMEGA (13.1 kHz)	LORAN-C (DANA) (100 kHz)	LORAN-C (FALLON) (100 kHz)
1	45425	5.80	(-)	(-)	62295.95	3947.01
2	45426	5.70	(-)	(-)	62295.83	3947.08
3	45427	5.70	+ 1.2	- 1.4	62295.84	3947.05
4	45428	5.71	+ 0.6	+ 0.1	62296.02	3947.05
5	45429	5.73	+ 0.9	+ 1.1	62295.93	3947.00
6	45430	5.69	+ 1.6	- 0.6	62295.94	3947.06
7	45431	5.79	- 2.2	+ 0.1	62295.89	3947.12
8	45432	5.70	+ 1.3	- 0.3	62295.78	3947.22
9	45433	5.72	- 0.8	+ 1.4	62295.64	3947.33
10	45434	5.74	- 0.3	- 0.1	62295.64	3947.48
11	45435	5.75	+ 0.1	- 0.6	62295.80	3947.57
12	45436	5.68	+ 0.9	- 1.5	62295.87	3947.57
13	45437	5.69	- 0.6	+ 1.1	62295.78	3947.61
14	45438	5.75	+ 2.3	- 0.8	62295.84	3947.62
15	45439	5.77	- 0.7	+ 1.4	62295.76	3947.66
16	45440	5.67	+ 0.1	- 0.2	62295.47	3947.80
17	45441	5.65	+ 0.9	+ 0.1	62295.47	3947.95
18	45442	5.64	- 0.3	0.0	62295.49	3947.95
19	45443	5.64	+ 1.2	- 0.4	62295.62	3948.02
20	45444	5.65	- 0.1	- 0.1	62295.56	3948.06
21	45445	5.72	- 0.4	- 0.3	62295.62	3948.12
22	45446	5.72	- 1.1	- 1.1	62295.71	3948.16
23	45447	5.72	- 1.9	+ 1.2	62295.42	3948.22
24	45448	5.72	+ 1.5	- 0.1	62295.41	3948.24
25	45449	5.72	+ 3.0	+ 0.6	62295.55	3948.28
26	45450	5.85	+ 0.8	- 0.7	62295.54	3948.34
27	45451	5.67	- 1.8	- 0.2	62295.58	3948.37
28	45452	5.66	- 0.9	+ 0.7	62295.57	3948.38
29	45453	5.68	+ 0.5	- 0.1	62295.63	3948.38
30	45454	5.68	+ 0.6	- 1.4	62295.52	3948.35

Values for Omega and Loran-C are as received at NBS, Boulder, Colorado. Values are 4-hour averages taken from 1600-2000 UTC daily. VLF data are reported exactly as seen. Some days show variations due to signal loss or station outages. The data are still useful.

-3-

FIGURE 7. AN EXAMPLE OF PUBLISHED FREQUENCY CALIBRATION DATA

AD P 002523

Specifying Performance for Atomic Standards

Joe White
U.S. Naval Research Laboratory
Washington, D.C. 20375

SUMMARY

As the needs for precise time, time interval and frequency continue to increase the demand for atomic frequency standards has also grown. This paper is designed to aid those whose program requirements now include these devices. There are several areas where specifications of atomic clocks vary significantly from crystal oscillator specifications. These include stability, environmental sensitivity, reliability, and testing. At this time, there is no general military specification for atomic standards.

The Naval Research Laboratory has extensive experience in procuring atomic standards in GPS and other programs. Examples based on experience are included. The level of detail required for particular applications is discussed.

INTRODUCTION

With the increased utilization of high technology hardware the demand for atomic frequency standards is increasing. For government agencies and system prime contractors who cannot use off the shelf commercial clocks there is no general military specification for use in procurements. As a result, each individual agency has generated unique specifications as required to meet its own needs. Of these, only one has been made into a MIL-SPEC (MIL-SPEC-28811(EC)). The purpose of this paper is to provide guidance, in the absence of a general specification, in the preparation of a specification for atomic clocks for a particular user application.

USING EXISTING SPECIFICATIONS

Many of the features desired in atomic clocks are sufficiently similar to quartz oscillators to allow use of portions of the existing crystal military specifications. These include such parameters as RF outputs, time ticks, pulse trains, synchronization, and quality control. Papers describing those specifications are included elsewhere in these proceedings.

It is also possible to use the framework from these specifications for guidance in parts selection, environmental stability, and testing. Problems generally arise in these areas due to the physics units in atomic clocks. These problems fall into four basic categories. They are

1. Nonstandard Parts
2. Environmental Limitations
3. Testing
4. Maintainability

NONSTANDARD PARTS

Unlike the crystal resonator used in oscillators, the physics packages used in atomic standards have yet to reach the degree of manufacturability which would allow it to be fully described in a military specification. There are a number of commercial vendors who maintain production lines of cesium or rubidium clocks and there are also vendors who are preparing to manufacture hydrogen standards. Each of these sources has a "recipe" for a reproducible physics unit. What actually goes into the finished product is a mixture of electronic and mechanical engineering and magic. The magic exists in the forms of trade secrets, patents and proprietary processes. It is these secrets which mean the difference between a viable commercial product and a laboratory curiosity. In most cases, the government did not pay for the development of the manufacturing technology and thus may have no clear claim to know its details sufficiently well to specify production.

Another problem that arises in this area is prediction of reliability. Reliability predictions are based on statistical evidence of performance. Many atomic standards are not produced in large enough numbers to produce a sound statistical base. If the proposed application requires either exceptional reliability or a good reliability prediction, only a few types of physics units which have been produced commercially in large numbers have a sufficient data base.

For example, consider a program which desires a 7.5 year lifetime for clocks. None of the cesium beam clocks currently under development for the program have been manufactured in a large enough quantity to calculate a MTBF with sufficient confidence. The approach for meeting the system reliability goal has two parts. The first is to require that the reliability for those subassemblies of the clocks for which conventional methods apply, be built under strict controls with the appropriate class of parts. The second is to apply what is known about life limiting factors in the cesium beam tube. A procurement package should include a requirement for potential vendors to show reliability estimates and the methods used to obtain them.

ENVIRONMENTAL PERFORMANCE

TEMPERATURE

Like crystal oscillators, the atomic standards show a change in output frequency with respect to ambient temperature. Over the normal operating range of the clock, the temperature coefficient is considerably lower than that achieved with a crystal. The operating range for some atomic standards is less than the full military

specification range. It is important that the specification match or exceed the expected temperature range for operating the clock and that the nature of the clock's performance near the expected extremes is specified and understood. Another area of concern in the specification is defining how the temperature coefficient is to be determined. The relationship between output frequency and ambient temperature is not usually linear nor can it always be expected to have the exact same shape in several clocks of the same design. It is possible that the temperature coefficient of a clock measured over a wide temperature range would show a much smaller change per degree than the same clock when measured over a smaller range. Some clocks will be used in environments where temperature changes are limited to a few degrees in an effort to improve performance. A narrow range instability would cause very poor performance if it fell within the operating range. The specification should include a method of measuring the temperature sensitivity that will ensure adequate performance.

ACCELERATION AND VIBRATION

Atomic clocks tend to be electronically and mechanically more complex than crystals. The physics units of rubidium and hydrogen standards use glass bulbs. Cesium and hydrogen clocks use precisely aligned magnetic state selectors. Rubidium and cesium clocks have been manufactured that meet the GPS launch environment vibration requirements and it appears likely that hydrogen clocks can also be made to survive those levels. It may be that it will not be possible to ruggedize these clocks sufficiently for extremely strong levels of vibration.

The limitation in applications where the clocks are required to operate and keep time through vibration and acceleration is not the physics package but the internal crystal oscillator. The servo control loops of atomic clocks are optimized for the time constant which gives the best performance over a broad range of conditions. When a rapidly varying rate of acceleration is applied to the crystal oscillator most servo systems cannot maintain frequency lock. It is possible to obtain improved performance by reducing the time constant, although the short term stability of the clock will be degraded. Specifications should include requirements of testing for stability during periods of vibrations if the application includes such operation.

VACUUM/ALTITUDE

There are several areas of concern when specifying operation in high altitude or vacuum conditions. Some commercial clocks designed for ground operation may not retain good thermal control in vacuum due to changes in the heat flow paths as the primary heat transfer mechanism changes from convection to radiation. Another problem is corona in the high voltage power supplies used to operate ion pumps in cesium and rubidium clocks. This problem is particularly serious if continuous operation is required from air to vacuum. Specifications should clearly state the intended operating pressure and should include requirements for testing in the expected operating conditions.

MAGNETIC FIELD

All atomic frequency standards show some change in frequency in the presence of external magnetic fields. To reduce sensitivity, all include magnetic shielding in the physics package. Applications should be designed to minimize exposure to strong fields. Specifications should indicate the strength of expected magnetic fields and should require testing to verify performance. If handling and storage specifications are included they should be designed to minimize possible magnetization of the internal magnetic shields.

STORAGE

Clocks using ion pumps in the physics units usually require periodic operation of the pump in order to maintain vacuum over a long period of time. Specifications should require vendors to provide a recommended schedule.

SHIPMENT

Hydrogen masers have a unique problem in shipment. Hydrogen gas is considered by all passenger air carriers to be hazardous cargo. In order to ship a maser, it is generally required that the hydrogen storage bottle be emptied and purged. Cargo only flights are slightly less restrictive. Specifications should encourage designs which facilitate meeting restrictions. Where practical surface transportation is recommended.

TESTING

Atomic standards present several challenges in the area of testing. There are specific problems related to testing individual areas of performance, there is a general problem of having suitable references available, and there is also a potential difficulty relating to the time required to adequately test some parameters. The factors relating to individual tests are detailed in the paragraphs related to the parameter in question.

The problem of having a proper frequency reference for testing varies in difficulty in inverse proportion to the stability required from the clock being tested. For instance, a quality crystal oscillator is an adequate reference for measuring phase noise and short term stability for some cesium and rubidium clocks. The only suitable reference for short term testing of a hydrogen maser is another maser. High performance cesium clocks are the best commercially available long term references. The testing procedure must specify the type of reference to be used for each test. By using LORAN C, TRANSIT, or GPS it should now be possible for all long term measurements to be traceable to the U.S. Naval Observatory.

Total test time is another factor that should be considered in developing a test plan. Measurements of frequency shifts due to environmental factors or aging can require a period of hours or a day for each data point. It may be necessary to strike a balance between acceptable levels of noise in the test data and delivery schedule.

MAINTAINABILITY

The primary concern in specifying maintenance

requirements for field use of atomic standards is understanding the level of capability needed for repairs or replacement of the physics unit. As a general rule replacement and most adjustments of the physics unit are high level maintenance operations which require skilled personnel. Potential vendors should be required to show the level of support personnel and ancillary equipment required for each level of maintenance.

PERFORMANCE SPECIFICATION

Atomic clocks show much better performance than crystal oscillators under most conditions. There are significant differences in performance among the several types in use. This section will attempt to outline those differences and indicate where care should be taken in specifying performance.

ACCURACY

The definition of the second is based on the cesium atom. Cesium beam standards are thus based on the same technology as the primary standards in use in some of the world's major timekeeping laboratories. The U.S. Naval Observatory uses an ensemble of high performance commercial cesium standards. Hydrogen masers are also capable of high accuracy. So far, there are not enough clocks of a given design in operation to make broad statements about accuracy, although individual clocks has shown excellent performance. Rubidium standards can be adjusted to operate at the correct nominal frequency but are not as inherently accurate as cesium. Specifications should state the accuracy required and should also state whether any comparison to external references is acceptable in order to meet the frequency requirement.

REPRODUCIBILITY

Along with accuracy it may also be desirable that the clock return to the same frequency after being turned off and then back on at some later time after exposure to particular environmental conditions. The range of performance in this respect is broader. In general cesium clocks are very good at returning to the same operating frequency for the same reasons as their accuracy. Depending on the conditions, hydrogen clocks can also show excellent reproducibility. The reproducibility requirements must state the conditions under which the performance is desired and should require the contractor to specify any alignment or tuning operations which might be required to obtain a given level of performance.

DRIFT

Low drift (change in frequency over time) is probably the single most popular reason for selecting atomic frequency standards over crystal oscillators. There is significant difference in drift performance between the various types. Cesium clocks do not produce measurable drift although frequency shifts may occur due to changing environmental factors or long term random walk phenomena. Hydrogen and rubidium clocks all exhibit some drift. The range is from parts in 10^{12} /day in some rubidium units to parts in 10^{15} per day or less in experimental masers. Drift specifications in rubidium and hydrogen clocks should include requirements for a test procedure which provides adequate measurement ability to detect the true

drift rate. If a burn-in period is required, the specification should allow for it.

SHORT TERM STABILITY

The hydrogen maser provides the best stability of the atomic clocks for measurement times of less than one day. Cesium clocks and hydrogen standards using automatic cavity tuning are the best long term performers. Rubidium clocks provide very good short stability for periods of less than 1 day but are usually less stable for longer periods due to drift. Specifications for short term stability should define the measurements periods of interest and the statistical method to be used to determine the stability from the raw measurement data. The Allan variance is the accepted method in the clock community for making stability estimations.

SETTABILITY

Settability is the capability of adjusting the output frequency of the clock to some desired offset from a reference. In the case where high intrinsic accuracy is the most desired feature, there may be little need to do such tuning. In other clocks it is usually required. The specifications should state to what precision the tuning is to be done and should also require that the method used not adversely affect the clock's performance.

WARM-UP

Warm-up times range from as little as 90 seconds in some rubidiums to about 1 day in some hydrogen masers. Warm-up times are based on the time required for the ovens to reach operating temperature and stabilize. If the application requires a rapid warm-up the specifications should state not only the desired warm-up time but also the acceptable power consumption during the warm-up period and the minimal acceptable performance at the end of the warm-up.

CONCLUSIONS

Atomic frequency standards are showing great potential as precision references of time and frequency in a variety of applications which will continue to expand. They are not, however, the answer to all problems nor will they operate at their best in all environments. Manufacturers of these clocks are particularly dependent on receiving thorough and complete specifications of the desired performance and the expected environment in order to deliver their best effort.

PLATFORM DISTRIBUTION SYSTEM SPECIFICATIONS

James A. Murray, Jr.

Naval Research Laboratory
Washington, D. C. 20375

Summary

Platform distribution systems form the link between a time dissemination system traceable to the U.S. Naval Observatory master clock system and the local systems that require precise timing. The distribution system serves the functions of establishing and maintaining accurate time and frequency and distributing it to the user systems.

Although the designs of these systems vary widely, there are certain technical factors that are repeatedly encountered. In this paper, these factors and other considerations that are important when designing or specifying platform distribution systems are discussed. Also presented are certain areas in which standardization could benefit not only platform distribution system implementation but the user systems development as well.

Introduction

Numerous platform distribution systems (PDS) have been developed and employed by users of precise time and frequency. For purposes of definition, a Navy description will be used: a platform distribution system maintains accurate time and frequency referenced to the U.S. Naval Observatory (USNO) and distributes them to user systems aboard the platform. While the size and complexity of a system that would perform those functions can vary widely, certain aspects are repeatedly encountered and occasionally neglected in specifying or designing platform distribution systems.

Standardization

One of the most important and difficult tasks is defining what services are to be provided. The source of the difficulty is a general lack of standardization in interfaces, from the standpoints of both format and performance requirements such as resolution, spectral purity, continuity, etc. Responsibility for meeting some of these performance requirements may be assigned either to the platform distribution system or to the user equipment.

At present, a multitude of time codes, pulse rates and frequencies are distributed aboard various platforms, which include land bases, ships, submarines and aircraft. Much of this proliferation is a result of designing for specific user applications instead of generalized requirements.

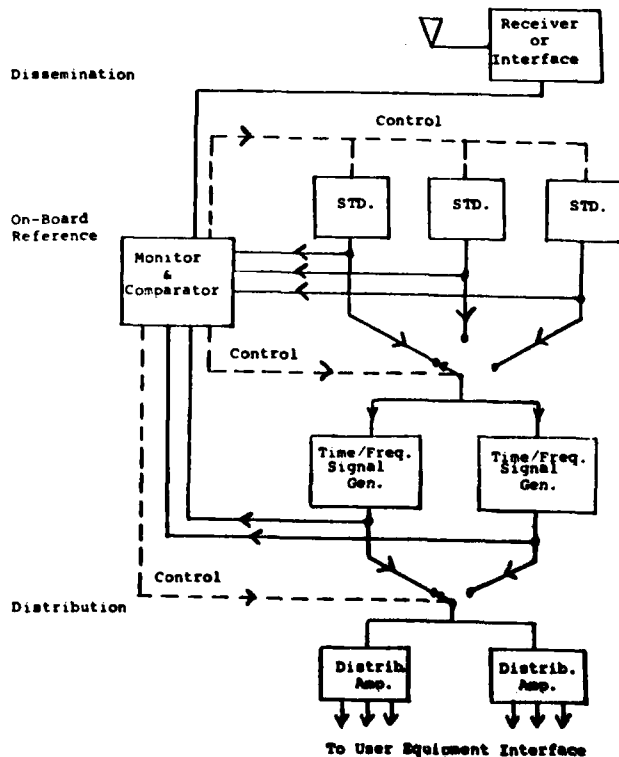
Some effort has been made towards standardization. DOD-STD-1399 (Navy) Section 441 (PTTI Interface Standard), for example, offers some guidance. The suggested formats are the ones employed by certain rather widely used Military timing equipment, but represent a de-factor standardization on 20-year-old requirements to serve then-existing user equipment. Although this standard should serve to reduce the proliferation of formats, present day technology can further reduce it and provide for more efficient distribution and more precise service to the user. MIL-STD-188-115 (Communications Timing and Synchronization) now under development may further refine the timing interface.

Functional Description

The three general functions of the platform distribution system (maintaining an on-board reference, receiving updates from the USNO through a dissemination system, and distributing timing to user equipments) are illustrated in an example of a PDS (Figure 1). While details may vary considerably, the example illustrates some functional areas that should be addressed in a more detailed specification. Multiple time and frequency standards are frequently used. Monitor and comparison functions are provided for manual or automatic selection of the on-line standard and for time or frequency adjustments. The dissemination equipment provides the connection with the USNO master clock. Distribution format generators (if used) may be redundant and are selected manually or automatically to drive the distribution amplifiers, which may also be redundant. Distribution cables or light pipes may be specified as part of the system when their routing is critical or redundant, and the dissemination equipment's antenna and cable may also form part of the specification.

Figure 1

PLATFORM DISTRIBUTION SYSTEM
(Example)



User System Considerations

User system concerns that affect the PDS specification include continuity of service; time and frequency accuracy and stability; timing resolution, timing update rate and ambiguity; and frequency spectral purity. The linkage between time and frequency services (phase relationships, how time is updated by steps or frequency adjustment, and magnitude of frequency or time adjustments) may also be a user concern. The effect on the user of a complete or intermittent loss of the time and frequency service can greatly influence the PDS specification and the design and cost of the PDS. Standardized interfaces, including signal format(s) and a complete characterization of the services provided by platform distribution systems can make the specification requirements much easier to arrive at both for the user system and the platform distribution system.

Platform Distribution System Considerations

These user equipment considerations translate into a set of system considerations for the PDS. However, the PDS considerations also include factors imposed by the type of platform and its operating modes and environment. Additionally, the special requirements of continuous operation for timekeeping, whether done in the PDS or the user system, impose requirements for redundancy, continuity, survivability, stability, accuracy and serviceability on part or all of the PDS. Of particular importance where redundant units of the PDS might be switched is the specification of allowable switching gaps or transients and the reliability of the monitor and control system that makes switching decisions.

Ordinarily, a PDS is specified in each factor to meet the requirements of the most demanding user. However, it may be prudent in some cases to transfer responsibility to the user for some especially difficult factors. Spectral purity and continuity through short interruptions, for example, may sometimes be recovered with a slaved clock or oscillator at or in the user equipment.

The required precision in time and frequency should be considered in terms of time accuracy, rate accuracy and stability and of frequency accuracy, stability and spectral purity. These are not necessarily functions of the onboard standards alone, but may also involve the dissemination and distribution systems.

Space and weight allowance may place limits on the performance of the PDS. However, it is frequently, but not always, true that platforms having shorter mission durations have less demanding timekeeping requirements.

Realizing that the dissemination system would generally be considered a non-continuous and perhaps vulnerable link, the timekeeping accuracy of the PDS during independent operation should be specified both as to free-running accuracy and the probability of maintaining continuous operation during such periods. A submarine, for example, may require on-board time standards of much higher stability and reliability than an aircraft on a short mission.

Unless the dissemination system employs two-way transmission or some other means of path-delay compensation, mobile platforms require some means of position determination. Such information may be obtained from the NAVSTAR Global Positioning System, Navy Navigation Satellite System or Loran C that also provide USNO time.

The time-signal distribution format should meet considerations of frame length, time resolution, and the allowable ambiguity interval. A minimum number of distributed frequencies chosen to meet generalized needs and not necessarily specific user inputs may be selected on the basis of maximum utility or special considerations such as reducing noise when multiplied to a microwave frequency. By providing a proper interface, the number of distributed signals can be minimized.

Another important consideration in distribution is electromagnetic compatibility (EMC). High-resolution time signals and high frequencies are not only degraded on coaxial lines, but also may radiate to other systems. Noise picked up by leakage and ground loops can also contaminate the distributed signals. Modern light-pipe technology should significantly improve EMC.

For timekeeping, redundancy is valuable not only internally within units, but by employing multiple onboard standards and distribution components. Redundant power to operate these systems is virtually standard, but may employ multiple battery backing with safeguards against common-mode failures.

When using multiple units or components, a number of techniques, including manual switching, are sometimes employed. Combining techniques have sometimes been used to overcome the principal disadvantages of the very reliable, but slow, manual methods. However, the complexity of combining systems, including the monitoring and decision-making portions of the combiner can easily reduce its reliability, in terms of purity, switching gaps or transients, and even complete failure, to less than that of a single, non-redundant component or unit. A loosely slaved clock or oscillator at the user equipment may be considered.

Maintenance of continuous or reliable service to the user systems can be aided by reliable alarms and monitors, definitive fault location provisions, and repair capability for redundant units or components during operation. Without these provisions in the larger installations, there is danger of disturbing a working unit of a redundant set after failure of another unit.

Management of the PDS is an important consideration. The question of updating the standards by time steps, frequency adjustment, or both and within what limits, may not have a single answer on a platform containing diverse user interests. While the method used may be of little concern to many users, it may be of critical concern to a few. Therefore, a clear understanding of the system management practices is needed. Options include (1) complete dependence upon the rate accuracy of onboard "primary" standards such as cesium-beams, with time-signal adjustments made in discrete steps, (2) periodic adjustment of frequency to maintain time, (3) no adjustment made during a critical operation, but offsets noted and relayed to critical users, and (4) other management philosophies. The method may be specified either to make maximum accommodation to a large user segment, while perhaps placing an extra burden on a few users, or to offer multiple service. Updating procedures should also take into consideration the confidence in or effects of noise on a measurement made through the dissemination system and how adjustment decisions would be made.

Performance Factors

Some performance factors of importance to the individual sections of a platform distribution system are enumerated below.

Dissemination System

- Availability & Survivability
- Accuracy
- Position, Velocity, Acceleration
- Format

Onboard Standards

- Frequency Accuracy
- Frequency Stability
 - Long Term
 - Short Term
- Spectral Purity
- Settability or Controllability
- Reliability
- Environment
 - Temperature and Humidity
 - Altitude
 - Acceleration
 - Magnetic Fields
 - Radiation

Monitoring

- Frequency and Time Comparison Accuracy
- Detection of Errors
 - Phase
 - Time
 - Frequency
 - Speed
 - Accuracy
- Signal Level
- Monitor Reliability

Control

- Control Philosophy
 - Frequency vs Time
 - Manual vs Automatic
- Signal Continuity
 - Phase Stability & Jitter
 - Switching Speed
- Reliability

Distribution

- Timing Resolution
- Signal Purity & EMC
- Reliability
- Survivability
- Continuity

The importance of these considerations varies widely among platforms of different types. The Navy is considering standardized platform distribution systems in classes suitable for large ships, small ships, submarines, shore stations, and aircraft. Standardization can promote not only more economical implementation, maintenance and support, but also can serve as a basis for more efficient and less expensive user equipment development and implementation.

Timekeeping

Timekeeping, either in the platform distribution system or in a user system served by a standard frequency, deserves special attention. The time error of the free-running clock after calibration through the dissemination system can be expressed as:

$$T(t) = T_0 + R_0 t + 1/2 D t^2 + \epsilon(t) \quad (1)$$

where,

t = elapsed time since calibration,

T_0 = time error of calibration,

R_0 = rate (frequency) error of clock at time of calibration,

D = fractional frequency drift, and for most practical uses:

$$\epsilon(t) \approx t \times \sigma_y (\tau = t) \quad (2)$$

where $\sigma_y(\tau = t)$ means the fractional frequency stability evaluated for a sampling time τ equal to the time t which elapsed after the last calibration.

The value of D and $\sigma_y(\tau)$ can usually be found in manufacturer's literature. Maximum values of D are normally given in parts per part per day for crystal oscillators and per month for rubidium standards. Cesium beam standards are not generally considered to have a systematic drift, and their long-term stabilities are bounded by accuracies or stabilities defined in terms of reproducibility or accuracy for the life of the standard. Values of $\sigma_y(\tau)$ are given for various sampling times, usually for τ in units of seconds, but sometimes for τ in days.

With T , T_0 and t given in seconds (R_0 and σ_y are dimensionless frequency ratios), D is in parts per part per second. A useful form of (1) is

$$T(t) = T_0 + 8.64 \times 10^4 (R_0 t + 0.5 D t^2) + \epsilon(t), \quad (1a)$$

where T and T_0 are in seconds, t is in days, and D is in parts per part per day. If D is given in terms of months, $1/30$ of that value might be used in (1a).

Obviously, drift can have a profound effect on timekeeping accuracy and must be considered in terms of the mission and survivability of the platform. However, it should also be recognized that the values used in (1) and (1a) are frequently quoted in terms of a rather benign environment and may require adjustment for the other environmental effects on stability, such as acceleration, temperature, and magnetic fields, that are usually quoted separately.

Maintenance of accurate time has become an essential part of some modern systems. Platform distribution systems can provide a valuable service to multiple user systems. However, the factors affecting performance and survivability must be addressed in the platform distribution system specifications. At present, there is relatively little guidance in doing this.

References

1. "Frequency Standards and Clocks: A Tutorial Introduction"; NBS Technical Note 616 (2nd Revision), U. S. Department of Commerce/National Bureau of Standards.

AD P 002525

MODERNIZATION OF THE MILITARY SPECIFICATION FOR QUARTZ CRYSTAL UNITS

Raymond L. Filler

US Army Electronics Technology and Devices Laboratory (ERADCOM)
Fort Monmouth, New Jersey 07703

Abstract

The military specification for quartz crystal units, MIL-C-3098G¹, is, at the time of this publication, woefully out-of-date. Several of the specified performance parameters and measurement techniques are inadequate for modern applications. In addition, many required inspections are not addressed, such as motional capacitance, activity dips and radiation hardness.

The philosophy behind the revision of the specification will be discussed and several of the major modifications will be described. Some of the modifications to be discussed are the elimination of the CI-meter as a standard test set and a description of the new or modified inspections for acceleration sensitivity, acoustical noise sensitivity, frequency-temperature behavior, thermal shock, aging and short-term stability.

Key Words: Quartz Resonators, Quartz Crystal, Military Specification.

Introduction

The Military Specification which covers the general requirements for quartz crystal units used for frequency control of electronic equipment is entitled, Crystal Units, Quartz, General Specification for. The current edition, MIL-C-3098G, was issued 12 October 1979. The intent of this paper is to discuss the rationale behind the proposed modification of "3098" as well as to describe several of these modifications.

Rationale

It has become increasingly apparent that MIL-C-3098G is inadequate for current and future military requirements. Modern systems require better performance over time, temperature, and other environmental stimuli than that for which "3098" was written. There are also many performance factors which are unaddressed, such as vibration and acceleration sensitivity while operating, radiation hardness, frequency-temperature and resistance-temperature anomalies ("activity-dips"), motional capacitance (C_1), short term stability (noise), thermal shock response and thermal-frequency repeatability. A primary motivation for revising this specification was to define test procedures for these performance factors and institute them as standard methods.

Two other areas of concern have been the inadequate level of quality assurance demanded

previously and the obsolete measurement equipment specified. The specification was originally written prior to the era of hybrid circuits, and encapsulated ovenized and temperature compensated oscillators. Complete parameter and frequency temperature data is needed on all crystal units in some applications. In addition, certain critical applications, such as space-borne equipment, require higher levels of quality assurance than that required for general use. Provisions, therefore, had to be included into "3098" to require higher levels of data reporting and tighter inspections, if required.

The "government reference standard test sets" specified by MIL-C-3098G were the now obsolete CI-meters. These equipments are no longer maintained by the government. They are also unsuitable for the precision measurements required both today and in the future. It was imperative, therefore, that the provisions included for measurement equipment reflect both the requirements of modern systems and the state-of-the-art of modern measurement technology.

Major Changes

The following is a summary of the major changes being considered. All of the items discussed are subject to change prior to final publication of the "new" 3098.

Etching

The requirement paragraph for etching of the crystal blank and the assurance paragraph for etching evaluation have been eliminated. These paragraphs were originally inserted to give assurance to the government that surface contamination and particulates were minimized. The idea is now to test directly for the performance criteria that etching is intended to improve. Therefore, the low drive level and aging inspections, to be discussed below, have been made more stringent to serve as functional inspection methods.

CI-Meter

In light of the facts that 1) the government does not support the CI-meter any longer, 2) the precision of the CI-meter is inadequate for many modern applications and 3) there are several commercial instruments capable of crystal unit measurement; all reference to the CI-meter has been deleted. Each manufacturer will be required to qualify a reference standard test set. The details of this qualification have yet to be established.

Quality Assurance Levels

The need for higher levels of quality assurance has been addressed by the introduction of two "product assurance levels". The product assurance level is denoted by a one letter, (S or B), suffix to the type designation. The difference between the two levels lies in the sampling plan used for the inspections. In the following table the Group A tests are typically for performance factors which vary from unit to unit while Group B inspections are typically design or lot dependent. In Table I, we have the sampling plan for the two product assurance levels and the two inspection groups.

TABLE I

Inspection Group	Product Assurance Level	
	B	S
Group A	II*	100%
Group B Subgroup I	S-3*	II*
Subgroup II	S-4*	III*

*in accordance with MIL-STD-105.³

For comparison purposes, Table IIA and Table IIB give the number of samples to be inspected for two lot sizes (above the slash) and the number of failures permitted for lot acceptance (below the slash). It should be noted that S level units are accepted on an individual basis for Group A testing.

TABLE IIA

Inspection Group	Sample Size/Failures Allowed	
	B	S
Group A	13/0	100%
Group B Subgroup I	13/0	13/0
Subgroup II	13/0	50/1

TABLE IIB

Inspection Group	Sample Size/Failures Allowed	
	B	S
Group A	80/2	100%
Group B Subgroup I	13/0	80/2
Subgroup II	13/0	125/3

Modifications to Existing Inspections

Several of the inspection procedures familiar to the industry from previous editions of "3098" have been modified. These modifications are intended to make it possible to specify a wider variety of performance parameters and to a higher degree of precision. The three major modifications are the Reduced Drive Level Test, the Aging Test, and the Frequency vs. Temperature test.

Resistance Non-Linearity

A more stringent low drive level test has been proposed to provide functional quality assurance in lieu of the former etching requirement. Resistance non-linearity has been reported to be a good indicator of surface imperfections (including scratches and particles on the surface.) The new reduced drive level inspection will be a measurement of the equivalent series resistance of the crystal unit as a function of drive current. It should be noted that throughout the revised specification all references to drive level will be in terms of current not power.

The experimental procedure for the resistance non-linearity inspection will be to condition the crystal unit at an elevated temperature in a quiescent state. After the high temperature conditioning, the crystal unit is brought to thermal equilibrium at room temperature and a resistance vs drive current measurement is made, using the lowest current level first. An ideal crystal unit will obey Ohms Law, i.e., the resistance will not be a function of current. A typical crystal unit will have a resistance which is more-or-less a function of the drive current. The specification sheet for the crystal type will specify the fractional resistance deviation which is permissible.

Aging

The accelerated aging test, which is 1 week at an elevated temperature, is not intended to be a predictor of long term performance. It is another check on process control. This test has been modified by raising the required conditioning temperature from 105°C to 125°C. It has also been suggested that the accelerated aging test be the thermal conditioning for the reduced drive level test.

The "regular" aging test has also been modified. Frequency measurements are required 5 times a week (every working day). Some typical aging data is shown in Figure 1. The data will be appropriately analyzed at the end of the 30 day aging period by performing a statistical fit to a suitable function to ascertain the level of conformance to the requirements. The requirements can be any or all of the following: the aging rate per day at day 30 (determined from the statistical fit), the total frequency change from day zero to day 30, and the maximum frequency shift in any 24 hour period.

The maximum frequency change for the 30 day period must be less than 2 parts per million for all crystal units. The temperature at which aging is to be performed will be 85°C ±2°C for crystal units with a 2 ppm/30 day aging specification. All other crystal units will be aged at a temperature to be specified, but it will be a temperature with a low frequency vs temperature slope, such as a turnover point.

An open question is the relevance of the fit of the 30 day data to the extrapolation of the aging to longer periods. For the present, the user will have to use the data supplied and some assumptions gathered from previous experience to predict aging behavior.

Frequency and Resistance vs Temperature

The inspection for temperature behavior is being greatly expanded. As in previous editions of the specification, the dependence of frequency and resistance on temperature will be considered. As contrasted to previous editions, the definitions related to frequency-temperature (F-T) characteristics depicted in figure 2 will be introduced. It will be possible to specify any one or combination of:

Frequency at upper turnover temp F(UTP)
Frequency at lower turnover temp F(LTP)
Lower turnover temperature LTP
Upper turnover temperature UTP
Frequency at the inflection temp F(T_i)
Slope at any temperature dF/dT
Maximum slope over temperature range
Slope at inflection temperature
F(LTP) - F(UTP)

or any other characteristic which may be important. The resistance will be limited by a specified maximum over a specified temperature range.

The temperature run will be performed much more slowly than in the past, to assure the elimination of thermal gradients in the crystal unit and to allow inspection for frequency-temperature anomalies, usually referred to as activity dips.

Additional Special Requirements

The inspections heretofore absent from "3098" will only be applied when specifically required by the specification sheet for a crystal unit type. These new inspections include:

Motional capacitance, C₁
Quality factor, Q
Short-term-stability (noise)
Vibration sensitivity
Acceleration sensitivity
Acoustic noise sensitivity
Radiation hardness
Condensibles
Thermal frequency repeatability
Thermal hysteresis
Thermal time constant
Thermal frequency overshoot

Some of these parameters are well known to the industry. The discussion here will center around the less common of these new inspections.

Thermal Frequency Repeatability

The definition of thermal frequency repeatability is the maximum fractional frequency deviation among the frequencies measured at each pass through the specified turnover temperature during a specified number of temperature cycles. A typical temperature cycle to be followed is shown in figure 3. The steps near the turnover temperature signify thermal equilibrium being reached at a number of temperatures to enable a determination of the turnover frequency by statistical means.

This test is mostly applicable to crystal units intended for ovenized application. It is intended to give a measure of the frequency

accuracy to expect after the low temperature storage of an oscillator with the oven unpowered. The corresponding inspection for non-ovenized oscillators is the thermal hysteresis inspection. This inspection compares the frequency retrace over the entire operating temperature range.

Vibration/Acceleration Sensitivity

An extremely important inspection for certain military applications, which has drawn a considerable amount of R&D interest lately, is the performance of a crystal unit while experiencing acceleration.⁴ This acceleration can be unidirectional, vibrational, or due to acoustic noise power impinging on the crystal unit. The current revision of "3098" contains only a before and after vibration measurement to determine the frequency and resistance offset.

The new "3098" will include vibration sensitivity to measure the frequency change during vibration and to search for mechanical resonances in the crystal unit. In addition "3098" will include acceleration sensitivity and acceleration induced frequency and resistance offset. This will measure the effects of a steady state acceleration such as would be experienced during missile launch. The acoustical noise inspection will measure the effects of sound energy on the crystal unit enclosure.

The vibration sensitivity will be measured using the sideband method which has been reported before.⁴ The acceleration modulates the frequency of the carrier at a rate determined by the driving signal to the shake table and at a magnitude depending on the crystal properties. The spectrum looks like that of figure 4. The relative magnitude of the carrier and the sideband along with the frequency of the carrier and the frequency of the driving signal are sufficient to determine the vibration sensitivity.

Time Constant/Overshoot

Another new inspection is the Thermal Time Constant/Thermal Frequency Overshoot Test. In this inspection, the ability of a crystal unit to follow a rapid temperature change is measured. The AT cut of quartz is very sensitive to thermal transients, while the SC-cut is relatively insensitive. In addition massive crystal blanks with poor thermal conductance in the mounting structure will have a long thermal response time.

The thermal time constant is measured by bringing the crystal unit to thermal equilibrium at 0°C while operating in a suitable test system. The unit is rapidly transferred to a 100°C liquid bath and the frequency is recorded until equilibrium is again reached. The time taken to reach the original frequency while at thermal equilibrium at 63°C is the thermal time constant.

The frequency overshoot is a measure of the deviation of the dynamic frequency-temperature characteristic (i.e., the F-T curve experienced during rapid temperature change) and the F-T curve obtained under quasi-thermal equilibrium conditions. The measurement method is depicted in figure 5. The upper curve is the F-T plot obtained "statically". The lower curve is the frequency vs time plot obtained during the thermal time constant test. The extremes of curve is denoted f_c. The

frequency overshoot is defined as

$$\text{Freq overshoot} = \Delta F_d - \Delta F_s \quad (1)$$

where $\Delta F_d = f_t \text{ (dynamic)} - f(100^\circ\text{C}) \text{ dynamic}$

$\Delta F_s = f_t \text{ (static)} - f(100^\circ\text{C}) \text{ static}$

and $f_t = \text{turnover frequency}$

$f(100^\circ\text{C}) = \text{frequency at } 100^\circ\text{C}$

Future Revisions

The current plan is to complete this revision and then tackle:

MIL-STD-683 - Crystal Unit, Quartz and Crystal Holders, Enclosures, Selection of and

MIL-H-10056 - Holders, Crystal, General Specification for.

Acknowledgements

The author wishes to thank Dr. John Vig, Mr. Vincent Rosati, Ms. Florence Briller and Mr. W. Quinones of ERADCOM and all participants in the joint government industry conferences for their many helpful discussions.

References

1. Military Specification: "Crystal Units, Quartz, General Specification for", MIL-C-3086, (1979), available from Naval Publications and Forms Center, 5801 Tabor Ave, Philadelphia, PA 19120
2. Symthe, R.C. and Horton, W.H., "Evaluation of Crystal Measurement Systems: These Proceedings (1983)
3. Military Standard: "Sampling Procedures and Tables for Inspection by Attributes", MIL-STD-1050, (1979), available from Naval Publications and Forms Center, 5801 Tabor Ave, Philadelphia, PA 19120
4. Filler, R.L., "The Effect of Vibration on Frequency Standards and Clocks", Proc 35th Annual Frequency Control Symposium, pp. 31-39. (1981) copies available from Electronic Industries Association, 2001 Eye Street, NW, Washington, D.C. 20006.
5. Filler, R.L., and Vig, J.R., "The Acceleration and Warmup Characteristics of Four-Point Mounted SC and AT-cut Resonators", Proc 35th Annual Frequency Control Symposium, pp. 110-116. (1981) copies available Electronic Industries Association, 2001 Eye Street, NW, Washington, D.C. 20006.

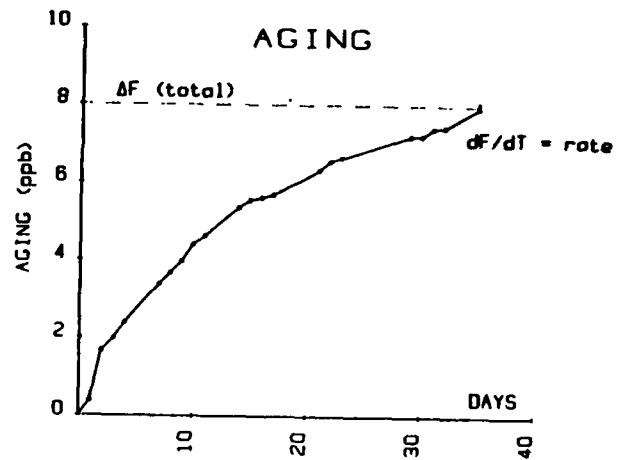
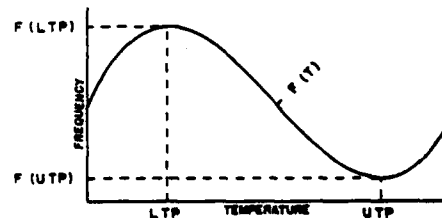


FIGURE 1: TYPICAL AGING DATA



slope: $S(T) = \frac{dF(T)}{dT}$

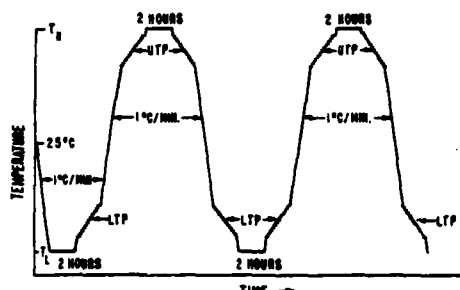
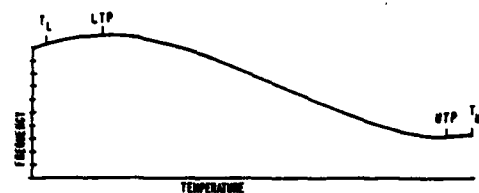
curvature: $K(T) = \frac{d^2F(T)}{dT^2}$

turnover point: $\frac{dF(T)}{dT} = 0$

inflection point: $\frac{d^2F(T)}{dT^2} = 0$

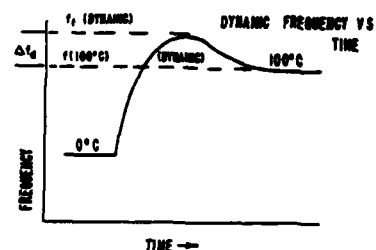
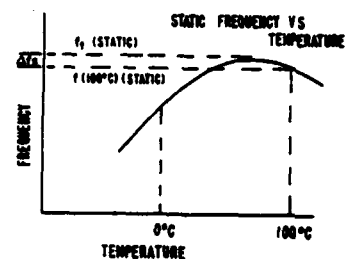
FREQUENCY VS TEMPERATURE DEFINITIONS

FIGURE 2



THERMAL FREQUENCY REPEATABILITY
(TWO-TURNOVER THERMAL CYCLE)

Figure 3



FREQUENCY OVERTHOOT
(LOWER TURNOVER TEMPERATURE BETWEEN 0°C AND 100°C; UPPER TURNOVER TEMPERATURE ABOVE 100°C.)

Figure 5

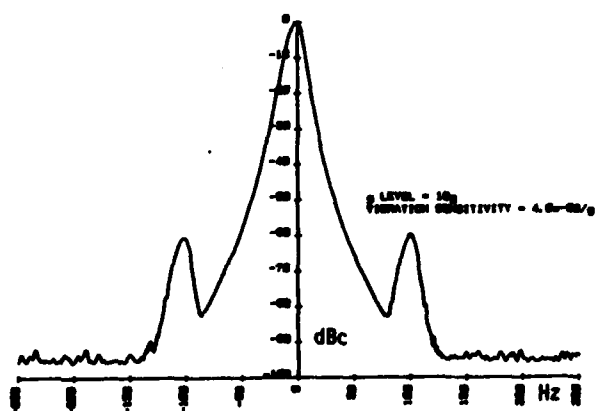


Figure 4: Typical Frequency Spectrum during Vibration

REVISION OF THE MILITARY SPECIFICATION FOR
QUARTZ CRYSTAL OSCILLATORS (MIL-0-55310)

Vincent Rosati and Stanley Schodowski

US Army Electronics Technology & Devices Laboratory (ERADCOM)
Fort Monmouth, New Jersey 07703

The current revision was started in mid-1982 with a request for comments sent to many oscillator users and vendors. The result was a flood of proposed changes, indicating that extensive revision is necessary. Two working group meetings, to which more than 260 representatives of the Military and industry were invited, were held at Fort Monmouth and in Philadelphia. Approximately thirty people attended each session. The group consisted of representatives from 13 companies and 3 Government agencies.

The initial goals of the working group were:

1. Modernize "55310", thereby making it more useful for precision oscillators.
2. Standardize definitions, requirements, and test procedures.
3. Conform, insofar as possible, to the IEC standard for crystal oscillators.

In addition, an important goal for the Government is that by upgrading the standard and making it more useful for precision oscillators, the number of non-standard oscillators could be reduced.

Still another goal emerged during the meetings. A user's guide to the selection, use, and specification of oscillators will be developed. It appears that many of the problems associated with oscillators that fail to meet desired performance stem from a lack of understanding on the part of the user as to what various classes of oscillators can do, how to specify them, and how to test them. A readable user's guide should remove these barriers.

Much time was spent, at both meetings, on the topics of the measurement of frequency-temperature stability and aging.

Regarding F-T stability, a consensus developed that temperature runs begin with a rapid cooldown to the lowest specified temperature, dwell at the low temperature for approximately 5 (oscillator) thermal time constants and then step the temperature at a rate commensurate with the stability class of the oscillator. For example, a 0.5 ppm-oscillator should be measured at one degree intervals. When the upper specified temperature is reached, the cycle is reversed.

Testing in this way will help uncover frequency anomalies due to condensibles, have a reasonable chance of exposing activity dips, and give a measure of the thermal hysteresis of the oscillator.

The measurement of frequency aging of TCXO likewise presents many problems. At what temperature should measurements be made? How often should data be taken? For how many days should one measure aging? Is it valid to fit a mathematical function to the data and use it to extrapolate future performance?

On a strictly scientific basis these questions cannot be answered in general. However, there is a practical need to develop methods which can prudently be applied.

It is realized and accepted that more extensive testing will increase the cost of initial procurement. The authors feel that, with the availability of automated test equipment, cost increases will be modest and that, in any case, life cycle costs of systems will significantly be reduced.

It is envisioned that as more is learned about the performance, testing and applications of oscillators, MIL-0-55310 will be revised periodically. The authors sincerely request that interested readers communicate their comments and suggestions to us at the address given under the title of this paper.

REVIEW OF NEW MILITARY SPECIFICATION ON SURFACE ACOUSTIC WAVE DEVICES

Elio A. Mariani

US Army Electronics Technology & Devices Laboratory
Fort Monmouth, New Jersey 07703

Summary:

A new military specification on surface acoustic wave devices (SAWD) has been drafted recently by Army personnel and is presently under review. This document will establish the general requirements for SAWD's as well as establish the quality and reliability assurance requirements which must be met in the acquisition of these devices. Requirements that are sensitive to particular device application shall be specified in other applicable detailed specifications. The contents and salient provisions of this document will be reviewed.

Introduction:

As a result of more than a decade of intensive development, SAW device technology has achieved high levels of device performance and system acceptance. At the present time, SAW devices have found application in a large number of military systems and that number continues to grow. Recognizing the relative maturity of SAW device technology, the U.S. Army Electronics Technology and Devices Laboratory, USAERADCOM, decided to initiate the preparation of a new military general specification dealing exclusively with SAW devices. In order to expedite the process the document was prepared as a limited coordinated specification, for mandatory Army use only; upon approval by Department of the Army, the specification will be coordinated with both the Navy and the Air Force.

After considerable time and effort, a finished draft document soon will be circulated to a number of interested people within the Army and private industry. This process will permit comment and any needed modifications in order to produce a consensus document. This will be ready in final form by the fall of 1983.

Discussion of Specification

The proposed general specification on SAW devices is organized into six major sections: (1) Scope; (2) Applicable Documents; (3) Requirements; (4) Quality Assurance Provisions; (5) Packaging and (6) Notes. Only those parts of special interest, that is the Requirements and Quality Assurance Provisions, will be addressed. The remaining sections are rather prosaic and will be available with the distribution of the complete draft of the general specification. Furthermore, the specification provides for two levels of quality and reliability assurance - Class B and Class S (space application).

The Requirements Section, which has eighteen (18) major subsections, begins with a discussion of detail specification sheets where the individual item requirements shall be as specified in the general specification and in accordance with the applicable detail specification. In the event of any conflict, the detail specification will govern. Also, unless otherwise specified, all SAW devices shall have an operating temperature range from -55°C to +125°C.

The Requirements section dealing with qualification indicates that SAWD's finished under this specification shall be the products which are qualified for listing on the applicable qualified products list at the time set for opening of bids. This section on qualification also discusses traceability documentation and change of qualified product. The remaining sections discuss the following topics: screening, materials, design and construction, hermetic seal, thermal shock, high temperature storage, shock, vibration, moisture resistance, life, resistance to solvents, solderability, short-circuit and open-circuit, electrical requirements, marking and workmanship. It is noted that under the section on electrical requirements, testing shall assure that the SAWD's tested meet the electrical requirements of the applicable detail specification and, as a minimum, shall include the measurement of the following parameters: center frequency of operation, bandwidth, insertion loss, feedthrough suppression, spurious echo suppression, and VSWR, when applicable.

The Quality Assurance Provisions section is divided into seven major subsections, the first of which addresses responsibility for inspection. It states that unless otherwise specified, the contractor is responsible for the performance of all inspections requirements.

Classification of inspection indicates three major categories: (1) material inspection; (2) qualification inspection; (3) quality conformance inspection. Also, all inspection conditions will be made in accordance with the general requirements of MIL-STD-883.

The subsection dealing with test routine states that qualification samples shall be subjected to the tests described under Qualification Inspection (Table I) which contains six (6) groups. All sample units shall be subjected to tests of group I. The sample SAWD's shall then be divided in five (5)

groups and subjected to the tests of groups II-VI in the order shown for the specific class of SAWD being qualified (i.e. Class B or S).

Quality conformance inspection will include inspection of product for delivery that consists of group A, B, and C inspections, as shown in Tables II, III, and IV.

The last subsection under the Quality Assurance Provisions deals with methods of inspection which correspond directly to each of the requirements paragraphs.

References:

1. Military Specification: General Specification for Microcircuits (Mil-M-38510D), 31 August 1977)

Acknowledgement:

The author gratefully acknowledges the advice and support of Mrs. F. Briller in the performance of this work.

TABLE II

GROUP A INSPECTION

EXAMINATION OR TEST	CLASS B*	CLASS S*
SUBGROUP 1 - WAFER		
MARKING	X	X
ADHESION OF METAL FILM	X	X
SHORT CIRCUIT	X	X
SUBGROUP 2 - SAWD'S		
VISUAL	X	X
MARKING	X	X
HERMATIC SEAL	X	X
INTERNAL WIRE BONDING	X	X

* CLASS B & CLASS S WILL UTILIZE SAME TESTS BUT DIFFERENT SAMPLING PLAN.

TABLE III

GROUP B INSPECTION

EXAMINATION OR TEST

ELECTRICAL CHARACTERISTICS:

- 0 CENTER FREQUENCY OF OPERATION
- 0 BANDWIDTH
- 0 INSERTION LOSS
- 0 FEEDTHROUGH SUPPRESSION
- 0 SPURIOUS ECHO SUPPRESSION
- 0 VSWR (WHEN APPLICABLE)

TABLE I

QUALIFICATION INSPECTION

- GROUP I
- 0 VISUAL CHECK
 - 0 MARKING
 - 0 DIMENSIONS
 - 0 INTERNAL WIRE BONDING
 - 0 ELECTRICAL CHARACTERISTICS
- GROUP II
- 0 SOLDERABILITY (WHEN APPLICABLE)
- GROUP III
- 0 HIGH TEMPERATURE STORAGE
 - 0 ELECTRICAL CHARACTERISTICS
- GROUP IV
- 0 LIFE
 - 0 SHORT CIRCUIT TEST
 - 0 ELECTRICAL CHARACTERISTICS (FINAL)

- GROUP V
- 0 HERMATIC SEAL
 - 0 SHORT CIRCUIT TEST
- GROUP VI
- 0 VIBRATION
 - 0 SHORT CIRCUIT TEST
 - 0 SHOCK
 - 0 SHORT CIRCUIT TEST
 - 0 THERMAL SHOCK (10 CYCLES)
 - 0 SHORT CIRCUIT TEST
 - 0 MOISTURE RESISTANCE
 - 0 SHORT CIRCUIT TEST
 - 0 ELECTRICAL CHARACTERISTICS (FINAL)

TABLE IV

GROUP C INSPECTION

EXAMINATION OR TEST

- SUBGROUP 1
- 0 HIGH TEMPERATURE STORAGE
 - 0 HERMATIC SEAL
- SUBGROUP 2
- 0 SOLDERABILITY (WHEN APPLICABLE)
 - 0 LIFE
 - 0 ELECTRICAL CHARACTERISTICS (FINAL)
- SUBGROUP 3
- 0 VIBRATION
 - 0 SHORT CIRCUIT
 - 0 SHOCK
 - 0 SHORT CIRCUIT
 - 0 THERMAL SHOCK (10 CYCLES)
 - 0 SHORT CIRCUIT
 - 0 MOISTURE RESISTANCE
 - 0 ELECTRICAL CHARACTERISTICS (FINAL)

AUTHOR INDEX

AUTHOR	PAGE	AUTHOR	PAGE	AUTHOR	PAGE
Adachi, T.	187	Itano, W.M.	37	Ramsey, N.F.	5
Allan, D.W.	55	Jespersen, J.L.	506	Raue, J.E.	361
Amstutz, P.	239	Katz, S.	181, 185	Reinhardt, V.	61, 97
Armington, A.F.	177	Kahan, A.	169	Roloff, R.	61
Ashrafi, D.	317	Kamas, G.	506	Ronen, M.	181
Ashrafi, S.	317	Kasai, Y.	376	Rosati, V.	501, 524
Aubry, J.P.	306	Kawatsu, T.	376	Rossman, H.	272
Balascio, J.F.	157	Key, P.L.	376	Rousseau, K.V.	428
Ballato, A.	194, 415	Kindell, R.	153	Salvo, H.	87
Beaver, W.D.	226	Kirchner, D.	394	Sawyer, B.	151
Bennett, C.	405	Kirk, A.	67	Schaumann, R.	371
Bollinger, J.J.	37	Kolb, E.D.	42	Schieber, M.	185
Bossoli, R.B.	164	Koehler, D.R.	153	Schluter, W.	61
Boulanger, J.-S.	78	Kong, A.M.	130	Schodowski, S.	501, 524
Bourke, B.	485	Konno, M.	428	Schwartzel, J.	239
Brunet, M.	454	Kosinski, J.A.	376	Sheats, L.	362
Buisson, J.	61	Kruger, R.	265	Shimoda, Y.	434
Carru, H.	239	Kudo, T.	61	Simpson, E.E.	153
Carter, R.L.	477	Kwon, T.M.	448	Sinha, B.K.	415
Chambers, J.L.	275	Lakin, K.M.	12, 18	Smythe, R.C.	290, 297, 349
Chao, H.L.	116	Larkin, J.J.	144, 320	Stahl, A.	454
Chuang, S.S.	248	Lamboley, J.	177	Stalder, T.	61
Cormier, J.E.	177	Lander, V.	111	Stebbins, S.	61
Costain, C.C.	78	Landin, A.R.	473	Stevens, D.S.	208, 325
Costlow, R.J.	97	Lau, K.F.	144	Stover, D.W.	27
Cutler, L.S.	32	Laudise, R.A.	428	Subramaniam, B.	164
Daams, H.	78	Lechopier, S.	153	Sullivan, V.	394
Dachel, P.	61	Lee, P.C.Y.	306	Takeuchi, T.	376
Darces, J.F.	111	Lias, N.C.	200	Tamura, H.	481
Datta, S.	423	Lipson, H.G.	157	Tamura, S.	481
Detaint, J.	239	Lister, M.	169	Tanski, W.J.	415
Detoma, E.	61	Lukaszek, T.	61	Taylor, J.H.	6
Douglas, R.J.	78	Mariani, E.	415	Tiersten, H.F.	208, 325
Dubois, B.	61	Martin, J.J.	525	Toki, M.	300
Dworsky, L.	232, 387	Martin, S.J.	130, 164	Tomikawa, Y.	376
EerNisse, E.P.	255	Mattison, E.M.	423	Toulouse, J.	125
Emin, C.D.J.	136	Mc Avoy, B.R.	49	Tsuzuki, Y.	187, 300
Facklam, R.	21	Mc Clelland, T.	87	Uberall, H.	317
Filler, R.	265, 501, 519	Mc Guire, M.D.	12	Uno, T.	434
Freon, G.	61	Melloch, M.R.	32	Uzunoglu, V.	91
Fujii, S.	448	Mengel, E.E.	423	Vessot, R.F.C.	49
Fujiwara, Y.	343	Merigoux, H.	27	Vig, J.R.	265
Gagnepain, J.J.	218	Miller, T.J.	111	Wakatsuki, N.	343
Gerard, E.	306	Mitsuoka, T.	423	Wakino, K.	481
Giffard, R.P.	32	Mizan, M.	300	Walls, F.L.	218
Gounji, T.	376	Moore, R.A.	194	Wang, H.T.M.	7
Graf, E.P.	492	Murray, J.A.	87	Wang, J.S.	144
Granveaud, M.	61	Nanamatsu, S.	516	Wardrip, C.	61
Green, E.R.	125	Nehring, D.	448	Watanabe, S.	337
Guinot, B.	61	Nishikawa, T.	459	Werner, J.F.	136
Gunshor, R.L.	423	Nottarp, K.	481	West, J.D.	164
Halliburton, L.E.	164	Nowick, A.S.	61	White, J.	513
Halperin, A.	181, 185	Oaks, O.	125	Wilkus, S.	354
Hanson, W.P.	261	O'Connor, J.	61	Williams, H.E.	12
Hahn, H.	284	Okazaki, M.	169, 177	Wilson, J.S.	442
Hahn, T.	18	Okzaki, M.	187	Wineland, D.J.	37
Hartmann, C.	354	Olivier, M.	337	Withers, R.S.	81
Haynes, J.T.	87, 272	O'Shea, T.	218	Wright, P.V.	81
Healey III, D.J.	284	Owens, J.M.	394	Yamada, S.	343
Horrigan, J.A.	177	Parekh, J.P.	477	Yen, K.H.	428
Horton, W.H.	290, 297	Parker, T.E.	473	Yong, Y.K.	200
Howard, M.D.	349	Paros, J.M.	116, 410	Yoo, K.B.	317
Imbier, E.	49	Peier, U.R.	255	Zecchini, P.	111
Ingold, J.	61	Penrod, B.	492		
		Pierret, R.F.	485		
		Powell, S.	423		
			284		

37th Annual Frequency Control Symposium - 1983
SPECIFICATIONS AND STANDARDS GERMANE TO FREQUENCY CONTROL

Institute of Electrical and Electronic Engineers

Order through: IEEE Service Center
445 Hoes Lane
Piscataway, NJ 08854
(201)-981-0060

- 176-1978 Piezoelectricity \$ 9.00
- 177-1966 Piezoelectric Vibrators, Definitions and Methods of Measurements for (ANSI C83.17-1970) \$ 4.00
- 180-1962 Ferroelectric Crystal Terms, Definitions of \$ 3.00
- 319-1971 Piezomagnetic Nomenclature \$ 4.00

Electronic Industries Association

Order through: Electronic Industries Assn.
2001 Eye Street, N.W.
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 older types.) \$ 6.80
- RS-367, Dimensional and Electrical Characteristics Defining Receiver Type Sockets. (Including crystal sockets.) \$20.20
- RS-417, Crystal Outlines (Standard dimensions and pin connections for current quartz crystal units - 1974.) \$ 7.80

(b) Production Tests

- RS-186-E, (All Sections), Standard Test Methods for Electronic Component Parts \$42.00

(c) Application Information

- Components Bulletin No. 6, Guide for the Use of Quartz Crystals for Frequency Control \$ 4.90

(d) RS-477, Cultured Quartz (Apr. 81) \$ 5.50

International Electrotechnical Commission

Order through: American National Standards Inst.
1430 Broadway
New York, New York 10018

*ANSI can quote prices on specific IEC publication on a day to day basis only. All IEC and ISO standards have been removed from its Standards Catalog. Call ANSI, NYC (212)-354-3300 for prices.

IEC Publication 122-1 (1976)

Quartz crystal units for frequency crystal and selection. Part 1: Standard values and test conditions. (Second edition)

IEC Publication 122-2 (1962) Section 3:
Guide to the use of Quartz Oscillator Crystals, including Amendment 1 (1969).

IEC Publication 122-3 (1977) Part 3: Standard
Outlines and Pin Connections. (Second edition)

IEC Publication 283 (1968) Methods for the Measure-
ment of Frequency and Equivalent Resistance of
Unwanted Resonances of Filter Crystal Units

IEC Publication 302 (1969) Standard Definitions and
Methods of Measurement for Piezoelectric Vibrators
Operating Over the Frequency Range up to 30 MHz

IEC Publication 314 (1970) Temperature Control
Devices for Quartz Crystal Units, including
Supplement 314A Contents: General Characteris-
tics & Standards; Test Conditions; Pin Connections

IEC Publication 314A (1971) First Supplement to
Publication 314 (1970) Contents: Guide to the Use
of Temperature Control Devices for Quartz Crystal
Units

IEC Publication 368 (1971) Piezoelectric Filters
Including Amendment 1, Amendment 2, and Supple-
ment 368A and 368B Contents: General Information
& Standards Values; Test Conditions

IEC Publication 368A (1973) First Supplement to
Publication 368 (1971) Contents: Guide to the
Use of Piezoelectric Filters

IEC Publication 368B (1975) Second Supplement to
Publication 368 (1971) Contents: Piezoelectric
Ceramic Filters

IEC Publication 444 (1973) Basic Method for the
Measurement of Resonance Frequency and Equivalent
Series Resistance Quartz Crystal Units by Zero
Phase Technique in a π - Network

IEC Publication 483 (1976) Guide to Dynamic Measure-
ments of Piezoelectric Ceramics with High Elec-
tromechanical Coupling

Department of Defense

Order through: Naval Publication & Form Center
5801 Tabor Avenue
Philadelphia, PA 19120

MIL-C-3098 Crystal Unit, Quartz. General Specifi-
cation For

MIL-H-10056 Holders (Enclosures), Crystal, General
Specifications For

MIL-STD-683 Crystal Units, Quartz/And Holders,
Crystal

MIL-F-28734 Frequency Standards, Cesium Beam,
General Specification For

MIL-O-55310 Oscillators, Crystal, General Specifi-
cation For

MIL-F-18327 Filters, High Pass, Low Pass, Band
Pass Suppression and Dual Functioning, General
Specification For

MIL-O-39021 Oven, Crystal, General Specification For

MIL-O-55240 Oscillators, Audio Frequency

MIL-F-28811 Frequency Standard, Cesium Beam Tube

PROCEEDINGS
ANNUAL FREQUENCY CONTROL SYMPOSIA

<u>NO.</u>	<u>YEAR</u>	<u>DOCUMENT NO.</u>	<u>OBTAIN FROM*</u>	<u>COST</u>
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	AD746210	"	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		SGC	25.00
37	1983		"	25.00

*NTIS-National Technical Information Service
Sillis Building
5285 Port Royal Road
Springfield, VA 22161

(All prices are as of June 1983 and are subject to change.)

* EIA-Annual Frequency Control Symposium
c/o Electronic Industries Association
2001 Eye Street
Washington, DC 20006

* SGC-Systematics General Corporation.
Brinley Plaza Route 38
Wall Township, NJ 07719

Remittance must be enclosed with all orders. For orders placed with NTIS and SGC from outside the United States, double the domestic price list per copy for handling and mailing.

A complete index to the Proceedings of the Frequency Control Symposia for 1956 (10th) through 1976 (30th) is available as part of the 1976 (30th) Proceedings.

The Proceedings of the 33rd (1979) and 34th (1980) Symposia contain a bibliography of the world-wide literature on precision frequency control and selection compiled by Dr. E. A. Gerber for the years 1968-1978 with part of 1980, respectively.