Annual Research Journal

Electrical Engineering
Research Experience for Undergraduates

Vol. II

August 2004

National Science Foundation
Grant No. EEC-0244030

Principal Investigators:
Ruyan Guo and Kenneth Jenkins

PENNSTATE

Department of Electrical Engineering
University Park, Pennsylvania

EEREU Annual Research Journal

Ruyan Guo (editor)
W. Kenneth Jenkins (co-editor)
Bo Wang (associate-editor)

Published in 2004 by
The Department of Electrical Engineering
The Pennsylvania State University
University Park, Pennsylvania, 16802

NSF EE REU Site Program Contact

316 Electrical Engineering East
The Pennsylvania State University
University Park, PA 16802
Telephone (814) 865-0184
Fax (814) 865-7065
E-Mail: eereu@engr.psu.edu
Web Site: http://www.ee.psu.edu/reu/

ISBN 0-913260-04-5

PREFACE

Summer 2004 marks the second year that the Department of Electrical Engineering (EE), Penn State University, hosted activities of the Research Experience for Undergraduates (REU) Site Program, sponsored by the National Science Foundation (NSF). This summer we had distinctive pleasure working with seventeen remarkable young men and women participating our EEREU program at the University Park Campus, Penn State. These REU students, mostly college sophomores or juniors, have outstanding academic preparations and were selected from nation-wide applicants interested in exploring research in electrical engineering related areas.

During the nine-weeks summer program, each REU student was engaged in a research project, working under the guidance of his or her faculty mentor(s). Seventeen research projects were carried out, individual findings were presented in our annual *EEREU Symposium*, and the technical reports, written by EEREU students as the primary authors, are compiled in this *Annual Research Journal*.

We are happy to present this year's *Annual Research Journal – Electrical Engineering Research Experience for Undergraduates, Vol.* ***II***, to interested readers. Nearly all the reports document original research and are of value for scientific dissemination and publication. We believe it showcases the achievement of our EEREU students and demonstrates the effective mentorship provided by the faculty and graduate student mentor teams.

Besides research activities, our EEREU program organized *Weekly Scientific Seminar Series* to have leading researcher/scientists at Penn State introducing the broad research frontier to our REU students. Students went on several *field trips* to visit prominent local, national and international industrial research sites. Students were also involved in debate and analysis on real ethical problems in engineering through *Weekly Workshop on Ethics* in Science and Research.

For more information about our Research Experience for Undergraduates Program in Electrical Engineering, please visit our website at: http://www.ee.psu.edu/reu/.

We are confident that this journal can serve a step-stone for our EEREU students at the beginning stage of their long research career. We are also hopeful that this journal may stimulate more college students to consider research career and pursue graduate studies in electrical engineering.

Ruyan Guo and W. Kenneth Jenkins
Co-Directors of the NSF EE REU Site Program
Dept. EE, The Pennsylvania State University

July 2004
University Park, Pennsylvania

TABLE OF CONTENTS

Research Articles (* indicates REU student author)

2004 NSF EEREU FACULTY AND STAFF MEMBERS

Faculty Mentors

Prof. Amar Bhalla
Prof. Sven Bilén
Prof. Charles Croskey
Prof. Craig Grimes
Prof. Ruyan Guo, Co-Director
Prof. Heath Hofmann
Prof. Ken Jenkins, Co-Director
Prof. I. C. Khoo
Prof. Mike Lanagan
Prof. Zhiwen Liu
Prof. Theresa Mayer
Prof. John Mitchell, Seminar Chair
Prof. Victor Pasko
Prof. Joan Redwing
Prof. Jeff Schiano
Prof. Srinivas Tadigadapa
Prof. Kenji Uchino

Prof. Andy Lau, Ethics Workshop Chair
Prof. David Salvia, Activities Chair
Ms. Amy Freeman, Director, Multicultural Program

Mrs. Linda Becker, Administration
Miss Amanda Skrabut, Assistant Director
Mr. Bo Wang, Publications Chair

2004 NSF EE REU STUDENTS, RESEARCH TOPICS, AND FACULTY MENTORS

REU STUDENTS, MAJOR, HOME INSTITUTION	RESEARCH TOPICS	FACULTY MENTORS AND AFFILIATIONS
Mr. James Berninghausen EE (Minor: Comp. Sci.) University of Portland	Nano Structure Influences on the Pyroelectric Detection	Prof. Amar Bhalla MRI and EE
Mr. Kevin Lin EE and Computer Science University of California Berkeley	Research and Development of the LionSat Communications Receiver Hardware	Prof. Sven Bilén EE and EDG
Mr. Kong Chhay Tep Electrical Engineering Michigan State University	Magnetoelastic Sensor Characterization and Measurements: Modeling, Simulation, and Experiments	Prof. Craig Grimes EE and MRI
Mr. Pape Sene Electrical engineering Polytechnic University	Electric Field Influenced Optical Modal Structure in Ferroelectric Single Crystal Fibers	Prof. Ruyan Guo EE and MRI
Mr. Joseph Tadduni EE and Economics Duke University	Piezoelectric Energy Harvesting	Prof. Heath Hofmann EE
Mr. Mohamed Abdulmagid EE/Wireless Engineering Auburn University	The Application of Adaptive Noise Canceling Filters in Intelligent Hearing Aids	Prof. Ken Jenkins EE
Mr. Xi Lin Electrical Engineering University of Illinois at Urbana-Champaign	Study of the Dynamics of Cross-Polarization Stimulated Orientation Scattering in Nematic Liquid Crystal Film	Prof. I. C. Khoo EE
Mr. Rehan Shariff Computer Engineering University of Illinois at Chicago	All Dielectric Meta-Materials	Prof. Mike Lanagan MRI and ESM

EE: Department of Electrical Engineering; **MRI**: Materials Research Institute; **MatSE**: Department of Materials Science and Engineering; **ESM**: Department of Engineering Science and Mechanics; **EDG**: Engineering Design & Graphics Program

2004 NSF EE REU STUDENTS, RESEARCH TOPICS, AND FACULTY MENTORS (Cont.)

REU STUDENTS, MAJOR, HOME INSTITUTION	RESEARCH TOPICS	FACULTY MENTORS AND AFFILIATIONS
Mr. Andrew Fontanella Electrical Engineering Penn State University	Study of Scattering from Nano-Structures	Prof. Zhiwen Liu EE
Mr. Ali Masnadi-Shirazi Electrical Engineering University of Texas, Arlington	Wavelet Analysis of Turbulence	Prof. J. Mitchell EE Prof. Charles Croskey EE
Miss Rachel Chisolm Physics/Engineering Washington and Lee University	Studies of Atmospheric Gravity Waves	Prof. Victor Pasko EE
Mr. John McGlade Electrical Engineering Penn State University	Studies of Electromagnetic Fields in the Atmosphere	Prof. Victor Pasko EE
Mr. Tyler Sullivan EE and Comp. Engineering The Pennsylvania State University	Studies of Lightning Induced Electrical Discharges in the Upper Atmosphere	Prof. Victor Pasko EE
Mr. Daniel Black Electrical/Computer Engineering University of Utah	Assembly and Characterization of Silicon Nanowires for Nanoelectronic Applications	Prof. Theresa Mayer EE Prof. Joan Redwing MatSE and EE
Miss Himani Suhag Electrical and Computer & Systems Engineering Rensselear Polytechnic Institute	Characterization of the Dynamic Behavior of Loss Mechanisms in Thin-Film High Temperature Superconducting Resonators	Prof. Jeff Schiano EE
Mr. Rakesh Reddy Electrical and Computer Engineering Carnegie Mellon University	Effect of Nanomaterials on the Resonant Characteristics of MEMS Structures	Prof. Srinivas Tadigadapa EE
Mr. Aria Pezeshk Electrical Engineering Michigan Technological University	The Development of an Ultrasonic Hypochlorous Acidic Humidifier for Disinfecting Application	Prof. Kenji Uchino EE and MRI

2004 NSF EE REU Summer Program
Weekly Scientific Seminar

Department of Electrical Engineering
Pennsylvania State University, University Park, PA 16802

Room 225 EE West, 11:00 am – 12:00 noon, Tuesdays
Chair: **Prof. John Mitchell** (jdm4@psu.edu)
(Followed by lunch break and **REU Weekly Ethics Workshop,** 12:30-1:30pm)

Date	*Topic*	*Speaker*
June 1, 2004	Welcome Meeting	Dr. Ruyan Guo
June 8, 2004	Digital Speech Signal Processing	Dr. Robert Nickel
June 15, 2004	Electrodynamic Tethers for Propulsion: Past, Present, and Future	Dr. Sven Bilén
June 22, 2004	Lightning-related Transient Luminous Events in the Middle Atmosphere	Dr. Victor Pasko
June 29, 2004	Sensors using Micro and Nanoscale Structures	Dr. Srinivas Tadigadapa
July 6, 2004	Energy Harvesting	Dr. Heath Hofmann
July 13, 2004	Nanoelectronic devices: single molecules to semiconductor nanowires	Dr. Theresa Mayer
July 20, 2004	What's Next? Graduate School, of Course!	Dr. Ken Jenkins

Suggestions and inquiries: eereu@engr.psu.edu

2004 NSF EE REU Summer Program
Weekly Ethics Workshop

Department of Electrical Engineering
Pennsylvania State University, University Park, PA 16802

Room 225 EE West, 12:30 pm –1:30 pm, Tuesdays
Chair: **Prof. Andrew Lau** (andylau@psu.edu)

Date	*Workshop Topic Addressed*	*Weekly*
June 1	(Welcome Meeting)	***Week 1***
June 8	Case study "Take-home exam" Intro to ethics: frameworks Project: ethical issues in engineering research	***Week 2***
June 15	Case study "data adjustment"; Research integrity	***Week 3***
June 22	Ethics in Engineering –Based on Piezoelectric Actuators, hosted by Prof. Kenji Uchino	***Week 4***
June 29	Dilbert™ Ethics Game	***Week 5***
July 6	Case study - trees; Environmental ethics: sustainability	***Week 6***
July 13	Student presentations: Ethics Issues in Engineering Research	***Week 7***
July 20	Student presentations: Ethics Issues in Engineering Research	***Week 8***

Suggestions and inquiries: eereu@engr.psu.edu

2004 NSF EE REU Bi-Weekly Field Trips and Industrial Sponsors

Nanofabrication National Facility
Penn State University
University Park, Pennsylvania

AccuWeather, Inc.
State College, Pennsylvania

The Learning Factory
Penn State University
University Park, Pennsylvania

Sound Technology, Inc.
State College, Pennsylvania

State of The Art, Inc.
State College, Pennsylvania

Applied Research Laboratory
Penn State University
University Park, Pennsylvania

Sullivan Research Park
Corning Inc., Corning, New York

2004 NSF EE REU SYMPOSIUM

8:55 AM to 3:45 PM, Thursday, July 29, 2004
Room 225 Electrical Engineering West Building
Pennsylvania State University
University Park, PA 16802

Time AM	Sessions and Topics	Chairs and Speakers
8:30-8:55	**Symposium Registration** (Refreshments Provided)	**Skrabut/Becker**
8:55–9:00	**Welcome**	**Jenkins/Mitchell**
9:00–10:15	**Session I** **Chairs**:	**Croskey/Mitchell**
9:00 – 9:15	RESEARCH AND DEVELOPMENT OF LIONSAT COMMUNICATIONS RECEIVER HARDWARE	Kevin Lin
9:15 – 9:30	STUDY OF ATMOSPHERIC GRAVITY WAVES	Rachel Chisolm
9:30 – 9:45	WAVELET ANALYSIS OF TURBULENCE IN THE MESOSPHERE	Ali Masnadi-Shirazi
9:45 – 10:00	EFFECTS OF SOLAR FLARES ON THE FIRST SCHUMANN RESONANCE FREQUENCY	John McGlade
10:00 – 10:15	PRINCIPLES OF ADAPTIVE NOISE CANCELING	Mohamed Abdulmagid
	10:15 – 10:30 am **Coffee Break**	
10:30 – 11:45	**Session II** **Chairs**:	**Grimes/Schiano/ Uchino/Hofmann**
10:30 – 10:45	PIEZOELECTRIC ENERGY HARVESTING USING ELEMENTS OF COUPLED MODE THEORY	Joseph Tadduni
10:45 – 11:00	ULTRASONIC PIEZOELECTRIC HYPOCHLOROUS ACID HUMIDIFIER FOR DISINFECTION APPLICATIONS	Aria Pezeshk
11:00 – 11:15	NONLINEAR BEHAVIOR OF MAGNETOELASTIC SENSORS: MODELING, SIMULATION, AND EXPERIMENT	K. Chhay Tep
11:15 – 11:30	CHARACTERIZATION OF THE DYNAMIC BEHAVIOR OF LOSS MECHANISMS IN THIN-FILM HIGH TEMPERATURE SUPERCONDUCTING RESONATORS	Himani Suhag
11:30 – 11:45	STEADY-STATE MODEL FOR THE CALCULATION OF THE ELECTRIC FIELD OF A SPRITE	Tyler Sullivan
11:45 - Noon	**(Group Photo)**	(Bo Wang)

2004 NSF EE REU SYMPOSIUM (cont.)

8:55 AM to 3:45 PM, Thursday, July 29, 2004
Room 225 Electrical Engineering West Building
Pennsylvania State University
University Park, PA 16802

12:00 – 1:15 pm Luncheon (Atherton Hotel)

Time PM	Sessions and Topics	Chairs and Speakers
1:30 – 2:30	**Session III** **Chairs**:	**Mayer/Bhalla Tadigadapa/Liu**
1:30 – 1:45	PLASMA-ASSISTED OXIDATION FOR SURFACE PASSIVATION OF SILICON NANOWIRES	Daniel Black
1:45 – 2:00	EFFECTS OF CARBON NANOTUBES ON MEMS STRUCTURES	Rakesh Reddy
2:00 – 2:15	ANALYSIS OF Ba0.5Sr0.5TiO3(BST):MgO BULK CERAMICS COMPOSED OF NANOSIZED PARTICLES FOR THE PYROELECTRIC IR SENSORS	James Berninghausen
2:15 – 2:30	MODELING SUPERCONTINUUM INDUCED GRATING DIFFRACTION PATTERNS USING RIGOROUS COUPLEDWAVE ANALYSIS	Andrew Fontanella
	2:30 – 2:45 pm **Coffee Break**	
2:45 – 3:45	**Session IV** **Chairs**:	**Pasko/Khoo /Lanagan**
2:45 – 3:00	STIMULATED ORIENTATIONAL SCATTERING IN NEMATIC LIQUID CRYSTAL WAVEGUIDE STRUCTURES	Xi Lin
3:00 – 3:15	INTENSITY PROFILES OF TRANSMITTED LIGHT IN FERROELECTRIC SINGLE CRYSTAL FIBERS	Pape Sene
3:15 – 3:30	ANALYSIS OF TUNABLE HORSESHOE RESONATORS USING DIELECTRIC LOADING TECHNIQUES	Rehan Shariff
3:15 – 3:30	AN INTRODUCTION TO MY LAB IN JAPAN AND A REPORT OF MY RESEARCH	Yasutaka Hiraki
3:30 – 3:45	CONGRATULATIONS AND CONCLUDING REMARKS	Guo/Jenkins /Skrabut

3:45 pm Adjournment

Picnic at Sunset Park, 5:00 pm – sunset, ALL INVITED

RESEARCH AND DEVELOPMENT OF LIONSAT COMMUNICATIONS RECEIVER HARDWARE

Kevin Lin* and Sven Bilén[#]
Department of Electrical Engineering
The Pennsylvania State University
University Park, PA 16802

*Undergraduate student of
Department of Electrical Engineering and Computer Sciences
University of California Berkeley
Berkeley, CA 94720

ABSTRACT

The LionSat Nanosatellite project requires a viable communications link that is low budget, lower power, and robust. Satellite telemetry introduces issues of frequency selection, operating temperatures, and radiation hardening that are not present in typical short-range wireless communications systems. There are several possible solutions for using a typical short-range wireless communications transceiver for a long-range satellite telemetry system. This article discusses these solutions, analyzes the power requirements, and provides a link budget analysis for the communications system. This article also examines the viability of using NASA's TDRSS (Tracking and Data Relay Satellite System) for LionSat communication purposes. Lastly, this article describes the testing of the design and evaluates the expected performance of the system for LionSat.

1. INTRODUCTION

The LionSat (Local *Ion*ospheric Measurements *Sat*ellite) is a student-lead design project at the Pennsylvania State University (Figure 1-1). The goals of the nanosatellite project include examining the wake and ram of the plasma surrounding the satellite as it flies through the ionosphere, and

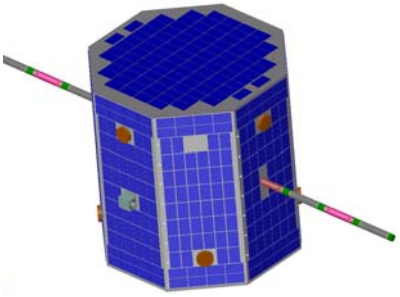


Figure 1-1. Conceptual Drawing of LionSat.

[#] Faculty Mentor

the testing of RF Ion thrusters to augment the spin of the NanoSat.[1] The purpose of the study is to research various designs for the communications uplink and to develop a working experimental model of the design. Because of the constraints imposed by the NanoSatellite III requirements, the communications link must be robust, require a small amount of power, function with a low-gain antenna that will be installed on LionSat, operate at the desired S-band frequency, and have a wide operating temperature range.

There are a number of integrated-circuit wireless transceivers that operate with a carrier 900 MHz and 2400 MHz. However, the LionSat uplink communications will operate at a non-commercial frequency between 2100 MHz and 2400 MHz. There are two approaches to using a commercial integrated circuit transceiver for our purposes. First, it is possible to couple the output of the signal into an external mixer, voltage controlled oscillator and phase locked loop to tune the frequency to our specifications. Second, the devices from some manufacturers may be able to support a carrier frequency lower than 2400 MHz, even though it is not within the specifications of the product. Furthermore, the transceiver on the satellite may be required to communicate with a third party transceiver because Penn State may not be allowed to transmit with the proposed LionSat Ground Station. Therefore, the transceiver needs to be adaptable and customizable so that communication with a third party transmitter can be achieved. This study analyzes various consumer integrated circuit transceivers and their suitability to the LionSat application.

LionSat may also use NASA's TDRSS (Tracking and Data Relay Satellite System) for satellite uplink and downlink communication. Currently, many satellites and systems use TDRSS, including the Hubble Space Telescope, the Space Shuttle, and the International Space Station. TDRSS would facilitate the frequency allocation process, because the frequencies are already allocated for TDRSS. This study examines the suitability of using TDRSS for LionSat.[2]

It is important to obtain an accurate estimate of the transmit power required to establish a healthy communications link between the satellite and the ground receiver. Because LionSat is a nanosatellite, it would be advantageous to use the least amount of power to operate. Various signal losses can be attributed to many factors such as free space loss, polarization loss, atmospheric loss, and antenna pointing loss, implementation loss.[3] Signal gain can also be acquired through antenna design and amplifier RF design. A detailed analysis of the link budgets of the TDRSS–LionSat system and the LionSat–Ground system is also presented in this study.

2. EXPERIMENTAL DESIGN

The first step in this process is to research all the commercially available integrated circuit transceivers and compare the specifications. All commercial transceivers are specified for 2400 MHz to 2500 MHz; however, some transceivers may be able to operate at a lower carrier frequency. Requesting more

detailed specifications from the manufacturers of the transceivers will confirm this. Information on TDRSS was obtained in a similar manner.

Previous LionSat studies[#] analyzed several factors resulting in signal loss. A refined link budget of the LionSat communications system can be defined by applying this information and using the specifications from the ChipCon CC2400 transceiver, the TDRSS system and information regarding the proposed orbit of LionSat. The AGI STK 5.0 (Satellite Tool Kit) simulator is used to estimate the range, the access window, and confirm the results of the link budget analysis.

The next step in the study is acquiring the transceiver board and performing tests to verify that the specifications are correct. Various frequencies and different FSK deviations are also analyzed to compare the spectrum output of different settings. The Agilent Technologies E4408B Spectrum Analyzer is the tool used for spectrum analysis. Power requirements are obtained and analyzed using various voltages to examine the power requirements for the transceiver system (Figure 2-1).

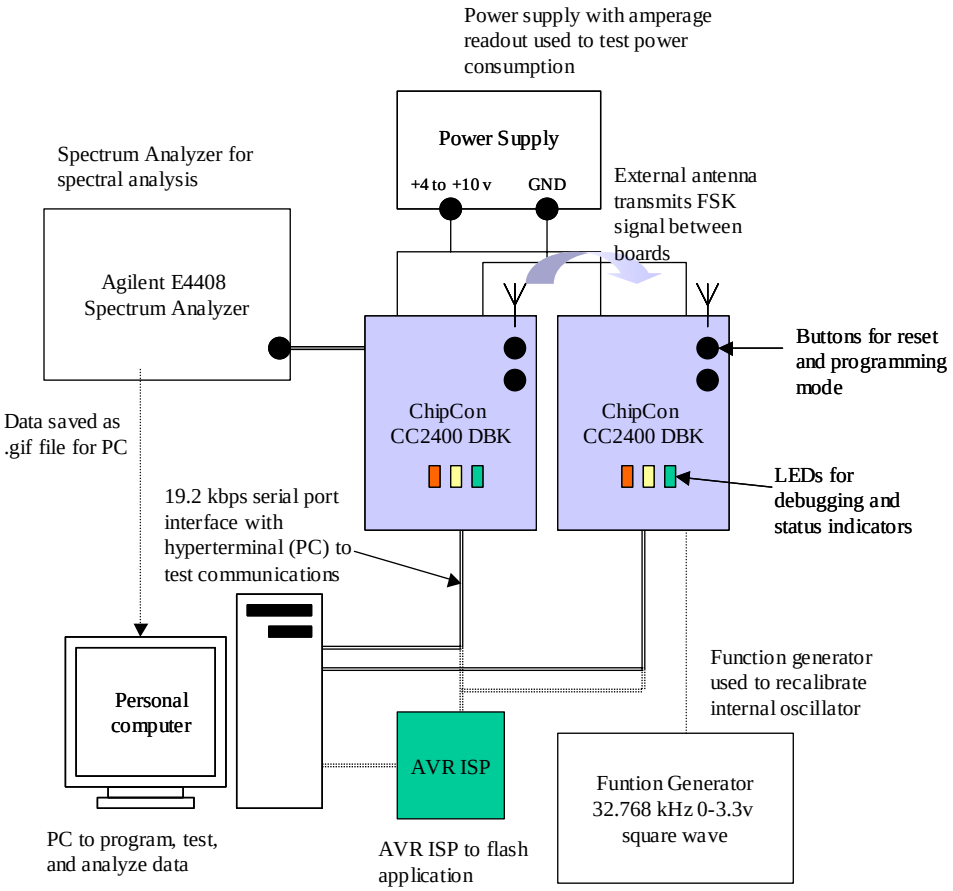


Figure 2-1. Experimental Setup to test the LionSat Uplink Communications System

[#] A Link Budget Analysis was performed by Dr. Charles Croskey, Penn State University.

Lastly, the chip needs to be adapted for LionSat purposes. Specifically, one transceiver needs to be "locked" in receive mode to send a signal output to the serial port at 19.2 kbps. The second transceiver needs to be "locked" in transmit mode to wait for an input signal from the serial port at 19.2 kbps. Also, the FSK deviation and carrier frequency needs to be changed for LionSat. This is accomplished by using and modifying sample C programs provided by ChipCon, using WIN-AVR GCC to compile the programs into machine code compatible with the ATMEGA8, and using AVR PROG to reprogram the flash memory of the ATMEGA8 microcontroller. The communications link is tested with a Hyperterminal connection and transmitting through the printed circuit-board wireless antennas.

3. RESULTS AND DISCUSSION

3.1 Selecting a Transceiver

The first approach to establishing a communications uplink with LionSat is to use a commercial 2.4 GHz transceiver and select a carrier frequency below 2.4 GHz. This is a simpler approach than coupling the output of a lower-frequency signal to a mixer, external voltage controlled oscillator, and a phase locked loop. The results of this research into several transceiver models (Table I and Table II) reveal that the MicroLinear ML2724 and the ChipCon CC2400 transceivers will both be able to transmit and receive at frequencies below 2.4 GHz. However, the MicroLinear transceiver can only transmit at a 1500 kbps baud rate, and as a result, has a less appealing receiver sensitivity than the ChipCon transceiver, which has a programmable baud rate of 10 kbps, 250 kbps, and 1000 kbps. The ChipCon transceiver is also more customizable and also provides a larger operating temperature range; thus, it is more suitable for LionSat.

3.2 Systems Design

The ChipCon CC2400 comes as a stand-alone chip, or as part of a circuit board package. The ChipCon CC2400 Demonstration Board Kit is most suitable for LionSat because it contains a 9-pin RS-232 serial port output interface and a programmable ATMEGA8 microcontroller with flash memory to store information regarding the frequency and modulation of the chip (Figure 3-1). The Penn State Ground Station (Figure 3-2) interfaces with the ChipCon CC2400 DBK with a power amplifier and band pass filter to boost the power to the requirements determined by the link budget discussed below. The LionSat communications system will be designed as follows: the output of the circuit is attached to a low noise amplifier, a postamp, and a band pass filter to improve performance and reduce the system noise temperature.[12] The board interfaces with the flight computer using a 19.2 kbps serial port interface with 8 data bits, 1 stop bit, no parity, and no flow control, which exceeds the 9.6 kbps data requirements (Figure 3-3).

Table I. 2.4 GHz Transceivers

Manufacturer		Honeywell[4]	ChipCon[5]	Micro Linear[6]	Toshiba[7]
Model		HRF-ROC094XC	CC2400	ML2724	TB32301AFL
Frequency (MHz)	**Min**	2400	2400	2400	2400
	Max	2500	2483	2485	2500
Op. Temperature (C)	**Min**	-40	-40	-10	-20
	Max	85	85	60	70
Storage Temperature (C)	**Min**	-40	-50	-65	-50
	Max	150	150	150	150
Power (mW)*	**Min**	62.4	31.51	135	
	Max	91	78.84	288	530
Data Transfer Rate (kbps)	**Min**	4	10		
	Max	128.8	1000	1500	1000
Rx Sensitivity (dBm)		-95	-87	-90	-85
Op. Voltage (V)		2.4 to 2.6	1.8/3.3	2.7 to 3.3 / 2.7to 4.5	2.7 to 3.3
Modulation		FSK	GFSK, FSK	FSK	
Can Operate <2400 MHz		**No**	**Yes**	**Yes**	**No**

Table II. 2.4 GHz Transceivers

Manufacturer		Atmel[8]	Motorola[9]	Nanotron[10]	Maxim[11]
Model		T2802	MC13190	nanoNET TRX	MAX2821
Frequency (MHz)	**Min**	2400 / 2289	2411	2380	2400
	Max	2500 / 2389	2473	2520	2499
Op. Temperature (C)	**Min**	-25	0	-40	-40
	Max	85	50	85	85
Storage Temperature (C)	**Min**	-40	-65	-40	-65
	Max	150	150	125	160
Power (mW)*	**Min**	172.8			94.5
	Max	391	160	TBD	396
Data Transfer Rate (kbps)	**Min**				1000
	Max				22000
Rx Sensitivity (dBm)			-71		-97
Op. Voltage (V)		3.2 to 4.6	4.8		2.7 to 3.6
Modulation				MDMA (Chirp)	
Can Operate <2400 MHz		**No**	**No**	**No**	**No**

Figure 3-1. CC2400 DBK board

Figure 3-2. Penn State Ground Station

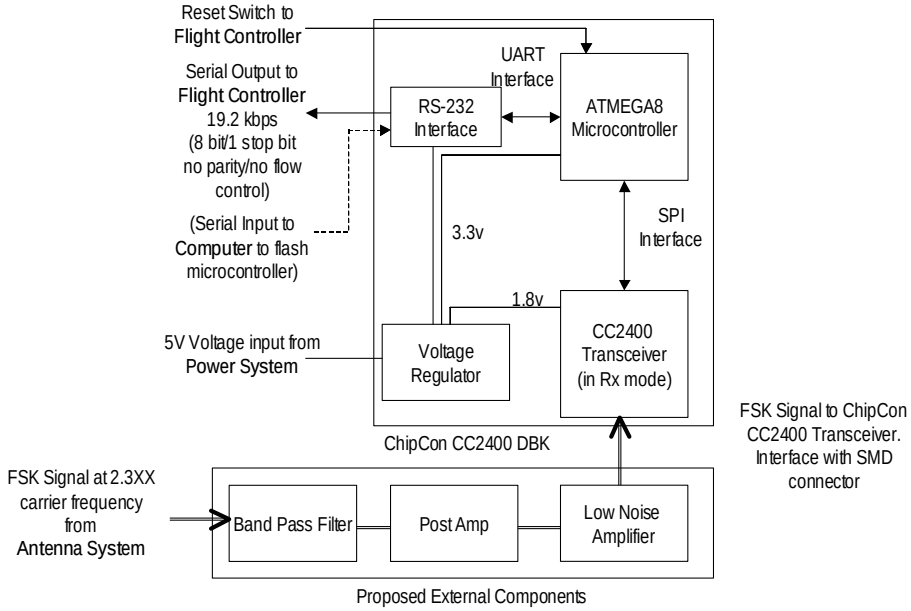


Figure 3-3. PSU-LionSat Communications Uplink System

3.3 The TDRSS Alternative

Alternatively, LionSat can use NASA's Tracking and Data Relay Satellite System (TDRSS) to communicate with Penn State. Since commercial TDRSS systems are expensive and power-hungry, NASA promises to give support for code to program a Xilinx FPGA to communicate with TDRSS (Figure 3-4).[13]

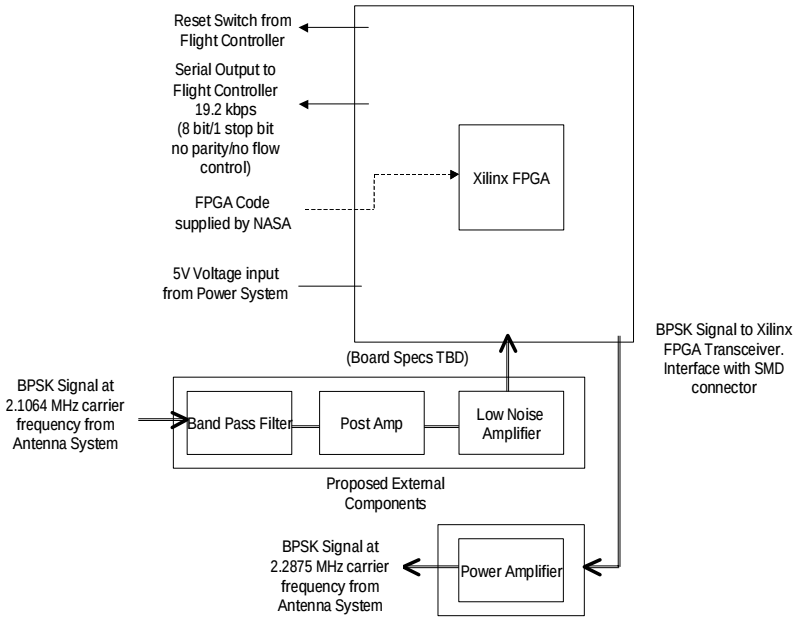


Figure 3-4. TDRSS-LionSat Communications System

TDRSS consists of several satellites in geosynchronous orbit. These satellites relay information to and from other satellites and spacecraft, such as the Hubble Space Telescope, the International Space Station, and the Space Shuttle. Currently, the TDRS-East satellite is serviced by F-4 at 41 W longitude, and the TDRS-West satellite is serviced by the F-5 satellite at 174 W longitude; both satellites communicate with NASA's White Sands Complex in New Mexico.[2] A third TDRS satellite, F-3 is located at 275W longitude and communicates with the Guam Remote Ground Terminal. LionSat can communicate with TDRSS by using the Multiple Access system, which transmits a BPSK signal at S-band frequencies. TDRSS eliminates the problem of frequency allocation and compatibility that we have with the conventional PSU-LionSat communication.

3.4 PSU-Lionsat Link Budget Analysis

A detailed Link Budget is calculated for the PSU-LionSat uplink system, using a method described by J. Wertz in Space Mission Analysis and Design.[14] The Equivalent Isotropic Radiated Power (EIRP) is the equivalent input power an isotropic radiator needs as input to match the performance of the LionSat Ground Station. Factors that contribute to the EIRP (Eq. 1 and Eq. 2) includes transmitter power (the amount of power the ground station can transmit into space), transmitter line loss (the losses through cable routing from the power amplifier to the antenna) transmit antenna gain (the gain of the Penn State ground station which has a 3.6 m parabolic antenna), and transmit antenna pointing loss (the misalignment of the ground station antenna to the satellite due to tracking errors).

$$EIRP = P + Lpt1 + Gt + Lpt2 \quad (1)$$

Transmitter Power (P) = 10 watts = 10 dB
Transmitter Line Loss (Lpt1) = -1 dB[15]
Transmit Antenna Gain (Gt) >=32.5 dB#
Transmit Antenna Pointing Loss (Lpt2) = 0.5 dB[15]

$$EIRP = 10dB - 1dB + 32.5dB - 0.5dB \quad (2)$$

$$EIRP = 41dB$$

Signal losses can be attributed to the Free Space Transmission Loss (Ls) (the loss of signal strength as a result of the spreading of the RF signal through space) (Eq. 3 and Eq. 4), the Polarization Loss (Lp) (the ground station is right-hand circular polarized# and the satellite is linearly polarized), the Atmospheric Loss (La2) (loss of signal strength due to air molecules). The ratio of received energy per bit to noise density (Eb/No) can be calculated by subtracting these losses from the EIRP and taking into account the system noise temperature (G/T) (the sum of all external contributions of noise) and the Receiver Antenna Gain (Grp) (Eq. 5 and Eq. 6).

$$Ls = 147.55 - 20\log(S) - 20\log(F) \quad (3)$$

Orbit Altitude = 400 km
Total Length = 2000 km (based on AGI STK 5.0 simulation)
Frequency = 2300 MHz

$$Ls = 147.55 - 20\log(2 \cdot 10^6 m) - 20\log(2.3 \cdot 10^9) \quad (4)$$

$$Ls = -165.705dB$$

$$Eb/No = [EIRP] + [Ls] + [La1] + [La2] + [Gr] - [Ts] + [228.6] - [10\log R] \quad (5)$$

[Ls] = -165.705 dB
[La1] = -3 dB[15]
[La2] = -.5dB
[Grp] = 0 dB (LionSat antenna is assumed to be omnidirectional in performance)
[Ts] = -26.5 dB[12]

$$Eb/No = 41dB - 165.705dB - 3dB - 0.5dB + 0dB - 26.5dB + 228.6dB - 10\log(1 \cdot 10^4) \quad (6)$$

$$Eb/No = 34.53dB$$

Specification for proposed Penn State ground antenna system, information from Operations and Maintenance Manual. Satellite Telemetry Tracking Antenna System. Model 100-12 Part No. 3036000-01. pp 1-8.

The Link Margin (Eq. 7 and Eq. 8) is calculated by determining the required Eb/No and taking into account the implementation Loss.

$$Margin = Eb/No_{actual} - Eb/No_{required} - Li \quad (7)$$

Eb/No_{actual} = 34.53 dB
$Eb/No_{required}$ = 14 dB (based on FSK modulation 1E-6 BER)
Li = 2 dB[15]

$$Margin = 34.53dB - 14dB - 2dB \quad (8)$$

$$Margin = 18.53dB$$

The Link Budget of the Penn State-LionSat System shows that a healthy link can be obtained using 10 watts of transmit power from the Penn State ground station. It is important to note that formal tests of the LionSat antenna system have not been performed, and the antenna performance estimates are optimistic because these estimates do not account for antenna nulls and losses that will decrease the link margin. However, the large link margin can compensate for some antenna loss.

3.5 TDRSS-Lionsat Link Budget Analysis

A different but similar approach is used to calculate the link budget of the TDRSS-LionSat system based on information from NASA.[16] First, the forward link budget of the system is computed (Eq. 9 to Eq. 16).

$$Ls = -(32.45dB + 20\log(R) + 20\log(f)) \quad (9)$$

R = 42,500 km (based on AGI STK 5.0 estimates)
f = 2106.4 MHz (frequency of TDRSS S-band return link)

$$Ls = -(32.45dB + 20\log(42500km) + 20\log(2106.4MHz)) \quad (10)$$

$$Ls = -191.487dB$$

$$Prec/No = [EIRP] + [Ls] + [Lp] + [L\theta] + [(G/T)] - 10\log(k) \quad (11)$$

Ls = -191.487 dB
Lp = -3 dB[15]
Lθ = 0 dB
G/T = -24.62 dB (assuming no antenna gain and system noise temperature of 290K)

$$Prec/No = 42db - 191.487dB - 3dB - 0dB - 24.62dB + 228.6dB \quad (12)$$

$$Prec/No = 51.493dB$$

$$(Eb/No)_{predicted} = Prec/No - 10\log Rd + c \quad (13)$$

Rd = 9.6 kbps (LionSat requirements)
C = -2 dB (estimate)

$$(Eb/No)_{predicted} = 51.493dB - 10\log(9600bps) - 2dB \quad (14)$$

$$(Eb/No)_{predicted} = 9.67dB$$

$$Margin = (Eb/No)_{predicted} - (Eb/No)_{required} \quad (15)$$

(Eb/No) required = -10.5 dB (BPSK signal, 1E-6 BER) (16)

$$Margin = 9.67dB - 10.5dB$$

$$Margin = -.83dB$$

The TDRSS return link budget is also calculated based on information provided by NASA[17] to determine how much power is needed to acquire a healthy link between LionSat and TDRSS (Eq. 19 to Eq. 28).

Prec(ideal) = 10 log *Rd* + *c* (17)
Rd = 200 kbps (LionSat specifications)
c = -221.8 dB (TDRSS specifications) (18)

Prec(ideal) = 10 log (200000 *bps*) – 221.8*dB*
Prec(ideal) = -168.79*dB*

Prec(predicted) = *Prec(ideal)* - *Lϕ* - *Lp* – Li – *Lnc* (19)
Lp = -3 dB[15]
Lθ = 0 dB
Li = -2 dB (TDRSS estimation)

Prec(predicted) = -168.79dB +0dB + 3 dB + 2 dB (20)
Prec(predicted) = -163.79 dB

$$Ls = -(32.45dB + 20\log(R) + 20\log(f)) \quad (21)$$

R = 42500 km (AGI STK 5.0 simulation)
F = 2287.5 MHz (TDRSS specification)

$$Ls = -(32.45dB + 20\log(42500km) + 20\log(2287.5MHz))$$

$$Ls = -192.205dB \quad (22)$$

EIRP(min) = *Prec(predicted)* – *Ls* (23)

$$EIRP(min) = -163.79\text{dB} + 192.205\text{dB} \quad (24)$$
$$EIRP(min) = 28.415$$

$$EIRP = P + G - Lpi \quad (25)$$
P = 1750 W, 32.43 dB (power needed to transmit)
G = 0 dB (assumed 0 dB for omni-directional antenna)
Lpi = 1 dB[15]

$$EIRP = 32.43\text{ dB} - 0\text{ dB} - 1\text{ dB} \quad (26)$$
$$EIRP(min) = 31.43$$

$$Margin = EIRP - EIRP(min) \quad (27)$$

$$Margin = 31.43 - 28.415 \quad (28)$$
$$Margin = 3.015\text{ dB}$$

The power required to transmit from LionSat to TDRSS is a function of the baud rate. The graph below (Figure 3-5) examines the power vs. baud rate of the LionSat-TDRSS system.

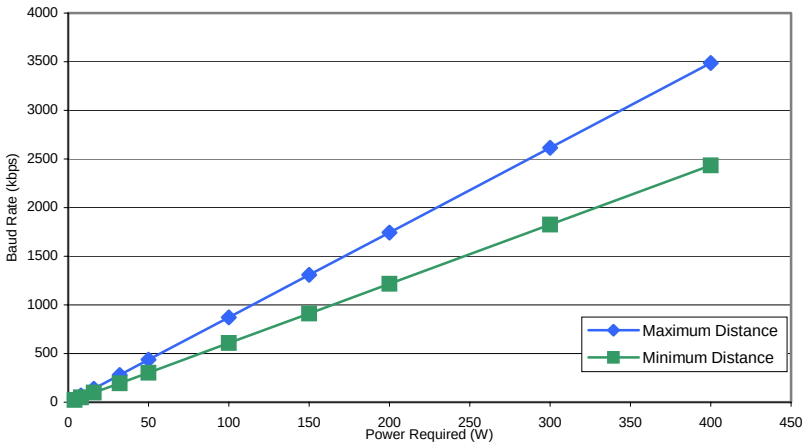


Figure 3-5. TDRSS Power vs. Baud Rate

The TDRSS link budget shows less promising results because it is more difficult to transmit through a greater expanse of free space. The forward service (uplink) has a –0.83 Link Margin for a bit error rate of 1E-6 and the return service (downlink) budget shows that a large amount of power will be needed to establish a link to TDRSS. Therefore, there will probably be a higher bit error rate for the forward link transfer, and the return link transfer may not be feasible for LionSat because of the outrageous power requirements for the link budget and the stringent power requirements of LionSat. On a positive note, LionSat is constantly in contact with TDRSS so it may be possible to reduce the baud rate of

the output (Figure 3-5), which will effectively decrease the power required to transmit with a healthy link.

3.6 Power Budget Analysis

The power output of the components of the PSU-LionSat uplink communications system is analyzed to make an estimated overall power budget (Table III). To verify the power requirements of the ChipCon CC2400 DBK, the board is connected to an external power supply, the output voltage is varied from 2V to 10V, and the current output reading is recorded from the power supply (Figure 3-6).

Table III. Power Budget of LionSat Communications Uplink

Component	Power Requirement
ChipCon CC2400 DBK	200 mW
Low Noise Amplifier (TBD)	32 mW
Band Pass Filter	0 mW
Post Amplifier	140 mW
Total Power Requirements:	372 mW

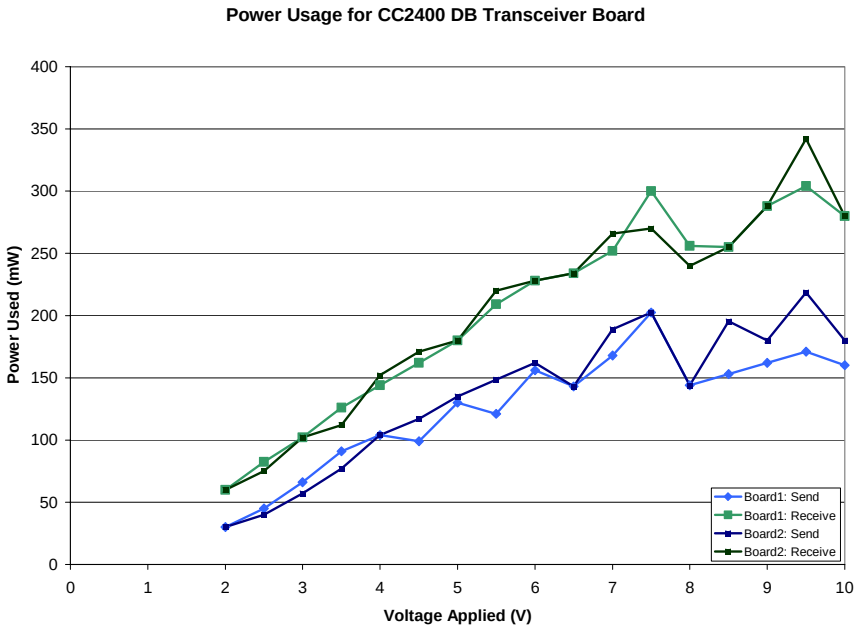


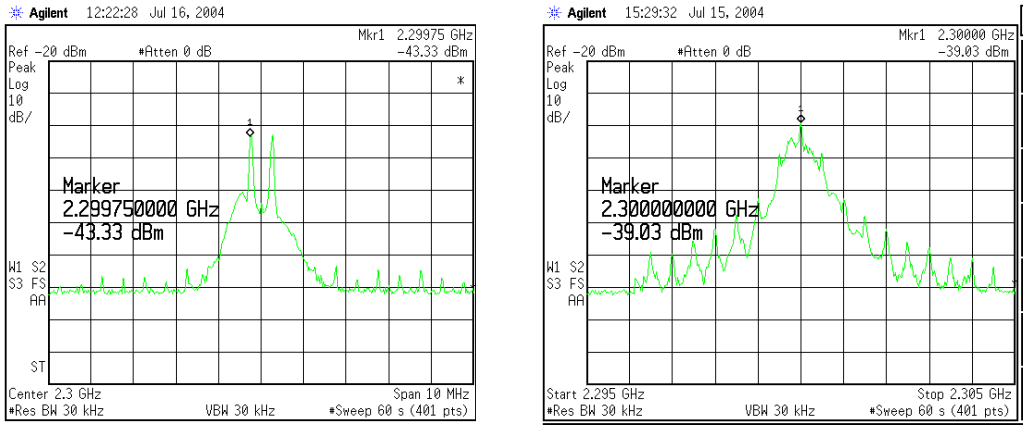
Figure 3-6. Power Usage for CC2400 DB Transceiver Board

Both ChipCon boards exhibit a similar power requirement curve, showing a linear decrease in power consumption as the voltage decreases. There are two voltage regulators that convert the input supply of 4 to 10 v to a usable voltage[18]; thus, the current requirements of system remains relatively constant while the voltage is varied. The results show that it may be most efficient to use a 3.3 volt

supply from the LionSat power system. However, since the EEPROM may be corrupted if the voltage drops below 2.7 volts[19], a 5 volt power supply may be a safer option for data integrity. Regardless of whether a 3.3volt or 5 volt supply is used, the design uses a small amount of power and meets LionSat and NanoSat III requirements.

3.7 PSU-LionSat Uplink Communications System Testing

In order to determine the feasibility of transmitting a signal at a frequency lower than its normal operating range, the CC2400 is tested under various frequencies using the "rfblinkled" program supplied by ChipCon. When button S2 is pressed, the green LED will blink on the receiving board and the yellow LED will blink on the transmitting board indicating a communications link between the transceivers. The results are observed using the Agilent Technologies E4408B Spectrum Analyzer connected to the SMA output of the ChipCon CC2400 DBK. The baud rate and FSK deviation can be decreased to suit LionSat applications. Specifically, the output of a 2300 MHz carrier frequency with 250 kHz FSK deviation and a 10 mbps baud rate is compared to the standard 2450 MHz carrier frequency with 250 kHz FSK deviation and 1000 mbps baud rate to determine the changes in spectral output (Figure 3-7).



(a) Fc = 2300 MHz; Fdev = 250 kHz; baud rate = 10 kbps

(b) Fc = 2450 MHz; Fdev = 250 kHz; baud rate = 1000 kbps

Figure 3-7: Comparison of spectral output by varying carrier frequencies and baud rate

The ChipCon CC2400 transceiver uses Frequency Shift Keying (FSK) modulation, where one frequency is used to represent a binary 1 and a second frequency is used to represent a binary 0.[3] The twin peaks that are visible in Figure 3-7a represent the two frequencies used in FSK modulation. There is an external antenna attached to the ChipCon CC2400 board, so a significant portion of the power output is delivered to this antenna; therefore the decibel readings from the spectrum analyzer do not indicate the total amount of power transmitted by the board. Nonetheless, the output indicated by the spectrum analyzer can be

used to compare the spectral width and patterns of varying parameters. The results from the Spectrum Analyzer confirm that the ChipCon CC2400 can be successfully tuned to any carrier frequency between 2250 MHz and 2500 MHz.

The spectral width can be decreased by several methods to improve performance and adhere to frequency spectrum regulations, if necessary. The frequency deviation of the FSK modulation can also be modified from 3.90625 kHz to 500 kHz; a smaller FSK frequency deviation produces a narrower power band. A Gaussian Frequency Shift Key modulation (GFSK) can be applied to limit the spectral width. The GFSK modulation is similar to a FSK modulation, but the GFSK passes the FSK signal through a Gaussian Filter to limit the spectral width.[20] Lastly, the spectral width can be reduced by decreasing the baud rate of the output, since the majority of the power output of a FSK modulation lies within twice the baud rate of the transmitted signal.[21] Since LionSat only requires an uplink baud rate of 9.6 kbps, it may be advantageous to use the 10 kbps baud rate for transmission for best results. Figure 3-7 shows that the customized FSK signal for LionSat has a narrower spectral width than the standard FSK signal, and confirms the ability of the ChipCon CC2400 to transmit at 2300 MHz.

The communications verification using the serial port had mixed results. Initially, one of the ChipCon CC2400 boards had a corrupt bootloader, so the software to perform communications tests between the two boards could not be loaded. Furthermore, problems persisted with serial port input and output due to EEPROM corruption, possibly due to abnormally low input voltages during several flash cycles. A recalibration of the internal oscillator of the ATMEGA8 using a 38.768 kHz reference signal was required for two of the boards. Using AVR ISP, these problems were resolved, and communication was established between the computer and the two transceiver boards using a 19.2 kbps serial port interface and Hyperterminal. The last test demonstrates that it is feasible to use the transceiver for LionSat purposes, provided that external components are added to amplify the signal and to improve receiver sensitivity. However, the snafus also raise the question of whether these components are reliable enough for a space mission.

4. CONCLUSION

The testing of the ChipCon CC2400 proves that it is a versatile board capable of meeting our requirements. LionSat requires low power consumption, an S-band frequency below 2400 MHz, and a 9.6 kbps minimum uplink transmission data rate, all of which are met or exceeded by the ChipCon CC2400 board. Furthermore, the CC2400 transceiver supports a range of frequency deviations, a choice of GFSK or FSK modulation, and a choice of three different baud rates. It also supports data structuring such as CRC and buffering. Thus, the transceiver can be configured to communicate with a third party transceiver; also, the transceiver can be configured to minimize the spectral width if the Penn State ground station is to be used.

Additional tests need to be performed to verify the communications uplink system for LionSat. Most importantly, a formal frequency allocation request needs to be submitted so that the Penn State ground station can be used to transmit to LionSat. Extensive Hyperterminal and serial port tests should be performed to confirm a low bit error rate. Once antenna designs are finalized, the antenna system should be tested to confirm the antenna gains and losses as well as the system noise temperature so that a more formalized link budget can be devised. Furthermore, range tests can be performed by disconnecting the external printed circuit board antenna from the CC2400 DBK transceiver board, and connecting attenuators to the SMA output. A low noise amplifier, band pass filter, postamp, and a power amplifier should be used to simulate range testing as well. The TDRSS option cannot be eliminated without further discussion with NASA representatives and the finalization of the LionSat power budget. It may be possible to negotiate for a longer communications window to decrease the data rate. LionSat may also have a higher power system budget than presently anticipated, which would aid in the communications uplink with TDRSS.

Regardless of which communications system is chosen, this study shows that it is feasible to build a viable communications system to satisfy all requirements of the LionSat nanosatellite.

ACKNOWLEDGEMENTS

I would like to thank the National Science Foundation for funding the research and the Pennsylvania State University for supplying the facilities necessary to conduct the research. I would like to thank Siri Namtvedt, Morten Bråthen, and Geir Eivind Jonsrud for offering technical support for the ChipCon CC2400 transceiver. Thank you to Jason Soloff, Bill Anselm, and John Staren of NASA for offering technical support regarding the TDRSS system and advice on satellite communications.

I would also like to thank those involved in the Penn State EEREU program, specifically Linda Becker, Amanda Skrabut, Bo Wang, Professor Ken Jenkins, and Professor Ruyan Guo for organizing an excellent program, sacrificing personal time for the happiness and well being for the undergraduate research students including myself, and providing an educational, exciting, and stimulating nine-week research experience. Thanks to Brendan Surrusco, the LionSat Project Manager, and the LionSat team for the help with various research related questions, the guidance through the design and test process, and the opportunity to work on the satellite. Thank you, Professor Mark Wharton, for helping with the design of the communications system. Thank you, Professor Charles Croskey, for your wisdom regarding satellite communications and help with the report. I would also like to thank my faculty mentor, Professor Sven Bilén, for answering questions, providing guidance, and helping me formulate the report and presentation. Lastly, I would like to thank my family for supporting me through all of my endeavors.

This material is based on work supported by National Science Foundation Grant No. EEC-0244030.

REFERENCES

[1] LionSat. *Penn State University*. 7 July 2004 <http://www.courses.psu.edu/ee/lionsat/LionSat_files/SubPages/mission.htm>.

[2] TDRSS Overview. *NASA Space Network Online Information Center*. 7 July 2004 <http://nmsp.gsfc.nasa.gov/tdrss/oview.html>.

[3] Roddy, Dennis. *Satellite Communications, 3rd Ed.* New York: Prentice-Hall Inc., 2001.

[4] HRF-ROC094XC Transceiver. *Honeywell*. 20 July 2004 <http://www.ssec.honeywell.com/microwave/products/HRF-ROC094XC.pdf>.

[5] ChipCon SmartRF CC2400. 5 Feb 2004. *ChipCon AS*. 20 July 2004 <http://www.chipcon.com/files/CC2400_Data_Sheet_1_2.pdf>.

[6] MicroLinear ML2724. April 2003. *Micro Linear Corporation.* 20 July 2004 <http://www.microlinear.com/downloads/DS/DS2724-F-01.pdf >.

[7] TB32301AFL 2.4 GHz Radio Communication IC. *Toshiba America Electronics Components Inc. 20 July 2004* <http://www.toshiba.com/taec/components/docs/ProdBrief/TB32301AFL_doc_fl.pdf>.

[8] 2.4 GHz WDECT/ISM Single-Chip Transceiver T2802 Preliminary. *Atmel Corporation.* 20 July 2004 <*http*://www.atmel.com/dyn/resources/prod_documents/doc4509.pdf>.

[9] Motorola. 2002. *Motorola Inc*. 20 July 2004 <http://www.motorola.com/mediacenter/news/detail/0,,1472_1134_23,00.html>.

[10] NanoNet TRX Datasheet Ver. 2.00 Preliminary. 10 May 2004. *Nanotron Technologies*. 20 July 2004 <http://www.nanotron.de/nanotron_intern/en/downloads/datasheets/001nanoNET_TRX_Datasheet.pdf>.

[11] Maxim 2.4 GHz 802.11b Zero-IF Transceivers. *Dallas Semiconductor*. 20 July 2004 <http://pdfserv.maxim-ic.com/en/ds/MAX2820-MAX2821A.pdf>.

[12] Personal Communication, M. Wharton, 2004.

[13] Schloff, Jason. Teleconference. 21 June 2004.

[14] Wertz, James R. and Wiley J. Larson ed. *Space Mission Analysis and Design*. 3rd ed. Torrance, CA: W. J. Larson and Microcosm, Inc., 1999.

[15] Personal Communication, C. Croskey, 2004.

[16] Forward Link Analysis. *NASA Space Network Online Information Center*. 16 July 2004 <http://nmsp.gsfc.nasa.gov/tdrss/fwdlnk.html>.

[17] Return Link Analysis. *NASA Space Network Online Information Center*. 16 July 2004 <http://nmsp.gsfc.nasa.gov/tdrss/rtnlnk.html>.

[18] User Manual Rev. 1.2 SmartRF CC2400 Demonstration Board Kit. *ChipCon AS*. 19 July 2004 <http://www.chipcon.com/files/CC2400DBK_User_Manual_1_2.pdf>.

[19] AVR053: Calibration of the internal RC oscillator. *Atmel* Corporation. 19 July 2004 <http://www.atmel.com/dyn/resources/prod_documents/doc2555.pdf>.

[20] GFSK differences and Advantages of FSK Modulation. *PaloWireless Bluetooth Wireless resource Center*. 19 July 2004 <http://www.palowireless.com/infotooth/knowbase/radio/109.asp>.

[21] Watson, Bob. FSK: Signals and Demodulation. *W. J. Communications.* 19 July 2004 <http://www.wj.com/pdf/technotes/FSK_signals_demod.pdf>.

STUDY OF ATMOSPHERIC GRAVITY WAVES

Rachel Chisolm*, Jonathan B. Snively[+], and Victor P. Pasko[#]

Department of Electrical Engineering
The Pennsylvania State University
University Park, PA 16802

*Undergraduate student of
Department of Physics/Engineering
Washington and Lee University
Lexington, VA 24450

ABSTRACT

Atmospheric gravity waves are traveling buoyant wave motions generated by mechanical oscillations of air in the upper troposphere associated with thunderstorms. They can propagate to altitudes as high as the mesosphere and lower thermosphere where they interact with chemically active regions of the atmosphere, modulating optical emissions commonly known as airglow. Gravity wave motions and their interaction with airglow chemistry can be simulated using a numerical model, the output of which models the airglow emissions for excited hydroxyl (OH) and atomic oxygen ($O(^1S)$). By integrating photon volume emissions rates from the simulated data for all heights of the atmosphere it is possible to calculate how much light would be collected by an airglow imaging system located on the ground. Reactions have characteristic wavelengths of airglow emissions and occur at specific altitudes; therefore each emission is considered separately. The intensity of the light that is actually recorded at the ground depends on the integration time of the camera and the integration over altitude of the airglow emissions. Numerical interpolation and integration methods are used to analyze simulated gravity wave-induced airglow perturbations, which may be compared with observed airglow data.

[#] Faculty Mentor
[+] Graduate Mentor

INTRODUCTION

Gravity Waves:

The Brunt-Väisälä frequency is the frequency at which a parcel of air will oscillate when displaced from its equilibrium. The Brunt-Väisälä frequency changes with altitude because it is dependant on several different parameters of the atmosphere, including the pressure, density, and temperature. The Brunt-Väisälä frequency (N) is defined as:

$$N^2 = \frac{g}{\theta}\frac{\partial\theta}{\partial z} \tag{1}$$

where θ is the potential temperature.

Atmospheric gravity waves are buoyant wave motions that carry momentum and energy throughout atmospheric regions. Short period, small scale gravity waves often originate in the upper troposphere and propagate into the mesosphere and lower thermosphere. Their origin around the tropopause is often the result of the motions of parcels of air that are heated during thunderstorms[1]. The pressure and temperature differences can cause a parcel of air to rise until the parcel reaches its equilibrium height[1]. It overshoots this height, and gravity acts as the restoring force. The gravity and buoyancy forces oppose each other, causing the parcel of air to oscillate at the Brunt-Väisälä frequency[2]. The energy from this oscillation can be dissipated in the form of gravity waves, which will thus have a frequency roughly equal to the Brunt- Väisälä frequency at the height where the wave is generated[1]. For a gravity wave to be able to propagate vertically, its frequency must be less than the Brunt-Väisälä frequency[2]. Any wave with a frequency greater than the local Brunt-Väisälä frequency would be reflected or lose energy to evanescence. The magnitude of a wave's velocity perturbation increases with altitude as it propagates vertically because its kinetic energy is dependant on the product of density and velocity squared. Because the density of the atmosphere decreases exponentially with altitude, the wave velocity must increase exponentially for energy to be conserved[2].

The dispersion relation is an equation derived from Euler's equations of motion that relates the vertical and horizontal wave number (m, k) with the frequency (ω) [3]. This relation is most easily derived under the assumption of an incompressible atmosphere, where the horizontal phase velocity is much less than the speed of sound. The incompressible atmosphere is a reasonable assumption and is beneficial because it eliminates acoustic wave solutions [2]. The dispersion relation can be stated for vertical wave numbers as:

$$\left(\frac{N^2}{\omega^2} - 1\right)k^2 = m^2 \tag{2}$$

The wave energy propagates along the line of the group velocity, the velocity at which a packet of waves travels, which is up and away from the source. The phase propagation, the direction in which individual waves

propagate, is downward and perpendicular to the direction of the group velocity. The horizontal wavelength and phase velocity are important because they are the characteristics of a gravity wave that are most easily studied from the ground; they can be seen in Figure 1, which is based on a diagram by Taylor and Hapgood[4]. The phase velocity is given as:

$$v_{px} = \frac{\omega}{k} = \sqrt{\frac{N^2 - \omega^2}{m^2}} \tag{3}$$

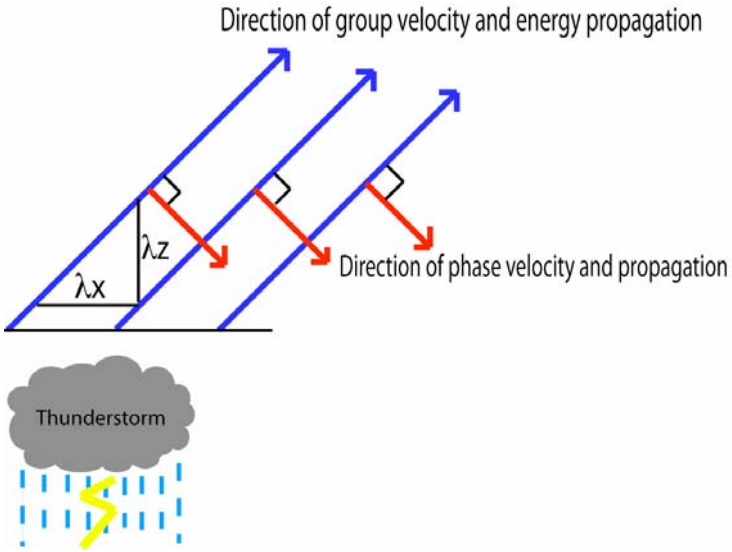


Figure 1. A diagram of the generation and structure of a gravity wave with a tropospheric source and the directions of the phase and group velocities

Airglow:

Gravity waves propagate into the upper atmosphere where they interact with background airglow emissions[4]. Each airglow emission has a different peak height and characteristic wavelength that depends on the species that is being excited. The airglow emissions studied in this paper are the near-infrared hydroxyl (OH) emission which has a peak height of 87 km and the 557.7 nm green line atomic oxygen (OI) emission that has a peak height of 96 km[4]. The peak height at which the airglow occurs is due to the composition of the atmosphere, and the peak height is where the atmosphere lends itself most readily to the reactions that create the airglow. Each airglow wavelength is due to a different chemical reaction in which the characteristic molecule or atom is excited. The numerical data we use for the hydroxyl emissions is for the OH (8-3) emission. The airglow is a result of the energy released as photons when the excited molecule returns to its unexcited state. The reaction that produces the vibrationally excited OH(v) is[5]:

$$H + O_3 \rightarrow OH(v) + O_2 \qquad (v=8) \qquad (4)$$

The excitation of $O(^1s)$ results in the 557.7 nm emission and is produced by a series of two reactions[6]:

$$O + O + M \rightarrow O_2 * + M \qquad (5)$$
$$O_2 + O \rightarrow O(^1S) + O_2 \qquad (6)$$

Airglow Imagers:

Ground-based airglow imagers have filters so that images can be taken of one airglow emission at a time. From the images taken, the structure of the wave can be seen, and the wavelength and direction of propagation can be deduced. Figure 2 shows an example of an all-sky airglow imager to demonstrate how the different components might be laid out. The equipment characteristics used in this paper are similar to those used in the ALOHA-93 campaign described by Taylor and Bishop[7]. The main components of the system are the telecentric lens arrangement, a filter wheel, and a cooled CCD camera. The lens system uses a fisheye lens for all-sky viewing capability (180°). The OI filter has a bandwidth of 2.65 nm centered on the OI wavelength of 557.7 nm with a transmission of 83%. The OH emission has wavelengths ranging from 715 to 930 nm, and the filter has a notch at 865 nm so that the O_2 emissions at that wavelength cannot pass through; the transmission for the OH filter is 80%. The OI images were taken over an integration time of 90 sec, and the OH integration time was 20 sec. The longer integration time is dictated by the camera sensitivity and the faintness of the airglow emission.

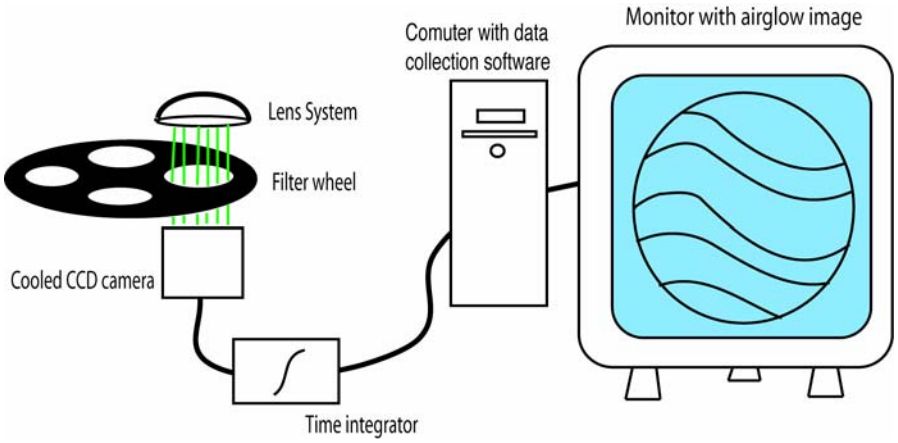


Figure 2. Diagram of an airglow imaging system that would be used to make ground-based observations of airglow.

The goal of this project is to create a virtual airglow imaging system that collects the data from a simulation and models what would be seen by a ground-based imager. This data is then compared directly with observations.

EXPERIMENTAL PROCEDURES

Theory:

From a model atmosphere, a steady-state emissions profile is produced, which shows the volume emissions rate (photons/cm^{3}s) over a given height range in the atmosphere (Figure 3). To find the photon emissions that can be seen from the ground, the steady state profile must be integrated over the entire height for a given column of air:

$$\int_0^\infty V_{OI}\,dz = \varepsilon_{OI} \tag{7}$$

$$\int_0^\infty V_{OH}\,dz = \varepsilon_{OH} \tag{8}$$

V_{OI} (cm^{-3}s^{-1}) is the photon volume emissions rate for atomic oxygen, and V_{OH} (cm^{-3}s^{-1}) is the photon volume emissions rate for hydroxyl. ε_{OI} (cm^{-2}s^{-1}) is the total photon emissions for OI for a given column of air, and ε_{OH} (cm^{-2}s^{-1}) is the total photon emissions for OH for a given column of air. This integration over height can be done using the trapezoid rule because there is a sufficient number of data points to be integrated that it is an accurate estimation of the integral.

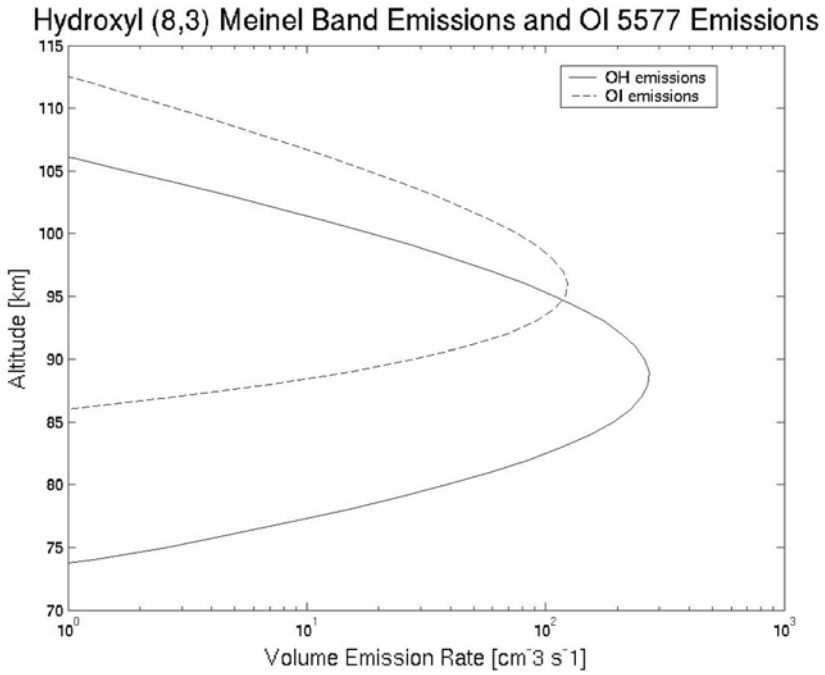


Figure 3. The steady state profile for the OH and OI emissions shows the height where the peak intensity of airglow emissions occurs.

The gravity wave simulation gives the volume emissions rate for every frame, and the frames are spaced at 30 second intervals. This project finds the volume emissions integrated over a time of 90 seconds for the OI emissions. To find the photon emissions that will be collected by an airglow camera, the emissions (ε_{OI}) must be integrated over time according to the following formulas:

$$\int_0^{90} \varepsilon_{OI}\, dt = \varepsilon_{collected} OI \tag{9}$$

$$\int_0^{30} \varepsilon_{OH}\, dt = \varepsilon_{collected} OH \tag{10}$$

The emissions in a vertical column of air for a period of 90 seconds may be quite different from a 30 second integration time because the integration time is approaching the period of very short period gravity waves.

Matlab can calculate a numerical integration using the trapezoid rule, but with very few data points, this integration is not very accurate. One solution was to interpolate four values of the photon emissions spanning a time period of 90 seconds. The interpolation was then evaluated at 20 points covering the full range of 90 seconds. With 20 points, the interpolated values could then be integrated numerically using the trapezoid rule to yield a more accurate value for the photon emissions.

Interpolation Methods:

Linear Interpolation: This method of interpolation is very simple. It consists of connecting each of the data points with lines. The resulting integral is essentially the same as using the trapezoid rule with just the three data points because the trapezoid rule simply forms trapezoids using the data points as corners and then calculates the enclosed areas.

Cubic Spline Interpolation: This method of interpolation finds the best fitting cubic polynomial in between points that fits certain parameters which are presented in Cheney and Kincaid[8]. The second derivative of the function must be continuous over the entire interval. The second derivatives of the first and last cubic polynomials must be equal to zero. The last constraint is what makes the function a natural cubic spline, and the function can be continued in either direction with a straight line.

Implementation:

Several different methods of interpolation were explored to find which was most accurate with only a few data points available. The method of testing the function to see how many initial data points are necessary to approach an error of zero was as follows: a set number of x values was used with equal spacing from 0 to π, and y=sin(x) was used for the y values. The sine function can describe most wave motions, and using an x range of 0 to π describes half of a wave period. Since a 90 second integration time with four data points was used

for a wave with a period of 4.4 minutes, this is approaching half of one period of the wave. It is therefore important that the interpolation be accurate for four or five points taken from sinusoidal data. Different interpolation methods were used to estimate the sine function, and then that function was evaluated over a set of 100 x values ranging from 0 to π. The trapezoid rule was then used to determine the integral of the interpolated sine function. This numerical integral was compared to the actual value of the sine integral from 0 to π, which is 2.

Figure 4 shows the plot of the error as a function of the number of x values used. The point where the difference between the actual integral and the interpolated numerical integral approaches zero was considered the point at which that specific interpolation method was considered to have enough points to be accurate. From Figure 4 it is apparent that cubic spline interpolation prior to integration using the trapezoid rule is a much better method, and it is relatively accurate for as few as 4 data points. For a period that spans 8 data points or more, using the trapezoid rule alone without interpolating first may be sufficient because Figure 4 shows that the linear interpolation begins to approach zero after 8 data points.

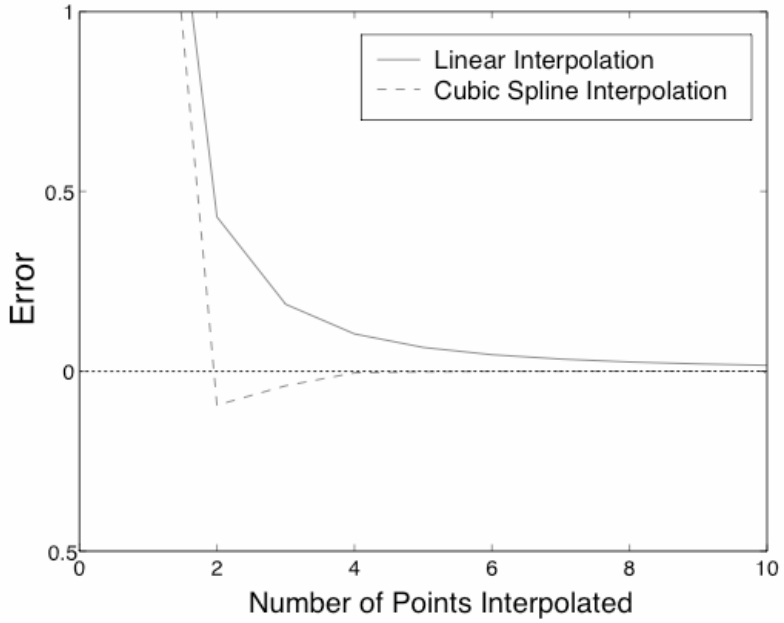


Figure 4. This graph shows the error versus number of data points two different interpolation methods.

RESULTS

The numerical model results used in this study are taken from Snively and Pasko[9]. These attempt to explain one aspect of an event reported by Taylor et al.[10] in which a 4.4 minute ducted gravity wave exhibited anti-phase perturbations in the OI and OH airglow layers. The simulated results depict a breaking gravity wave with period of 8.8 minutes exciting a ducted gravity wave

with period of 4.4 minutes. The OH and OI airglow emissions rates for the 4.4 minute wave are shown to be approximately 180 degrees out of phase due to delays in the OH excitation reaction. Here we study the effects of time integration of the OI emission to determine if other phase shifts can be introduced by the airglow imager.

The result of the integration of airglow emissions can be shown in a figure describing the photon emissions over a horizontal spread of 600 km (800 to 1400 km of the full simulation domain). In Figure 6 the wave has been propagating for 10,800 seconds. Figure 6a shows an image of the vertical scaled velocity over the horizontal spread, which shows the structure of the gravity wave. Figures 6b and 6d show the photon emissions over the camera integration times for the OI and OH emissions as would be seen from a ground-based imager that has an integration time of 90 seconds for the OI airglow and 30 seconds for the OH airglow. The OH photon emissions plot is simply the photon emissions rate, as calculated by the Matlab code, multiplied by 30 to get the total emissions for 30 seconds. Figures 6c and 6e show the photon emissions rate for OI and OH. Figure 7 shows the same plots as Figure 6 but for the time step where the wave has been propagating for 12,000 seconds.

For the OI plot, two different methods were used to get the 90 second integration. First, a trivial integration was done, as is shown by the dashed line, by multiplying three different time steps by 30 seconds and adding them together. This is the equivalent of having three different data points, drawing boxes around them and then adding the areas of the boxes, as is illustrated in Figure 5. The solid line is the time integration that is done by taking four different time steps that span a time of 90 seconds and taking a cubic spline interpolation of those four points and then integrating. The spline was evaluated at 20 different data points, and then the numerical integral was evaluated using the trapezoid rule. With 20 points, the integration should be a more accurate indication of the actual amount of photons that would be collected by an airglow imager. There doesn't appear to be much difference between the two different integration methods, but the difference becomes more apparent when the plots are enlarged. The cubic spline integration creates light and dark peaks with slightly higher magnitudes. The result of the trivial integration also seems to be phase shifted a few kilometers to the left of the cubic spline integration.

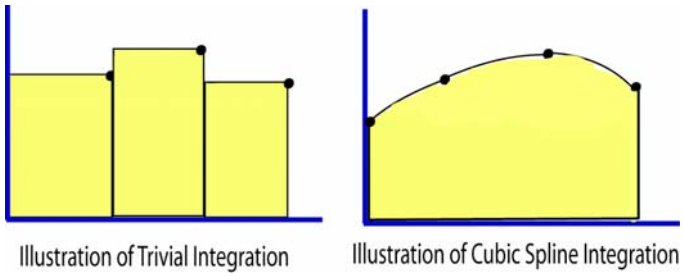


Figure 5. Illustration of the two different methods used to integrate over time for the OI emissions

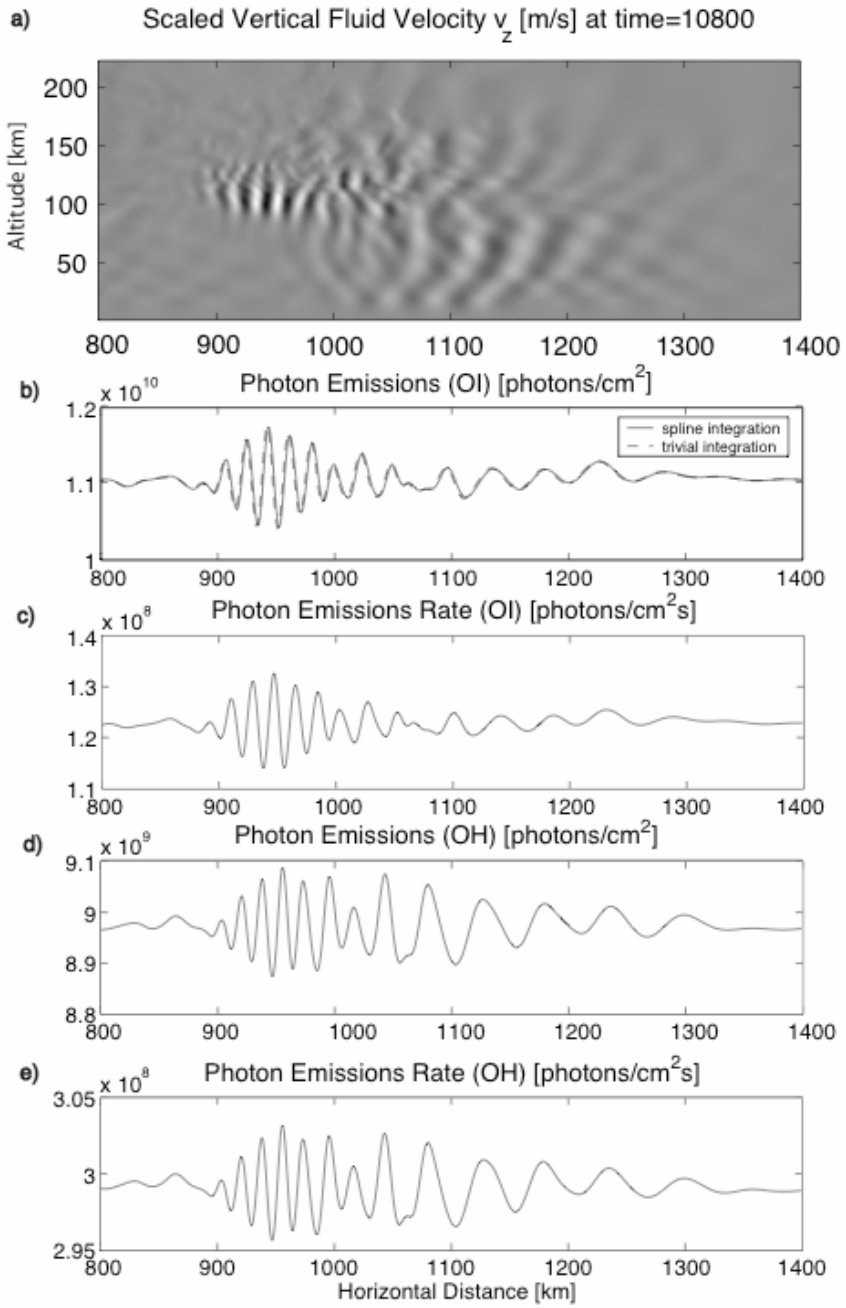


Figure 6. 6a) The vertical scaled velocity, representing gravity wave motion after the wave has been propagating for 10,800 s, with the light and dark regions showing the differences in velocity, 6b) Photon emissions for atomic oxygen with a 90 s integration time as they could be seen by a ground-based airglow imager, 6c) OI photon emissions rate, 6d) Photon emissions for hydroxyl over a 30 s integration time, 6e) OH photon emissions rate

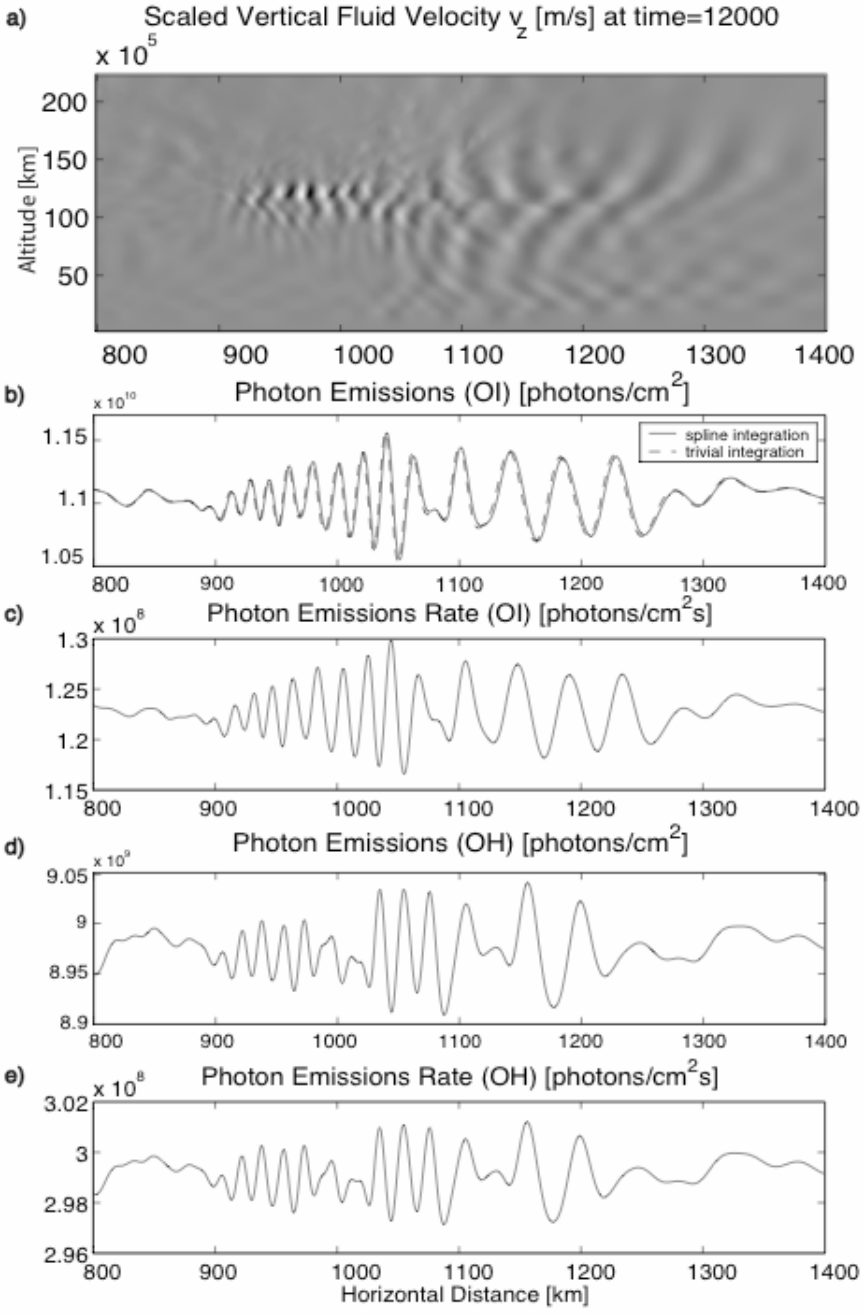


Figure 7. 7a) The vertical scaled velocity, when the wave has been propagating for 12,000 s, with the light and dark regions showing the differences in velocity, 6b) Photon emissions for atomic oxygen with a 90 s integration time, 6c) OI photon emissions rate, 6d) Photon emissions for hydroxyl over a 30 s integration time, 6e) OH photon emissions rate

DISCUSSION

The cubic spline interpolation method should produce a more accurate picture of the number of photons that would be seen by a ground-based airglow imager. The integration time does make a difference in the structure of the airglow that is imaged. The 90 s integration time is a significant portion of the period of the gravity wave. The photon emissions rates for OI and OH appear to be nearly anti-phase for the 10,800 s time step, however there is a slight phase shift in the integrated OI emission due to its reliance on the previous 90 seconds of data. At the 12,000 s time step, the integrated OI and OH emissions are not in phase but do not have a complete phase reversal. The phase variations between OI and OH emissions rates result from wave dynamics and chemistry, however additional phase variations in the total emissions arise due to the integration effects and non-simultaneous imaging of the airglow layers. So, because recorded OI and OH images must be taken sequentially rather than simultaneously, and because integration times vary between layers, phase variations between layers are somewhat harder to interpret for very short period waves.

One thing that would make the simulation different from what would be seen from a ground-based imager is the effect that the atmosphere would have on the airglow emissions. Some of the emissions will be absorbed by the atmosphere as it travels to the ground. The OI airglow emission peaks at an altitude of about 96 km and the OH airglow peaks around 87 km[5], at which heights the airglow must travel through most of the atmosphere before it reaches the ground. At the 557.7 nm OI emission wavelength, about 20% of the photons will be absorbed if they travel through the entire depth of the atmosphere[11]. About 10% of the near-infrared OH emissions will be absorbed by the atmosphere[11]. The only effect of this absorption by the atmosphere would be to decrease the photon count, but since the emission loss would be uniform, the structure of the airglow images should be unaffected.

CONCLUSION

It is important to consider the camera integration time when studying the photon emissions given by the simulation. Due to camera integration times, a slight phase shift occurs in addition to those caused by wave dynamics or chemistry alone. Although the overall effect may be small for longer period waves, integration times along with the total time taken between images of separate airglow layers should be taken into consideration when studying short period gravity waves.

ACKNOWLEDGMENTS

I would like to thank Jonathan Snively for teaching me everything I know about gravity waves and for his guidance throughout the whole process. I would also like to thank Dr. Victor Pasko for giving me the opportunity to work on this project. This material is based upon work supported by the National Science

Foundation under grants EEC-0244030 and ATM-01-23020 to Pennsylvania State University

REFERENCES

[1] Pierce, A. D. and S. C. Coroniti, "A Mechanism for the Generation of Acoustic-Gravity Waves During Thunderstorm Formation," *Nature*, 210, 1209-1210, 1966.

[2] Taylor, L. L., *Mesospheric Heating Due to Intense Tropospheric Convection*, NASA Contractor Report 3132, 1979.

[3] Snively, J. B., *Tropospheric Forcing as a Source of Quasi-Monochromatic Short-Period Gravity Waves Observed in the Upper Mesosphere and Lower Thermosphere*, MS Thesis, 19-22, 2003.

[4] Taylor, M. J. and M. A. Hapgood, "Identification of a thunderstorm as a source of short period gravity waves in the upper atmospheric nightglow emissions," *Planet. Space Sci.*, 36 (10), 975-985, 1988.

[5] McDade, I .C., E. J. Llewellyn, D. P. Murtagh, and R. G. H. Greer, "Eton 5: Simultaneous Rocket Measurements of the OH Meinel $\Delta v=2$ Sequence and (8,3) Band Emission Profiles in the Nightglow," *Planet. Space Sci*., 35 (9), 1137-1147, 1987.

[6] Bates, D. R. "The Green Light of the Night Sky," *Planet. Space Sci*., 29 (10), 1061-1067, 1981.

[7] Taylor, M. J. and M. B. Bishop, "All-sky measurements of short period waves imaged in the OI (557.7 nm), Na (589.2 nm) and near infrared OH and $O_2(0,1)$ nightglow emissions during the ALOHA-93 campaign," *Geophys. Res. Lett*., 22 (20), 2833-2836, 1995.

[8] Cheney, Ward and David Kincaid, "Natural Cubic Splines," 330-332, *Numerical Mathematic and Computing*. 4th ed., Brooks/Cole, Pacific Grove, 1999.

[9] Snively, J. B. and V. P. Pasko, "Thermal Ducting and the ALOHA-93 Spectacular Gravity Wave Event," CEDAR Workshop 2004, *Poster Session Booklet*, 43, Santa Fe, NM, July 2004.

[10] Taylor, M. J., D. N. Turnbull, and R. P. Lowe, "Spectrometric and Imaging Measurements of a Spectacular Gravity Wave Event Observed During the ALOHA-93 Campaign," *Geophys. Res. Lett*., 22 (20), 2849-2852, 1995.

[11] Gary, B. L. "All Sky Photometry Tutorial," [online resource] http://reductionism.net.seanic.net/CCD_TE/AllSkyPhotometry.html

WAVELET ANALYSIS OF TURBULENCE IN THE MESOSPHERE

Alireza Masnadi-Shirazi*, Charles L. Croskey[#], and John D. Mitchell[#]

Department of Electrical Engineering
The Pennsylvania State University
University Park, PA 16802

*Undergraduate student of
Department of Electrical Engineering
University of Texas at Arlington, Arlington, TX 76013

Abstract

During the past several decades considerable progress has been made concerning the efforts to characterize the presence and force of turbulence in the mesosphere. While ground based studies using radars were successful providing a preliminary qualitative overview of turbulent activity in the upper middle atmosphere, the deployment of in situ sounding rocket experiments has provided us with significantly more quantitative measurements of turbulent eddy motions.

In an effort to study the global atmosphere and weather we consider the analysis of mesospheric turbulence measured by rocket borne ionization instruments. The standard Short-Time Fourier approach has already been used to process the existing data that would give us information about the frequency features vs. altitude of turbulent regions in the mesosphere. However, wavelet analysis is introduced as an alternative suitable tool to derive spectral information of the data. Due to wavelet analysis ability to isolate short-duration processes embedded in highly-structured harmonic phenomena, it is believed that it might provide us with additional information on the actual vertical extent of the observed turbulent structures.

[#] Faculty Mentor

Introduction

Turbulence in the upper atmosphere plays a vital role in the atmosphere's energy budget. Turbulence both heats the atmosphere by dissipation of kinetic energy and cools it at the same time by heat conduction, and it is not clear which effect is stronger, heating or cooling.[1] In an effort to unravel the ambiguities regarding mesospheric turbulence, many attempts have been made in the recent years to measure turbulence in the upper atmosphere and, more specifically, to deduce geophysically relevant turbulence parameters such as the dissipation rate of turbulent energy, ϵ. Both remote and in situ techniques have been used. Compared to remote sensing methods, in situ measurements of rocket born ionization tools have the advantage of measuring a spectrum of spatial fluctuations and not simply one scale of the spectrum. The spectrum of the fluctuations can then be studied to decide whether or not the conditions for turbulence are satisfied.[2]

Fritts et al. [3] have shown that both gravity wave breaking events and Kelvin Helmholtz instabilities are expected to produce sharp local variations in turbulent kinetic energy generation and dissipation. Currently utilized techniques to analyze and derive spatial spectrums from in situ measurements, however, rely on the Short-Time Fourier analysis with a windowing resolution of about 1 km which is known to smear out any existent sharp changes. Thus, the Short-Time Fourier method does not allow extracting the actual vertical extent of the turbulent regions, a quantity that is crucial for the derivation of the geophysically relevant turbulence parameters.[2]

In this paper we present a relatively new method of wavelets for analyzing turbulence data gathered from in situ measurements of the electron density fluctuations in the mesosphere. Wavelet analysis is similar to the Short-Time Fourier analysis, but it uses an adaptive windowing technique, incorporating a narrower window for analyzing sharp local changes and a wider window for analyzing gradual changes. This approach allows a higher resolution in the time-frequency spectrum. Wavelet analysis will be used as a more accurate way to analyze electron density fluctuations caused by turbulence without the disadvantage of loosing important information on the vertical extent of turbulent structures. This technique is applied to the data from a sounding rocket named MMMW-23 of the international MACWAVE-campaign, launched from Andøya, Norway at 00:47:00 UT on July 5, 2002. We compare the results with the Short-Time Fourier technique and demonstrate that the wavelet method indeed reveals features not obtained from the Fourier approach.

Experimental Description

The Short-Time Fourier Transform (STFT) is used to analyze only a small section of the signal at a time; it involves a technique called windowing the signal, which maps a signal into a two-dimensional function of time and

frequency. The STFT represents a sort of compromise between the individual time- and frequency-based views of a signal. It provides some information about both when and at what frequencies a signal event occurs. However, one can only obtain this information with limited precision, and that precision is determined by the size of the window. While the STFT compromise between time and frequency information can be useful, the drawback is that once you choose a particular size for the time window, that window is the same for all frequencies. Many signals require a more flexible approach: one where we can vary the window size to determine more accurately either time or frequency.

Wavelet Analysis represents the next logical step: a windowing technique with variable-sized regions. Wavelet analysis allows the use of long time intervals where we want more precise low-frequency information, and shorter regions where we want high-frequency information. A wavelet is a waveform that is effectively localized in both time and frequency space and has an average value of zero. So far, many wavelets have been made and each has its own characteristics and waveform. Fourier analysis is the breaking up of a signal into sine and cosine components with different frequencies. Similarly, wavelet analysis is the breaking up of a signal into shifted and scaled versions of the original wavelet (mother wavelet). Figure 1 illustrates the steps taken to perform the 1-D wavelet transform on a signal. The wavelet transform involves plotting the coefficients of the wavelet transform versus time and scale.

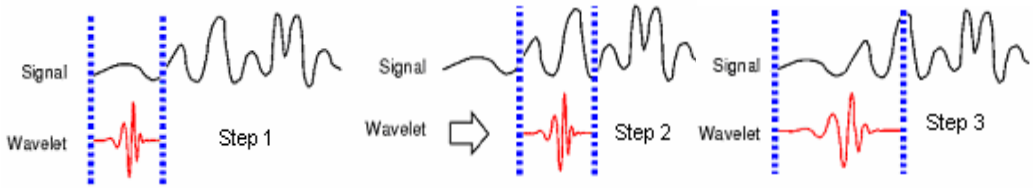


Figure 1. Wavelet analysis Steps: The correlation factor between the mother wavelet and part of the signal in the wavelet window is first computed, then the signal is continuously shifted to the right and the same thing is repeated. The mother wavelet is then scaled and steps 1 and 2 are repeated to the end. Then, Steps 1 through 3 are repeated for other scales. [9]

In recent years the wavelet technique has been used in numerous geophysical studies seeking to identify localized power events in time series. In our study, we apply the complex Morlet-wavelet consisting of a complex sinusoidal wave modulated by a Gaussian:

$$\psi_0(\eta) = \pi^{-1/4} e^{i\omega_0 \eta} e^{-\eta^2/2} \qquad (1)$$

where ω_0 is a nondimensional angular frequency of the mother wavelet (we have chosen $\omega_0 = 2\pi f_0 = 2\pi$, where f_0=1 is the mother wavelet frequency of oscillations) and η is a nondimensional time parameter. Figure 2 shows the real and imaginary parts of this mother wavelet and its corresponding complex sinusoidal wave.

The continuous wavelet transform of a discrete time series signal x_n (with n=0 to N-1 being the index of the elements in the time series) is defined as: [4]

$$W_n(s) = \sum_{n'=0}^{N-1} x_{n'} \psi^* \left[\frac{(n'-n)\delta t}{s} \right] \qquad (2)$$

where $(*)$ indicates the complex conjugate, s is the wavelet scale, δt is the sampling time interval of the time series, and n is the localized time index.

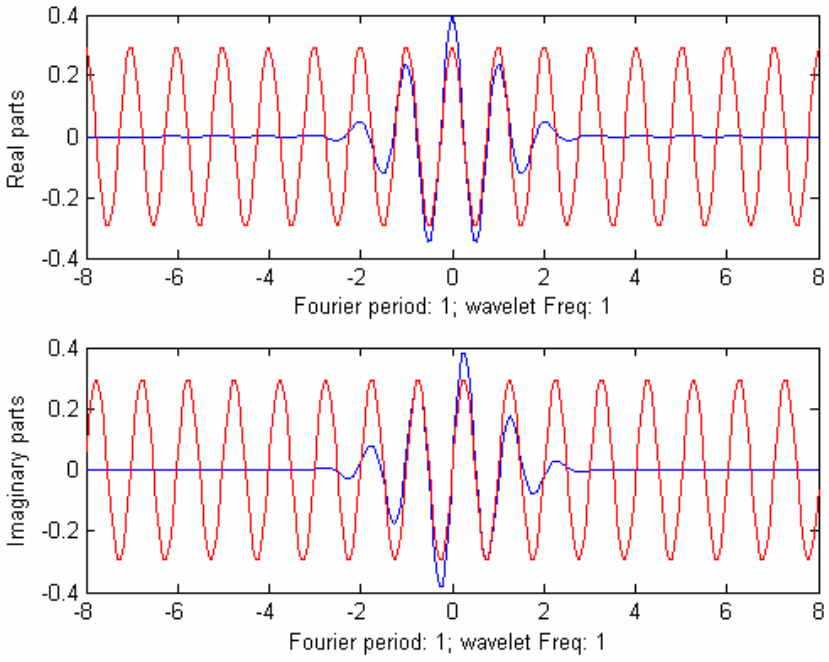


Figure 2. Real and imaginary parts of the complex Morlet mother wavelet (blue) and its corresponding sinusoidal wave (red), with $f_0 = 1$.

To achieve an equal spacing between wavelet transform coefficients versus scale in a Bode diagram, the wavelet scales are constructed as fractional powers of two:

$$s_j = s_0 2^{j\delta j}, \quad j = 0, 1, \ldots, J \tag{3.1}$$

$$J = \delta j^{-1} \log_2(N\delta t / s_0) \tag{3.2}$$

where s_0 is the smallest resolvable scale, given by $2\delta t$, δj is relative to the spacing between scales in logarithmic units (we have chosen δj=1/4), and *N* is the number of elements in the time series to be analyzed. Hence, by varying *s* and *n* (i.e., varying frequency and time), one can construct a 3-dimentional plot showing the amplitudes of any feature versus the scale (related to the frequency) and how this amplitude varies with time. [4]

Because we want to compare our results to the Fourier method results, we would like to also convert the scales used in the wavelet transform to exact Fourier frequencies used in the Fourier method. For the complex Morlet wavelet, ω_0 and *s* are related to the Fourier Frequency f by:

$$f = \frac{\omega_0 + \sqrt{2 + \omega_0^2}}{4\pi s} \tag{4}$$

Having known that $\omega_0 = 2\pi f_0 = 2\pi$, the relationship between Fourier frequency and wavelet scale comes out to be $f = 1.0125/s$.

The main aim of this analysis is to ascertain the existence of turbulence in the mesosphere from the data gathered by rocket borne ionization tools. In order to do so, we will compare our results to the mathematical model of turbulence (i.e., Heisenberg's theory of Turbulence[5]). Once the existence of turbulence has been determined, geophysically relevant parameters, for example the energy dissipation rate ϵ, can be extracted by fitting the theoretical spectral model to the calculated turbulent spectrum. Heisenberg's theoretical model on turbulence states a -5/3 slope in the inertial sub range and then a smooth transition to a -7 slope in the viscous subrange for the Fourier Bode power plot in a given altitude range in which turbulence is present.

Results and Discussion

Figure 3 illustrates the relative electron density fluctuations with respect to Altitude obtained from the upleg flight of the MMMW-23 payload.

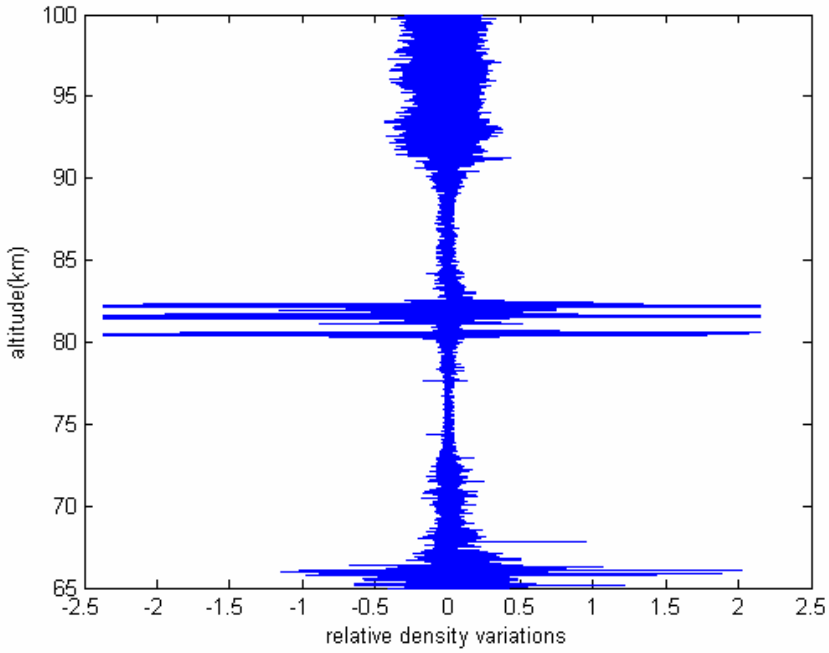


Figure 3. Relative electron density variations of the MMMW-23 rocket payload ranging from 65 to 100 km in altitude.

Before applying the wavelet transform to the raw data, some preliminary processing of the data should be done. Due to the rocket rotation at a constant rate, a residual AC signature at approximately 5.3 Hz is picked up by the measuring probes, thereby, creating sharp spikes in the Fourier spectrum at 5.3 Hz and its harmonics. As these roll-produced harmonics can cause some artifacts in the wavelet analysis, we have utilized some consecutive notching filters to remove these undesired frequencies. Figure 4 shows a sector of the Fourier spectrum before and after notching.

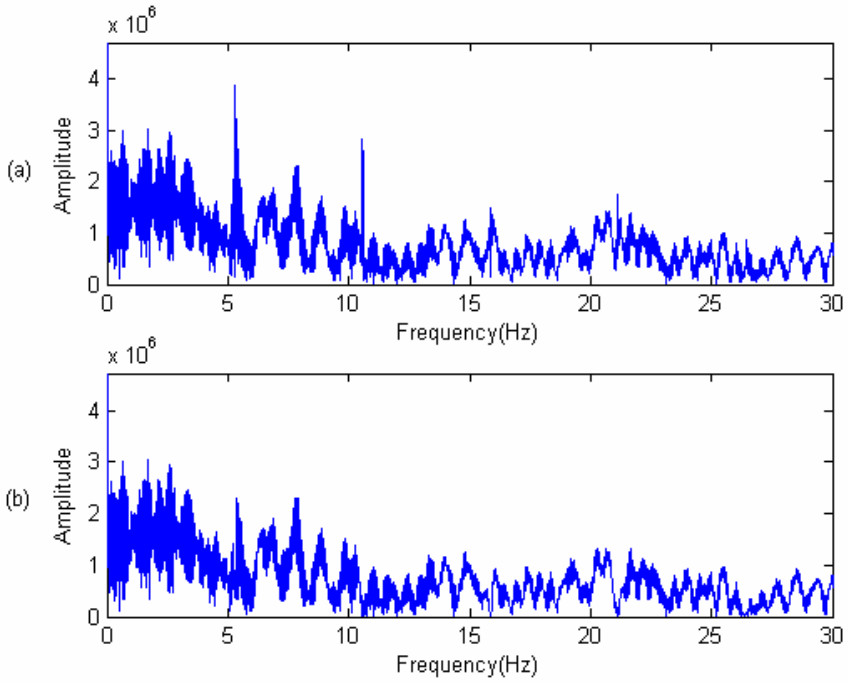


Figure 4. Sections of the Fourier spectrums of the relative electron density fluctuations: (a) before applying notch filters; (b) after applying notch filters.

Because the rocket electronics measuring the electron density fluctuations have a certain measuring frequency bandwidth between 20 to 600 Hz, an inverse filter is utilized to compensate the attenuated frequency components in the band reject region. This filter is applied to the wavelet coefficients and results in the amplification of the frequencies in the band reject space. In Figure 5 we present the 3-D Bode plot of the wavelet power spectrum (dB) of the measured electron air density variations after notching and inverse filtering for an altitude range between 75 and 90 km. Note that equation 4 is used to convert the wavelet scales into exact Fourier frequency values. Figure 5 is a composite of approximately 90000 individual wavelet power spectra (our number of data points in this altitude region). Each of these wavelet power spectra can be studied to see if there are any turbulent signature features that follow a -5/3 slope (-16.7 *dB/dec)* in the inertial subrange and a -7 slope (-70 *dB/dec*) in the viscous subrange[5]. However, due to the fact that turbulent structures are formed on a vertical extent (known as the inner scale) of between 20-50 m[6], each 500 of these wavelet spectra will be averaged, thereby defining an altitude resolution of approximately 100 m.

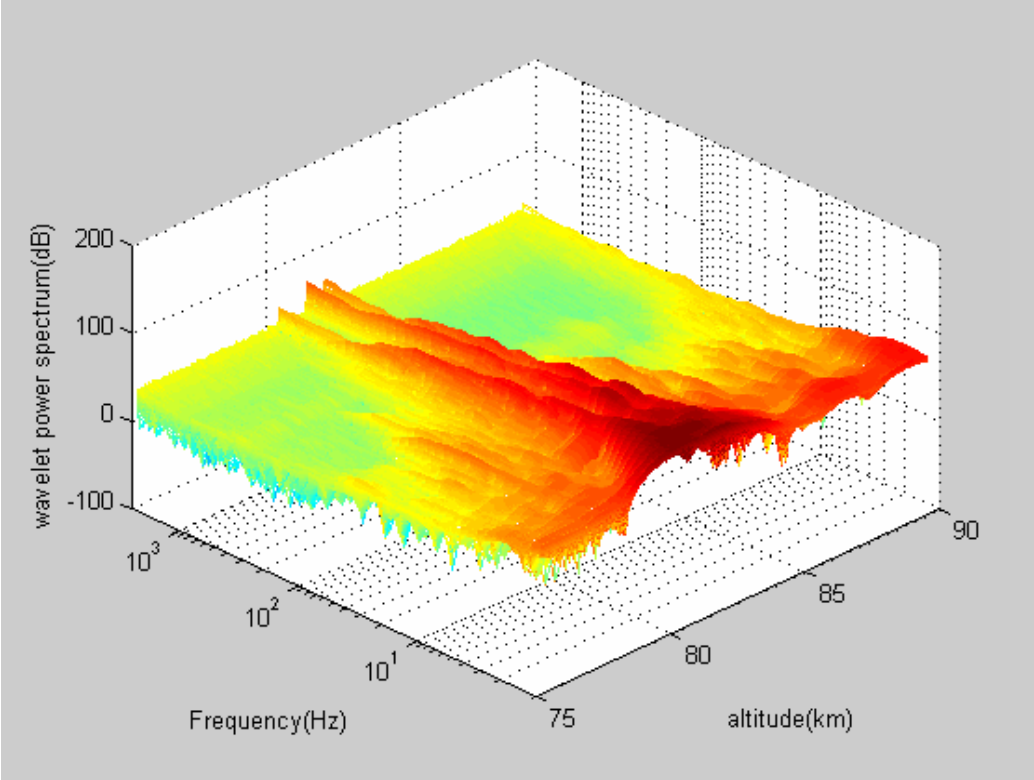


Figure 5. The 3-D Bode plot of wavelet power spectrum (dB) of the measured electron air density variations after notching and inverse filtering. The flat plane marks the noise level of our measuring instrument.

Some altitude regions indicating a -5/3 and a -7 slope in the wavelet power spectra have been selected. These regions turn out to be roughly around 77, 82 and 88 km for the MMMW-23 rocket measuring the electron density variations. Above 92 km, as it can be seen in Figure 3, some sharp local fluctuations have been measured, which are known to be caused by the presence of plasma.[7] In Figures 6a-c we illustrate three altitude regions that match the *Heisenberg* model of turbulence. Figure 6d (taken with permission from Croskey et al.[7]) shows the relative power density using the Short-Time Fourier analysis and is displayed for an indication of consistent results between the wavelet technique and the Fourier technique. Figure 7 demonstrates two altitude regions that do not contain the -5/3 and -7 slopes.

At altitude regions between 80 and 85 km we have attained some results indicating a reasonable fit between the wavelet power spectrum and the *Heisenberg* model. However, due to the large transition frequency from the inertial subrange to the viscous subrange, these results seem to be unrealistic. This high transition frequency is mainly due to the damping effect that some charged

ice particles present in this altitude region have on the electron movement[8]. An example of such a region is shown in Figure 6b.

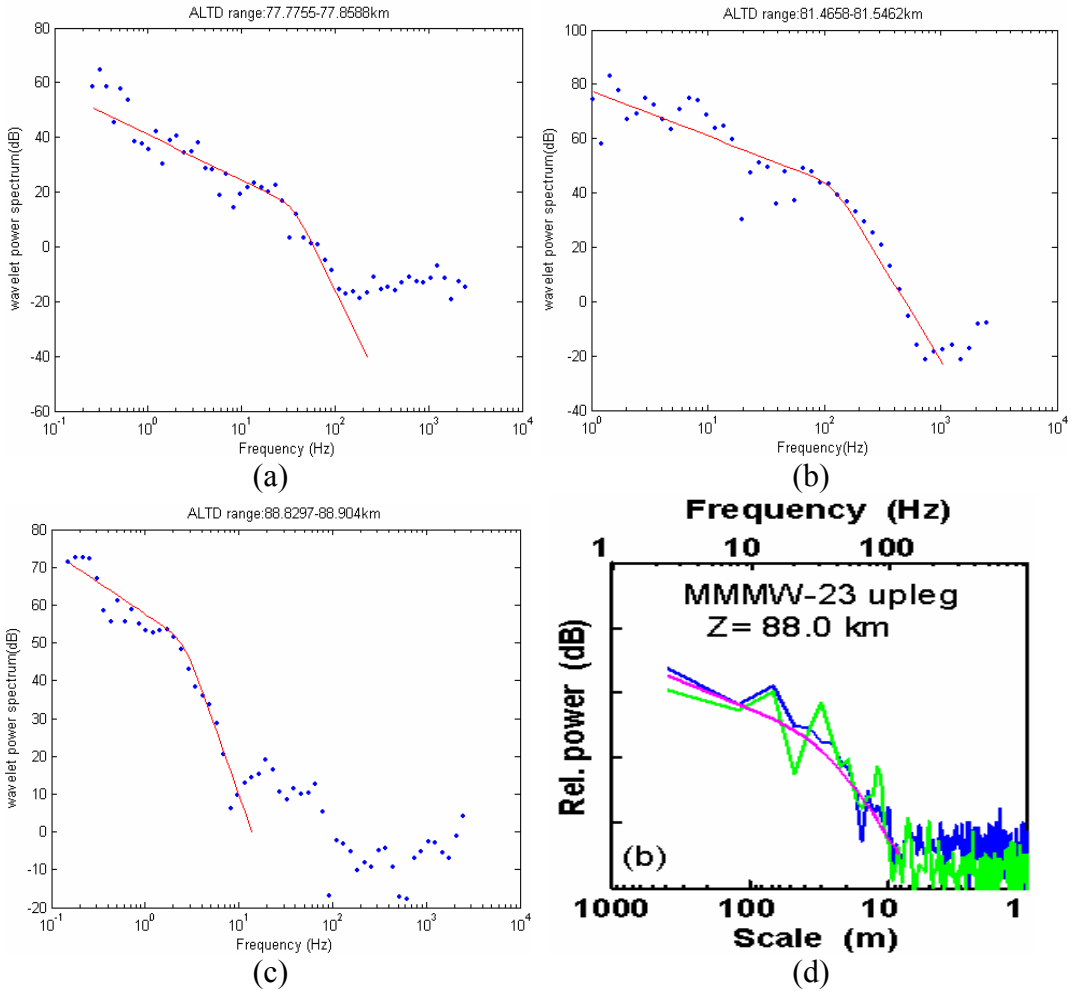


Figure 6. Power spectral density plots with comparative Heisenberg fits (straight lines) at different altitudes: (a) wavelet power spectrum at altitude range 77.77-77.86 km with transition frequency around 30 Hz; (b) wavelet power spectrum at altitude range 81.47-81.55 km with transition frequency around 100 Hz (ice particles are present in this altitude range resulting in a high transition frequency); (c) wavelet power spectrum at altitude range 88.83-88.9 km with transition frequency around 3 Hz; (d) relative power density using the Short-Time Fourier analysis shows consistent results with Figure 6c on the turbulence detection aspect. [7]

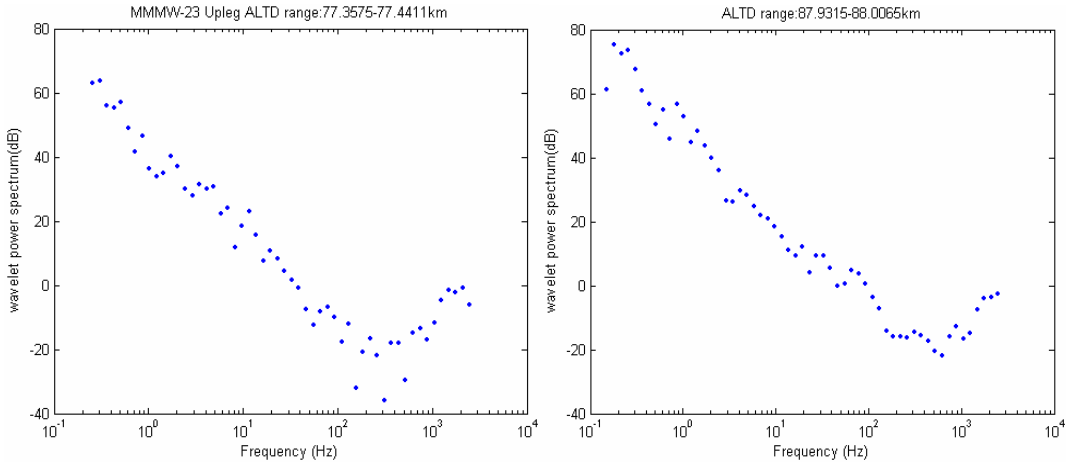


Figure 7. Wavelet power spectrums that do not correspond to turbulent activities; (a) at altitude range of 77.36-77.44 km; Note that in less than 0.5 km above this altitude (Figure 6a) turbulence is present; (b) at altitude range of 87.93-88 km; Note that in less than 1 km below this altitude range (Figure 6c) turbulence is present.

One might argue that wavelet analysis introduces some artifacts due to the properties regarding the specific wavelet that is being used, hence, not being able to correctly detect slopes which are present in a Fourier power spectrum. To study this case we have implemented our wavelet analysis on colored noise with known corner frequency and filter degree (representing the slope in the Fourier Bode power spectrum), in order to observe any consistency with the Fourier Power Spectrum. These results, shown in Figure 8, are an indication of the complete consistency between the Fourier-based method and the wavelet-based method.

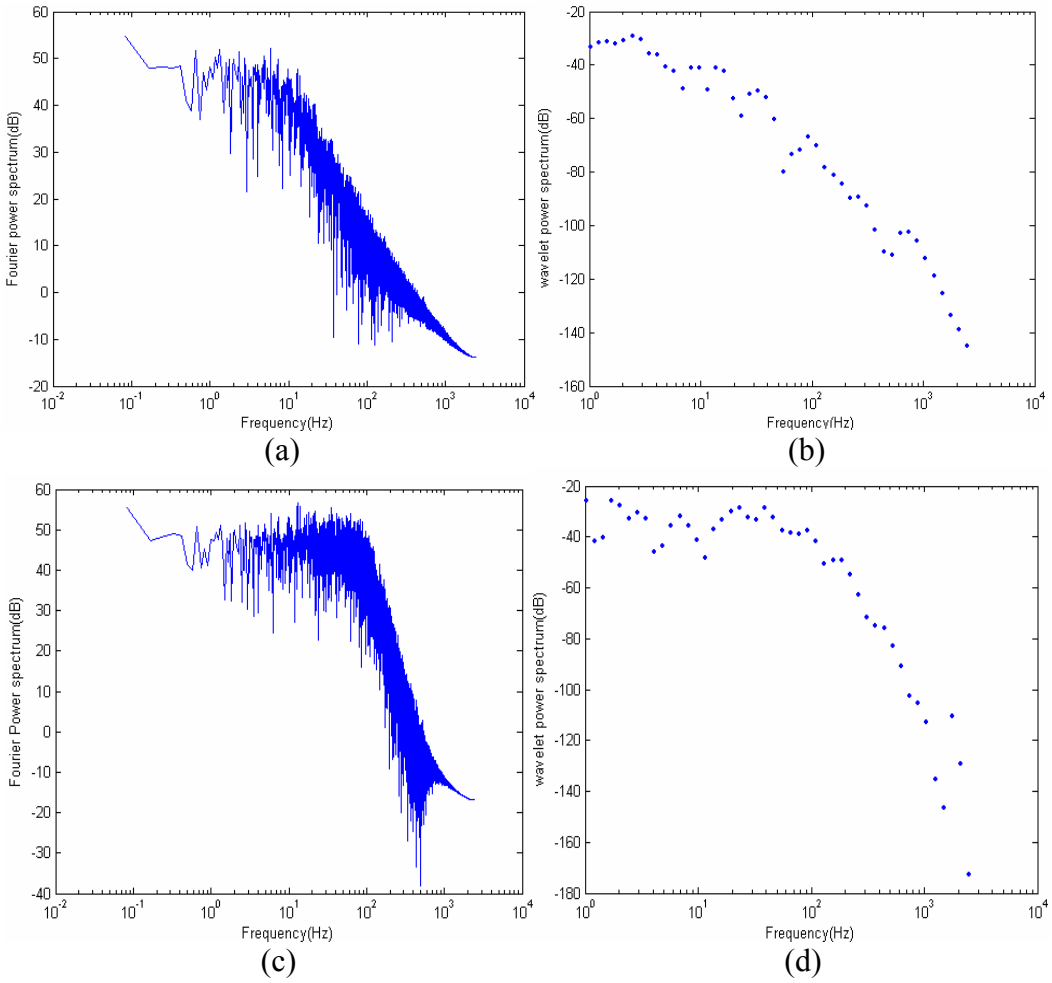


Figure 8. Fourier and wavelet power spectrums for colored noise: (a) Fourier power spectrum of a white noise passed through a first order butterworth low pass filter with corner frequency of 10 Hz; (b) wavelet power spectrum of a white noise passed through a first order butterworth low pass filter with corner frequency of 10 Hz; (c) Fourier power spectrum of a white noise passed through a second order butterworth low pass filter with corner frequency of 100 Hz; (d) wavelet power spectrum of a white noise passed through a second order butterworth low pass filter with corner frequency of 100 Hz and.

Summary

We have introduced the wavelet technique as a relatively new method for processing electron density fluctuations measured by in situ rocket probes. Unlike

the traditional Fourier-based method, wavelet analysis can result in a finer altitude resolution of about 100 m. Having compared the results from the standard Fourier method and the wavelet method, we find general agreement between both techniques. However, we were able to extract turbulent signature features not seen using the Fourier method. This was because the Fourier method bears the risk of smearing out the turbulence pattern features through an altitude range of about 1 km. As a final check of the consistency between the Fourier method and wavelet method, an analysis on colored noise was done. This investigation verified that the wavelet transform can indeed identify the spectral slopes present in the Fourier transform.

Acknowledgement

This material is based upon work supported by the national science foundation Grant No. EEC-0244030.

References

[1] F.-J. Lubken, "On the extraction of turbulent parameters from atmospheric density fluctuations," *Journal of Geophysical Research*, **97** (20) 385-95 (1992).

[2] B. Strelnikov, M. Rapp and F.-J. Lubken, "A new technique for the analysis of neutral air density fluctuations measured in situ in the middle atmosphere," *Geophysical Research Letters*, **30** (20) ASC 7-1 to 7-5 (2003).

[3] D. Fritts, C. Bizon, J. Werne and C. meyer, "Layering accompanying turbulence generation due to shear instability and gravity-wave breaking," *Journal of Geophysical Research*, **108** (8) 8452 (2003).

[4] C. Torrence and G.P. Compo, "A practical guide to wavelet analysis," *Bulletin of theAmerican Meteorological Society*, **79** 61-78 (1998).

[5] W. Heisenberg, "Zur statisischen Theorie der Turbulenz," *Z. Physik*, **124** 628-57 (1948)

[6] F.-J. Lubken, "seasonal variations of turbulent energy dissipation rates at high latitudes as determined by in situ measurements of neutral density fluctuations," *Journal of Geophysical Research*, **102** (13) 441-56 (2003).

[7] C.L. Croskey, J.D. Mitchell, R.A. Goldberg, T. A. Blix, M. Rapp, R. Latteck, M. Friedrich and B. Smiley, "Coordinated investigation of plasma and neutral density fluctuations and particles during the MacWAVE/MIDAS summer 2002 program," *Geophysical Research letters*, **31** In Press (2004)

[8] M.C. Kelly, D.T. Farely and J. Röttger, "The effect of cluster ions and anomalous VHF backscatter from the summer polar mesosphere," *Geophysical Research letters*, **14** 1031-34 (1987).

[9] Matlab Documentation, "W*avelet Toolbox*," Mathworks Inc., 2003. http://www.mathworks.com/access/helpdesk/help/toolbox/wavelet/wavelet.shtml

EFFECTS OF SOLAR FLARES ON THE FIRST SCHUMANN RESONANCE FREQUENCY

John McGlade, Heng Yang[+], and Victor P. Pasko[#]

Communications and Space Sciences Laboratory (CSSL)
Department of Electrical Engineering
The Pennsylvania State Unversity
University Park, PA 16802

Abstract

Variations in the exact Schumann Resonance frequencies occur due to changes in the conductivity profile of the Earth-ionosphere cavity. Aside from seasonal and diurnal cycles, these changes are also caused by factors such as X-ray bursts and high-energy particle precipitation from the sun[1]. The discrete data we get from an FDTD model[2] does not allow us to obtain the Schumann Resonance frequencies directly, and the resolution in the frequency domain is contingent upon sampling time. In this paper, the method of choice in determining the Schumann Resonance frequencies is through exponential approximation. Exponential approximation is effective because it can detect the slight shift in Schumann Resonance frequencies based on a short sampling time, which is desirable in order to reduce computer calculation time. In this paper we use the Prony Method of Exponential Approximation[3] to find this slight frequency shift in a short sampling time.

[#] Faculty Mentor
[+] Graduate Mentor

Introduction

At altitudes above approximately 30 km, collisions between cosmic rays/solar radiation and neutral atmospheric particles establish a region of significant ion/electron densities. This region, called the ionosphere, acts as a good, but imperfect conductor of electromagnetic waves. The surface of the Earth is also a good conductor; thus, the space between the Earth and the ionosphere serves as a waveguide and is aptly named the Earth-ionosphere waveguide*. The region below approximately 80 km (D through lower E region) is of particular interest in this study because the EM properties of the Earth-ionosphere cavity are determined by the conductivity profile in this region. *Schumann*[4] first predicted the Earth-ionosphere cavity could possess resonance properties, and in subsequent years they were experimentally detected[5]. These resonances are thus referred to as Schumann Resonances, and comprise most of the natural background EM spectrum between 6 and 50 Hz[6]. The first three intensified amplitudes fluctuate around approximately 8, 14, and 20 Hz.

Global lightning activity is the biggest contributor of natural ELF electromagnetic signals†, and this occurs at a rate of approximately 100 strokes per second[6]. For typical lightning storms, individual lightning strokes are not discernable from Schumann Resonance data unless the storms are close to the observation site. However, the intensity produced by sprite inducing lightning is much larger than the typical ELF band intensity. Thus, researchers are able to remotely sense large lightning events that are often precursors of sprites, jets, and elves[7].

Schumann resonances vary not only with internal sources but also with propagation factors; hence, Schumann resonances are a good tool for monitoring both Global lightning and ionospheric properties[8, 9]. Directly monitoring the lower ionosphere is difficult from ground sites and from satellites, so a perfected system based on Schumann Resonances would be an effective solution. Yet another reason for studying this phenomenon is that SR amplitude follows closely with tropical temperature variations as well, and so we can monitor Global Tropical Temperatures[10].

The purpose of this study is to examine the effects of conductivity changes in the ionosphere. Diurnal and seasonal cycles both contribute to the fluctuation of SR frequencies[5, 11], but more recent research has led teams to observed slight frequency shifts during solar flares. This phenomenon can be attributed to perturbations in conductivity[12]. Depending upon the altitude of the perturbation, frequencies can increase or decrease. Using a model generated by the FDTD method, this paper establishes the reliability of the Prony algorithm in determining Schumann resonance frequencies, and compares the data results from conductivity

---

* This waveguide is also referred to as a cavity resonator: completely enclosed with resonant electric and magnetic fields.

† Commonly known as “spherics”.

perturbations at different altitudes to find the Schumann frequency variations. We can then extrapolate how X-ray bursts and Solar Proton Events (SPEs) change SR frequencies.

Theory:

The simplest model to determine Schumann Resonant frequency depicts the ionosphere and the surface of the Earth as perfectly conducting spheres separated by free space, and yields angular frequencies

$$\omega_n = c/a * \sqrt{(n(n+1)(1 - h/a))} \quad (1)$$

where n is the mode number, c is the speed of light, a is the radius of the Earth, and h is the distance between the Earth and the ionosphere[4]. In reality, however, the mode frequencies are lower than the predicted values. The discrepancy is due to the fact that the ionosphere cannot be accurately modeled as a perfect conductor. ELF waves dissipate energy and shift in frequency when traveling in the ionosphere, and so the ionosphere's content determines the exact Schumann Resonant frequencies at any given time. An appropriate model of the ionosphere's conductivity was proposed by *Sentman*[13] and uses an exponential equation to model the profile at any altitude.

$$\sigma(z) = 10^{-16} * \exp\left[z/3.1 + 2.303\beta*\exp-((z-z_0)/a)^2\right] \quad (2)$$

where β = amplitude of the perturbation (increase/decrease in conductivity in pwrs of 10), z_0 = altitude of the perturbation (km), and a = width of the perturbation (km). With this model, we can simulate an increase or decrease in conductivity at any altitude. Figure (1) graphs the exponential model. The blue line signifies the conductivity profile in an unperturbed state. Galactic Cosmic Rays and soft X-rays are contributors to the exponential conductivity profile. Perturbations of increasing conductivity with ß = 3 are simulated at altitudes of 40 km and 70 km to mimic Solar Proton Events and X-ray bursts, respectively.

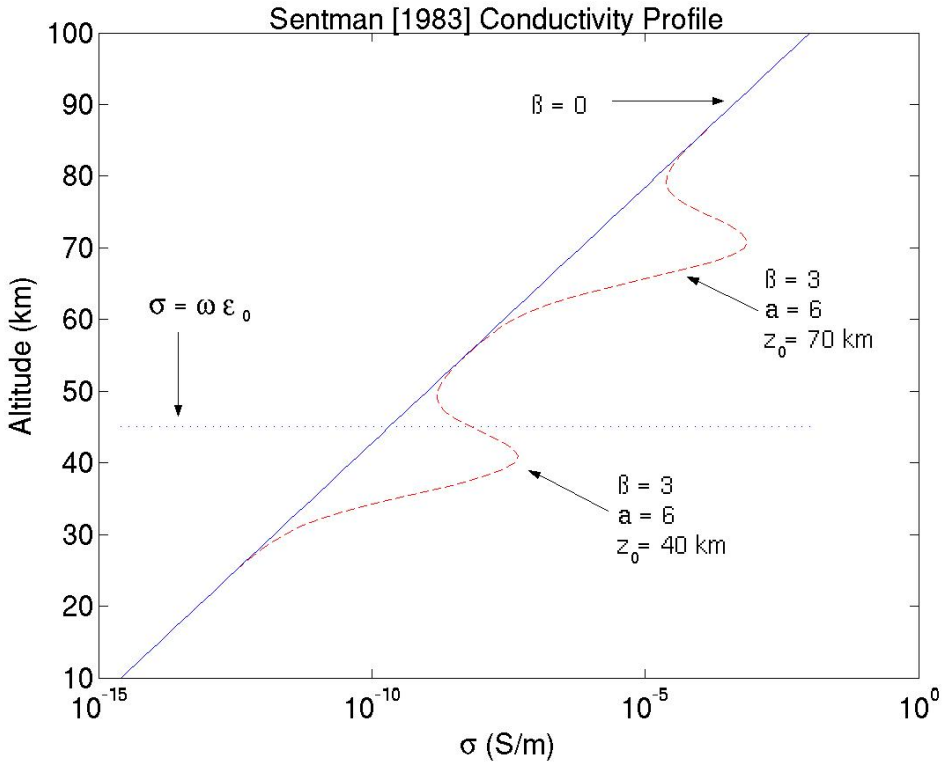


Figure 1: [Sentman, 1983] conductivity model using an exponential equation with a perturbation parameter, ß. The line $\sigma = \omega\varepsilon_0$ signifies the altitude in which the ionosphere becomes conductive.

Experimental Description:

Discrete data from an FDTD model does not allow us to get the Schumann Resonance frequencies directly, and the data is composed of many damped harmonics of unknown periods. So we employed complex exponential functions and used Prony's method to determine the first Schumann Resonant frequency and its Q value. To account for dissipation effects[14], Schumann resonances are modeled as standing waves with a damping parameter φ and angular frequency ω' such that:

$$\omega = \varphi + i\omega' \qquad (3)$$

The Prony method is detailed as follows:

$$f(t) = C_1*\exp(\omega_1 t) + C_2*\exp(\omega_2 t) + C_3*\exp(\omega_3 t) + \ldots \qquad (4)$$

Or, if $\mu_i = \exp(\omega_i)$ where i = 1, 2, 3,…

$$f(t) = C_1 * \mu_1^{\,t} + C_2 * \mu_2^{\,t} + C_3 * \mu_3^{\,t} + C_4 * \mu_4^{\,t} + \ldots \qquad (5)$$

A fundamental property of this method is that t increments at equally spaced points. This is essential for the model to then satisfy difference equations with constant coefficients[15]. For example, with t equal to integer values of 1, such that t = 0, 1, 2, 3, 4, …, N –1 (where N equals the number of sample points), we can generate a list of equations of the form:

$$
\begin{aligned}
f_0 &= C_1 + C_2 + C_3 + C_4 + \ldots + C_n \\
f_1 &= C_1 * \mu_1^{\,1} + C_2 * \mu_2^{\,1} + C_3 * \mu_3^{\,1} + C_4 * \mu_4^{\,1} + \ldots + C_n * \mu_n^{\,1} \\
f_2 &= C_1 * \mu_1^{\,2} + C_2 * \mu_2^{\,2} + C_3 * \mu_3^{\,2} + C_4 * \mu_4^{\,2} + \ldots + C_n * \mu_n^{\,2} \\
f_3 &= C_1 * \mu_1^{\,3} + C_2 * \mu_2^{\,3} + C_3 * \mu_3^{\,3} + C_4 * \mu_4^{\,3} + \ldots + C_n * \mu_n^{\,3} \\
f_4 &= C_1 * \mu_1^{\,4} + C_2 * \mu_2^{\,4} + C_3 * \mu_3^{\,4} + C_4 * \mu_4^{\,4} + \ldots + C_n * \mu_n^{\,4} \\
f_5 &= C_1 * \mu_1^{\,5} + C_2 * \mu_2^{\,5} + C_3 * \mu_3^{\,5} + C_4 * \mu_4^{\,5} + \ldots + C_n * \mu_n^{\,5} \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
f_{N-1} &= C_1 * \mu_1^{\,N-1} + C_2 * \mu_2^{\,N-1} + C_3 * \mu_3^{\,N-1} + C_4 * \mu_4^{\,N-1} + \ldots + C_n * \mu_n^{\,N-1}
\end{aligned}
\qquad (6)
$$

The difficulty at this point is that both the C's and the μ's are unknown, and thus 2n equations are required to solve for the variables. Also, the μ's are the exponential values, thus nonlinear. *Prony*[16] observed that for any C and μ there exist complex constants α depending only on μ such that each μ is the root of the algebraic equation:

$$\mu^n - \mu^{n-1}\alpha_1 - \mu^{n-2}\alpha_2 - \ldots - \mu\alpha_{n-1} + \alpha_n = 0 \qquad (7)$$

The left hand side of the equation then has the solution $(\mu - \mu_1)(\mu - \mu_2)\ldots(\mu - \mu_n)$. To determine the coefficients $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \ldots \alpha_n$, multiply the left side of the first equation in (6) by α_n, the second by α_{n-1},… the nth equation by α_1, and the (n+1)th equation by –1. The result is shown below:

$$f_n - \alpha_1 f_{n-1} - \ldots - \alpha_n f_0 = 0. \qquad (8)$$

Successive equations can be developed in the same manner by starting with the second, third, fourth,…, $(N-n)^{th}$ equations of (6).

$$
\begin{aligned}
f_n &= \alpha_1 f_{n-1} - \alpha_2 f_{n-2} - \ldots - \alpha_n f_0 \\
f_{n+1} &= \alpha_1 f_n - \alpha_2 f_{n-1} - \ldots - \alpha_n f_1 \\
f_{n+2} &= \alpha_1 f_{n+1} - \alpha_2 f_n - \ldots - \alpha_n f_2 \\
f_{n+3} &= \alpha_1 f_{n+2} - \alpha_2 f_{n+1} - \ldots - \alpha_n f_3 \\
f_{n+4} &= \alpha_1 f_{n+3} - \alpha_2 f_{n+2} - \ldots - \alpha_n f_4 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
f_{N-1} &= \alpha_1 f_{N-2} - \alpha_2 f_{N-1} - \ldots - \alpha_n f_{N-n-1}
\end{aligned}
\qquad (9)
$$

Since the f's are known, we can solve these equations either directly if N = 2n or by the method of least squares if N > 2n. Once the α's are solved we can determine the roots of (7) and then solve (6) for the constants.

An important value in analyzing energy dissipation from the system is the Q Value. It is inversely proportional to the rate at which the electric or magnetic field energy is dissipated[13]. The Q value is defined as the resonant frequency divided by the FWHM (Full Width Half Maximum) bandwidth. The bandwidth is determined by the damping parameter. If the damping parameter is small, the peak will be sharper and the Q value will be larger. Once the complex frequencies are determined, the Q-value can be calculated rather easily:

$$Q_n = \text{Im}\{\omega_n\} / (2 * \text{Re}\{\omega_n\}) \quad (10)$$

To perform the data analysis, I developed MATLAB code that read the FDTD data, employed the exponential approximation method, and determined the resonant frequencies. Prony Approximation assumes that the number of components is known. An interesting challenge within the MATLAB code is that since the exact number of sinusoid components was unknown, every component possibility up through 20 was tested. Stable data was established typically between 10 and 17 components, and we used this data as our determined values. Also, the Prony method relies on time increments of equal value, but the simulation also depended upon the size of that increment. An increment of 0.0065 seconds was necessary to obtain accurate results. This could plausibly be due to the scaling factor problem inherent in the Prony Method[15].

Results:

The values from the Prony Method (Table 1) correspond very well with the analytical solution for the ideal case, and the nonideal solutions are very near to the results of *Sentman*[13].

Table 1

Mode (n)	Analytic Ideal	Prony Ideal	Sentman NonIdeal	Prony NonIdeal
1	10.47	10.39	7.61	7.53
2	18.13	17.9	13.4	13.4

Figure (2) shows the ideal Earth-ionosphere resonant frequencies (blue) and contrasts them with the nonideal case (red). Clearly visible is the downshift of the resonant frequencies due to energy dissipation from the system.

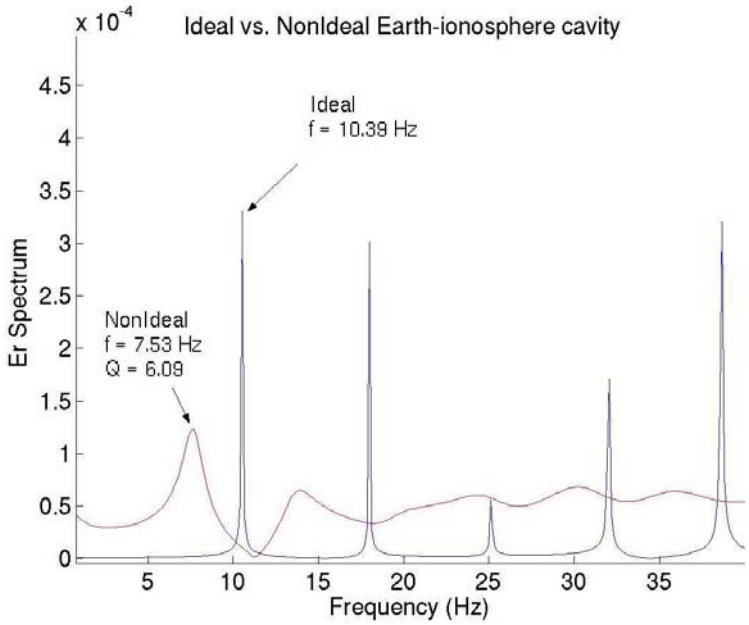


Figure 2: Ideal (blue) and Non-Ideal (red) cavity resonator.

Figure (3) shows the frequency results from a simulation with a conductivity perturbation ß= 3 and Figure (4) is its corresponding Q-value. The blue, solid lines correspond to data determined by *Sentman*[13], and the red, dashed lines are the values determined from this experiment.

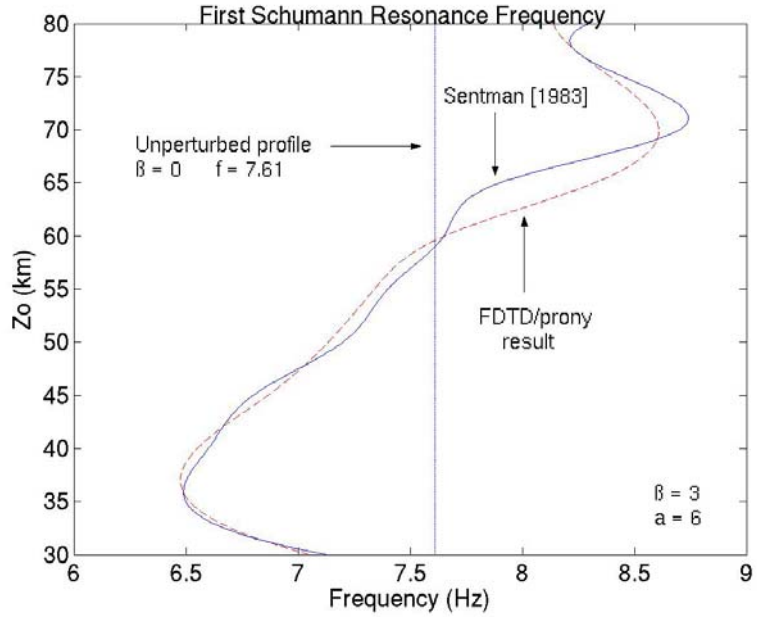


Figure 3: When perturbations occur above or below 60 km, the frequency increases or decreases from the nominal value, respectively.

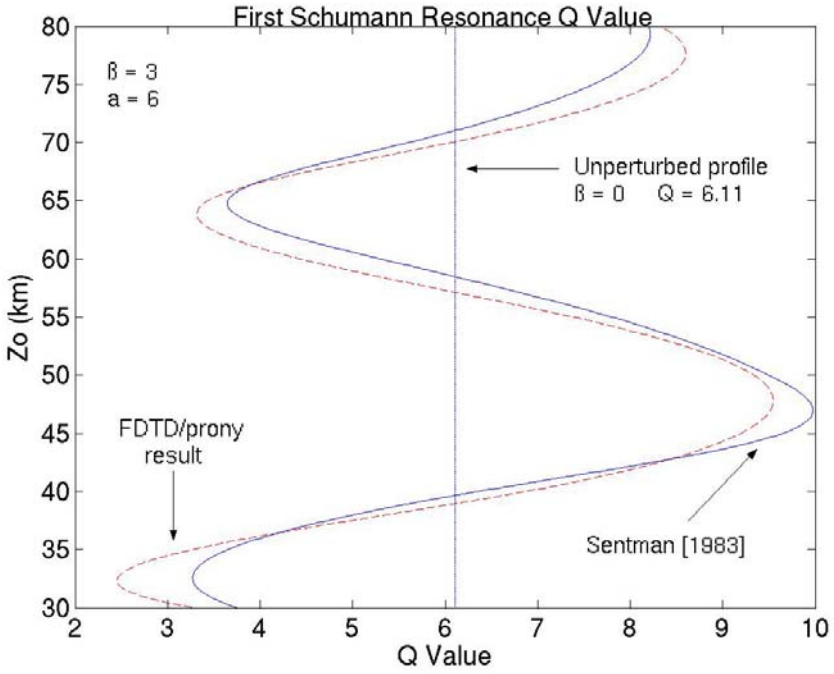


Figure 4: Q Value fluctuation adheres closely to the Joule dissipation, and changes drastically with the conductivity profile.

Discussion:

Our data reveals that the first Schumann Resonance frequency decreases when a perturbation occurs below an altitude of ~ 60 km, and the frequency increases when a perturbation occurs at an altitude above ~60 km. In Figure (3), a maximum increase of ~ 1 Hz was observed at an altitude of 70 km, and a maximum decrease of ~1.1 Hz occurred between 30 and 35 km. An explanation for the contrasting effects is that at low altitudes the ionosphere is a poor conductor and increasing the conductivity dissipates more energy and thus lowers the eigenfrequency. At high altitudes the ionosphere is a good conductor, and increasing conductivity makes it a more reflective surface, thereby decreasing energy loss[13].

The peculiarity of the Q Value depicted in Figure (4) led *Sentman*[17] to develop a two-scale height conductivity profile and factor in tesseral and sectoral modes at the reflective boundary as well as the zonal mode. In general, the conductive boundary begins at approximately 45 km, and the Q Value achieves a maximum between 45 and 50 km. This max concurs with the Joule Dissipation maximum for the radial electric field. The second maximum above 70 km corresponds with the transverse dissipation layer.

Changes in ionosphere conductivity at altitudes in the D and lower E region are often the result of X-ray bursts and Solar Proton Events[18]. Solar X-ray bursts typically last for less than 1 hour and are usually followed by SPEs, and these change the conductivity profile at different heights. While *Schlegel and*

Füllekrug[20] concluded that SPEs increase Schumann Resonance frequencies, those values were daily averages and could be flawed. Our results correspond with the conclusions of *Roldugin et. al*.[18, 19] that X-ray bursts cause an increase in SR frequencies and SPEs cause their decrement. X-ray bursts interact with the ionosphere at higher altitudes and cause a slight increase in resonant frequency, while SPEs increase ionization at lower altitudes and cause a decrease in resonant frequency.

With an ultimate goal being able to monitor ionospheric conductivity from Schumann Resonance data, changes in eigenfrequency and Q Values cannot do the job alone[13]. While this model corresponds well with observed perturbations, the reverse process is not easy to extrapolate. The complexity of the Q Value leaves ambiguity as to the altitude and size of perturbations, and current sampling techniques are not refined enough to yield definitive answers.

Acknowledgement:

I would like to offer thanks to Professor Victor Pasko for affording me this opportunity to learn and help with this exciting topic. I would especially like to thank Heng Yang for his instruction and support. I would also like to thank Dr. Ruyan Guo, Dr. Kennith Jenkins, Linda Becker. This material is based upon work supported by the National Science Foundation under Grant No. EEC-0244030.

References:

[1] Roldugin, V.C., Y.P. Maltsev, A. N. Vasiljev, and E.V. Vashenyuk, Changes of the first Schumann resonance frequency during relativistic solar proton precipitation in the 6 November 1997 event, *Ann. Geophys*., 17, 1293-1297, 1999.

[2] Yang, H., and V.P. Pasko, Three-dimensional FDTD Modeling of Earth-ionosphere Cavity Resonances, *Eos Trans.* AGU, 84(46), Fall Meet. Suppl., Abstract 2427-3304-3324, 2003.

[3] Hildebrand F.B., *Introduction to Numerical Analysis*, McGraw-Hill, New York, 1956.

[4] Schumann, W. O., On the radiation free self-oscillations of a conducting sphere, which is surrounded by an air layer and an ionospheric shell (in German),*Z. Naturforsch*., 7, 149-154, 1952.

[5] Balser, M., and C. A. Wagner, Observations of earth-ionosphere cavity resonances, *Nature* (*London*), 188, 863, 1962.

[6] Sentman, D.D., "Schumann Resonances"; pp. 267–295 in *Atmospheric Electrodynamics*, Edited by Hans Volland, CRC Press, New York, 1995.

[7] Sato, M., and H. Fukunishi, Global sprite occurrence locations and rates

derived from triangulation of transient Schumann resonance events, *Geophys. Res. Lett.*, 30, 1859, 2003.

[8] Shvets, A. V., A technique for reconstruction of global lightning distance profile from background Schumann resonance signal, *J. Atmos.Solar-Terr. Phys*., 63(10), 1061, 2001.

[9] Cummer, S.A., Ionospheric E region remote sensing with ELF radio atmospherics *Radio Science*, 35, 1437-1444, 2000.

[10] Williams, E.R., The Schumann Resonance: A Global Tropical Thermometer, *Science*, 256, 1184-1187, 1992.

[11] Satori, G., Monitoring Schumann resonances–ll. Daily and seasonal frequency variations, *J. Atmos. Terr. Phys*., 58, 1483, 1996.

[12]Cannon, P.S. and M.J. Rycroft, Schumann resonance frequency variations during sudden ionospheric disturbances, *J. Atmos. Terr. Phys*., 44, 201-206, 1982.

[13]Sentman, D. D., Schumann resonance effects of electrical conductivity perturbations in an exponential atmospheric/ionospheric profile, *J. Atmos.Terr. Phys.*, 45, 55-66, 1983.

[14] Füllekrug, M., Schumann resonances in magnetic field components, *J. Atmos. Terr. Phys*., 57 (5), 479-484, 1994.

[15]Kahn, M.H., M.S. Mackisack, M.R. Osborne, and G.K. Smyth, On the Consistency of Prony's Method and Related Algorithms, *J. Comput. Graph. Statist*., 1, 329-349, 1992.

[16]Prony, R., Essay experimental et Analytique, *J. École Polytechnique*, Paris 1, 24-76, 1795.

[17]Sentman, D. D., Approximate Schumann resonance parameters for a two-scale height ionosphere, *J. Atmos. Terr. Phys.*, 52, 35-36, 1990.

[18]Roldugin V. C., Y. P. Maltsev, A. N. Vasiljev, A. V. Shvets, and A. P. Nikolaenko, Changes of Schumann resonance parameters during the solarproton event of 14 July 2000, *J. Geophys. res*., 108 (A3), 1103, 2003.

[19]Roldugin V. C., Y. P. Maltsev, and A. N. Vasiljev, Schumann resonanc frequency increase during solar X-ray bursts, *J. Geophys. res*., 109, A01216, doi:10.1029/2003JA010019, 2004.

[20]Schlegel, K. and Füllekrug, M, Schumann resonance parameter changes during high-energy particle precipitation, *J. Geophys. res*., 104 (A5), 10111, 1999.

PRINCIPLES OF ADAPTIVE NOISE CANCELING

Mohamed A. Abdulmagid*, Dean J. Krusienski[+], Siddharth Pal[+], and
William K. Jenkins[#]

Department of Electrical & Computer Engineering
The Pennsylvania State University, University Park, PA 16802

*Undergraduate student of
Department of Electrical & Computer Engineering
Auburn University
Auburn, AL 36830

ABSTRACT

The purpose of this paper is to learn about adaptive filtering theory, learning algorithms, and to develop an appreciation for some of the applications of adaptive filters. The basic idea is to use the LMS (Least-Mean-Square) algorithm to develop an adaptive filter that can be used in ANC (Adaptive Noise Cancellation) applications. The method uses a noisy signal as primary input and a reference input that consists of noise correlated in some unknown way with the primary noise. By adaptively filtering and subtracting the reference input from the primary input, the output of the adaptive filter will be the error signal, which acts as a feedback to the adaptive filter. With this setup, the adaptive filter will be able to cancel the noise and obtain a less noisy signal estimate.

INTRODUCTION AND BACKGROUND

The first work in adaptive noise canceling was done by Howells and Applebaum and their colleagues at General Electrical Company. Their design was a system for antenna side lobe canceling that used a reference input derived from an auxiliary antenna and a simple two-weight adaptive filter[1]. The first adaptive noise canceling system at Stanford University was designed and built in 1965. The purpose was to cancel 60 Hz interference at the output of an electrocardiographic amplifier and recorder[1].

[#] Faculty Mentor
[+] Graduate Mentor

The basic concept of adaptive noise canceling is shown in figure 1^2. A signal s(n) is transmitted over a channel to a sensor that also receives noise **N**o(n) uncorrelated with the signal. The combined signal and noise (s(n) + **N**o(n)) form the primary input to the canceller. A second sensor receives a noise **N**1(n) uncorrelated with the signal but correlated in some unknown way with the noise **N**o(n). This sensor provides the reference input to the canceller. The noise **N**1(n) is filtered to produce an output y that is as close a replica as possible of **N**o(n). This output is subtracted from the primary input [s(n) + **N**o(n)] to produce the output of the system (e(n) = s(n) + **N**o(n) – y(n)).

In the system shown in figure 1^2, the reference input is processed by an adaptive filter, which differs from the regular or fixed filter. The adaptive filter automatically adjusts its own impulse response through an algorithm that responds to an error signal, which depends on the filter's output. In the noise canceling application the objective is to produce an error signal that is a best fit in the least squares sense to the signal s(n). This is accomplished by feeding back the system output to the adaptive filter using an LMS algorithm to minimize the error signal until it reaches the value (e(n) = s(n)).

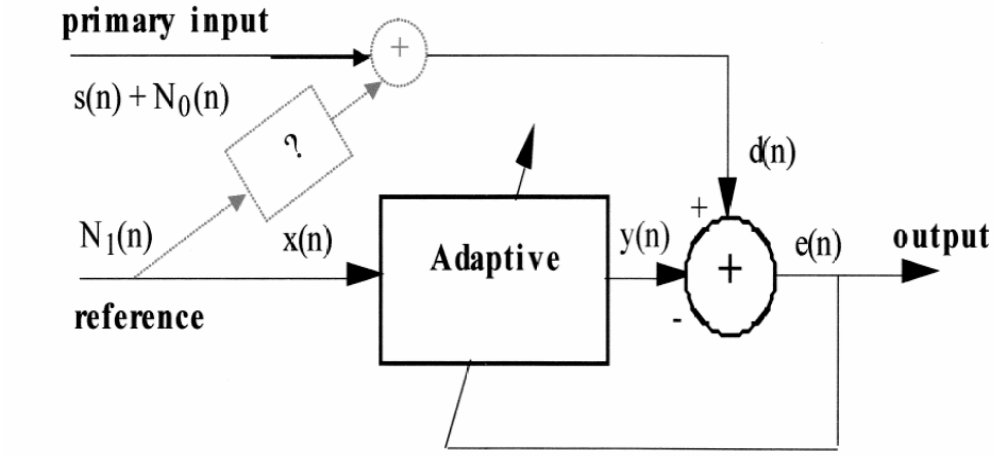


Fig. 1. Adaptive Noise Canceller

Assume that s(n), **N**o(n), **N**1(n), and y(n) are statistically stationary and have zero means. Assume that s(n) is uncorrelated with **N**o(n) and **N**1(n), and suppose that **N**1(n) is correlated with **N**o(n). The output e(n) is

$$e(n)=s(n)+\mathbf{N}o(n)-y(n) \tag{1}$$

Squaring both sides

$$e^2(n)=s^2(n)+[\mathbf{N}o(n)-y(n)]^2+2s(n)[\mathbf{N}o(n)-y(n)] \tag{2}$$

Taking the expected value for both sides, and realizing that s(n) is uncorrelated with **N**o(n) and **N**1(n) and with y(n) results in

$$E[e^2(n)] = E[s^2(n)] + E[[(\mathbf{N}o(n) - y(n))^2] + 2E[s(n)[\mathbf{N}o(n) - y(n)]]$$

$$= E[s^2(n)] + E[(\mathbf{N}o(n) - y(n))^2] \quad (3)$$

The signal power $E[s^2(n)]$ will be unaffected as the filter is adjusted to minimize $E[e^2(n)]$. The minimum output power will be given by

$$\min E[e^2(n)] = E[s^2(n)] + \min E[s(n)[\mathbf{N}o(n) - y(n)]^2] \quad (4)$$

When the filter is adjusted so that $E[e^2]$ is minimized, $E[(\mathbf{N}o(n) - y(n))^2]$ is also minimized. The filter output y(n) is then a best least squares estimate of the primary noise **N**o(n). Also when $E[(\mathbf{N}o(n) - y(n))^2]$ is minimized, $E[(e(n) - s(n))^2]$ is also minimized, and that leads to

$$[e(n) - s(n)] \approx [\mathbf{N}o(n) - y(n)] \quad (5)$$

The output e(n) will contain the signal s(n) plus noise. From (1), the output noise is given by [**N**o(n) – y(n)]. Since minimizing $E[e^2]$ minimizes $E[(\mathbf{N}o - y)^2]$, and since the signal in the output remains constant, minimizing the total output power maximizes the output signal-to-noise ratio. From (3), it seems that the smallest possible output power is $E[e^2] = E[s^2]$, and that will lead to $E[(\mathbf{N}o - y)^2] \approx 0$. Therefore, y(n) = **N**o(n), and e(n) = s(n). In this case, minimizing output power causes the output signal to be noise free.

EXPERIMENTAL DESCRIPTION

In this section, all experiments, equipments, and methods are described. The setup of the experiment is to have a white noise wave file played through the speakers of the desktop. One of the speakers is close to someone who is talking into the microphone (the primary input with noise), and the other is close to another microphone but they are inside a box (to make the reference noise different from the primary noise but at the same time correlated to it).

The code is developed in Matlab using the LMS algorithm. The input of the adaptive filter x(n), which is a vector of input samples (the reference noise), is multiplied with a weight vector, w(n), corresponding to the input vector to obtain the output y(n)

$$y(n) = \mathbf{W}^T(n) * \mathbf{X}(n) \quad (6)$$

This output is subtracted from the desired signal d(n) (the primary input) to obtain the output error signal

$$e(n)=d(n)–y(n) \quad (7)$$

Then the LMS equation will be

$$\mathbf{W}(n+1)=\mathbf{W}(n)-\mu*e(n)*\mathbf{x}(n), \quad (8)$$

where: $\mathbf{x}(n) = [x(n, x(n-1),\ldots,x(n-N+1)]^T$

$\mathbf{W}(n) = [w_0(n), w_1(n),\ldots,w_{N-1}(n)]^T$

μ is the step size

The error signal is used as a feedback to the adaptive filter. The weight vector is adjusted (by the error signal) to cause the output y(n) to agree as closely as possible with the desired signal d(n). This is accomplished by comparing the output y(n) with the desired signal d(n) (the primary input) to obtain the error signal and then adjusting the weight vector to minimize this signal.

RESULTS

Two different configurations for ANC filters were investigated experimentally: 1) the adaptive line enhancer and 2) the more general ANC filter that depends on an appropriate reference noise signal. The adaptive line enhancer is effective in situations where the primary signal of interest is narrow band (perhaps even a single sinusoid). The general ANC requires that a reference noise can be acquired that is correlated with the primary noise on the input signal, and which contains no components of the primary information signal.

Adaptive Line Enhancer

The adaptive line enhancer (figure 2)[2] is a good example to start with to apply the adaptive filtering theory. The adaptive line enhancer simply predicts a sinusoid wave. In other words, it finds a sinusoid wave in noise. The input of the system is the sinusoid wave with some noise in the background. When we pass a delayed version of that wave through the adaptive filter and compare it to the original wave, the output of the filter is going to be the sinusoid wave and the output of the system will be the background noise.

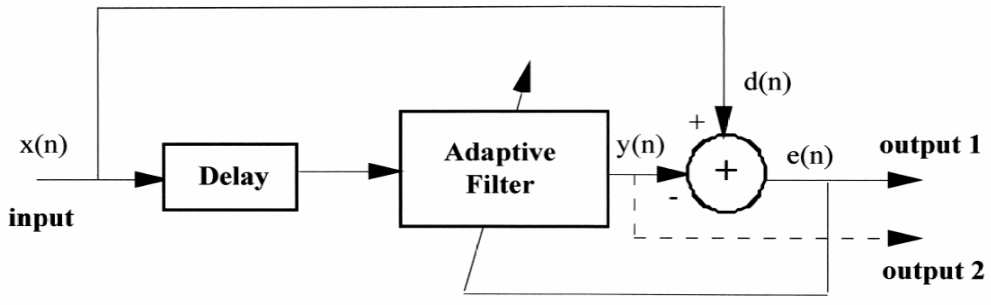


Fig. 2. Adaptive Line Enhancer

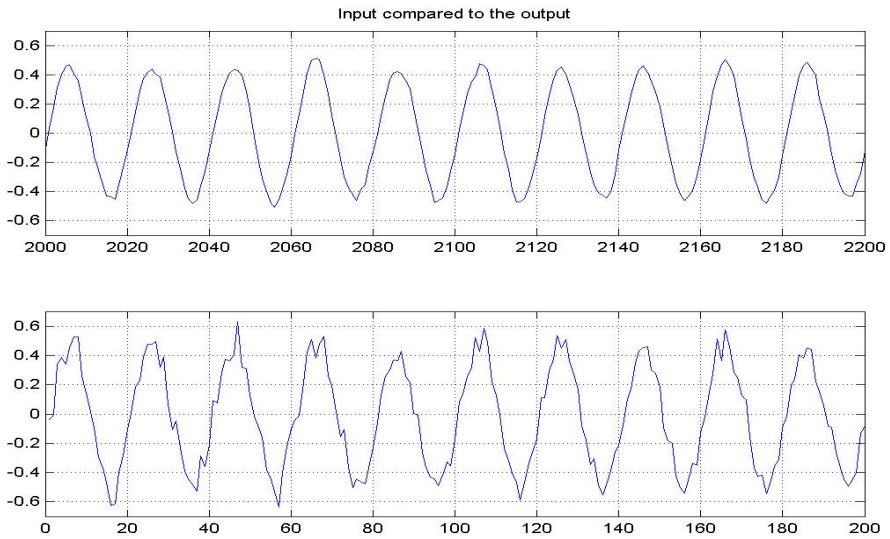


Fig. 3. Noisy signal compared to the cleaned signal

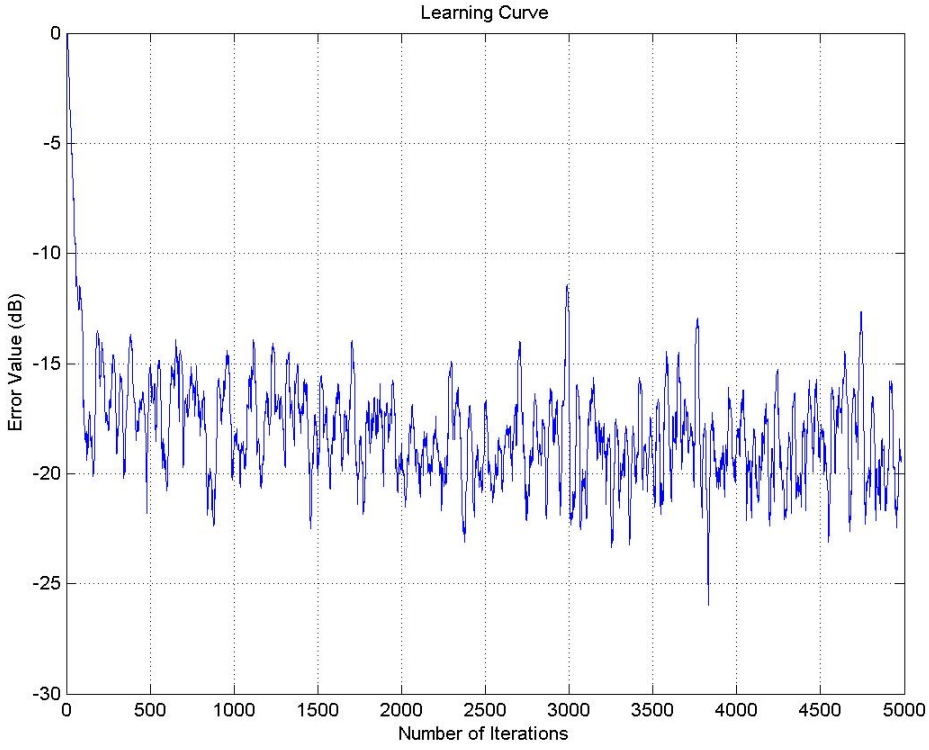


Fig. 4. Learning curve of the adaptive line enhancer in dB

Figure 3 shows a sinusoid with some additive white noise being passed through the adaptive line enhancer. The one at the bottom of the figure is the input of the filter, and the one in the top is the output of the filter. Figure 4 shows the learning curve of the adaptive line enhancer. The following parameters are used to obtain this result: n = 100, delay = 1, and the input signal is a sinusoid signal with white noise added to it. The X-axes shows the number of iterations and the Y-axes shows the value of the error at that iteration.

General ANC Experiment

In the noise canceling experiment (main experiment), we passed the reference noise through the adaptive filter to be filtered and subtracted from the noisy signal (the primary input). The input signal was a speech signal with white noise in the background; the reference input was another kind of noise that correlated with the primary noise (the primary noise is been passed through a coloring filter, which has length of 4 taps). The coloring filter was a 4-tap FIR filter with tap values given by h(0) = 1, h(1) = -0.7, h(2) = 0.4, h(3) = -0.5, and the step size was 0.1.

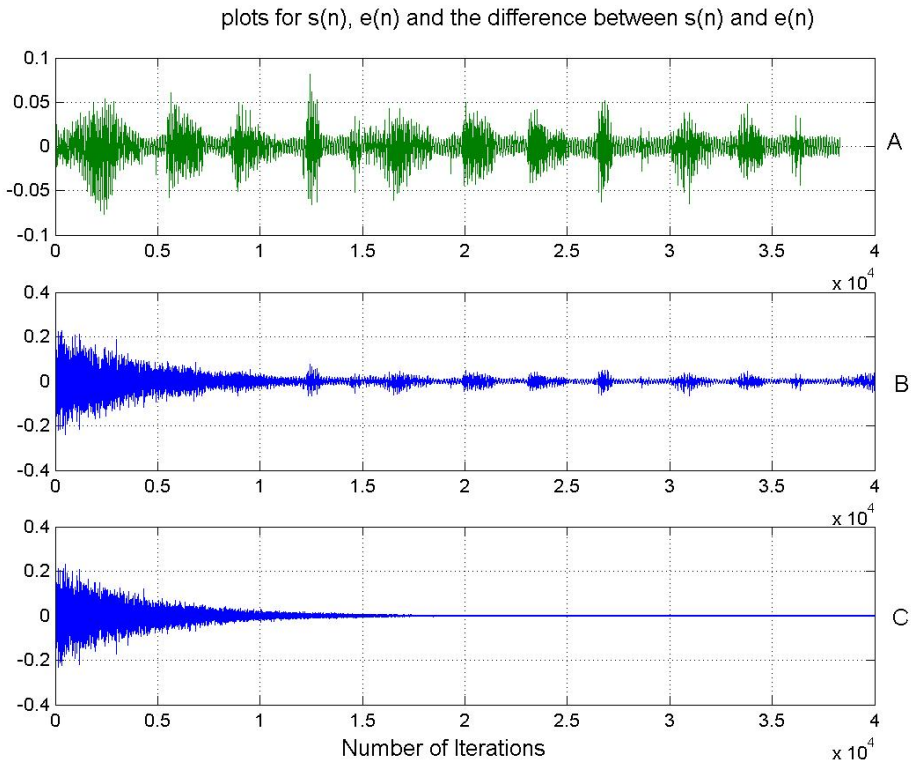


Fig. 5. The input signal, the output signal, and the error signal

Figure 5 shows the speech signal, the output signal, and the error signal, which was used as a feedback signal to adjust the adaptive filter. The X-axes shows the number of iterations and the Y-axes shows the value of the error at that

iteration. Figure 6 shows the weights of the adaptive filter and how they converge from random values to the correct values of the coloring filter. As we can see from figures 5 and 6, the adaptation is complete after 2000 iterations and the amplitude value of the error signal becomes zero.

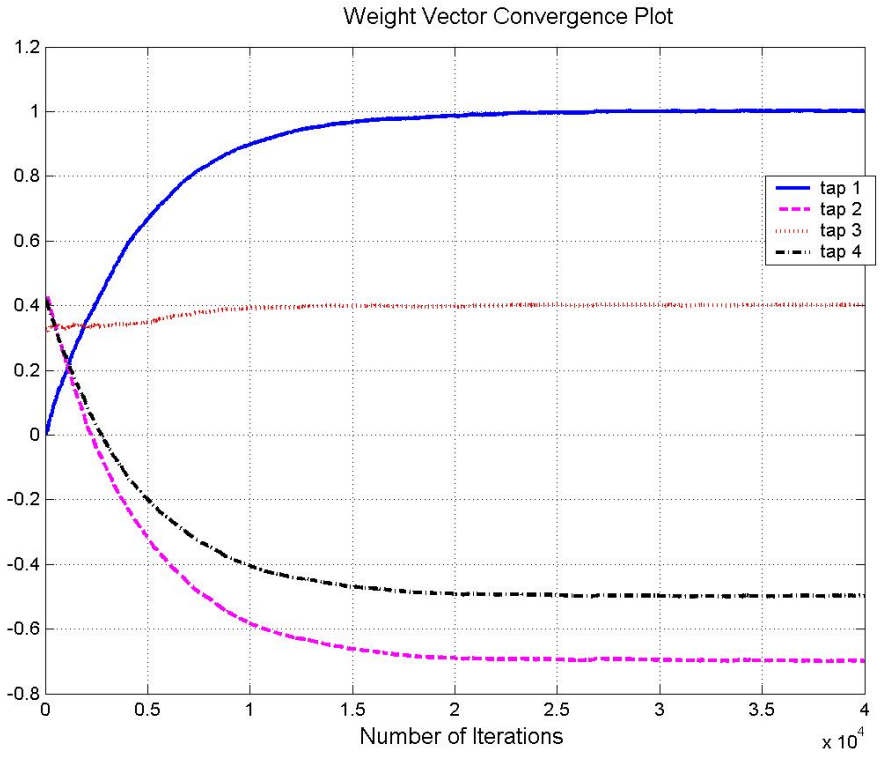


Fig. 6. Weight vector convergence

DISCUSSION

As theory implies, the adaptive line enhancer uses a training signal, which is the current value of the input, and a delayed version of that signal so past values of the input are used to predict the present input value. The assumption is that the noise samples which are more than M samples apart are uncorrelated with one another. As a result, the predictor can only make a prediction about the sinusoidal components of the input signal, and when adapted to minimize the output MSE, the line enhancer will be a filter tuned to the sinusoidal components. The maximum possible rejection of the noise will also be achieved, since any portion of the noise that passes through the prediction filter will enhance the output MSE whose minimization is the criterion in adapting the filter tap weights.

In the experiment of canceling noise in a speech signal the interference was a white noise wave containing many harmonics that varied in amplitude, phase, and waveform from point to point in the room. The noise canceller was able to reduce

the output power of this noise by 20 to 30 dB, making the interference barely perceptible to the listener. No noticeable distortion was introduced into the speech signal, and convergence times were on the order of seconds.

CONCLUSION

Adaptive noise canceling is a method of optimal filtering that can be applied whenever a suitable reference input is available. The advantages of this method are its adaptive capability, its low output noise, and its low signal distortion. It leads to a system that automatically turns itself off when no improvement in signal-to-noise ratio can be achieved. It is clear from the experiment that the output noise and signal distortion are generally lower than can be achieved with optimal (fixed) filter configurations.

The experimental data presented in this paper demonstrated the ability of the adaptive noise canceling to reduce additive periodic or stationary random noise in both periodic and random signals. One of today's applications of the adaptive filtering theory is Hearing Aids, which could be improved by the digital feedback noise cancellation. Basic ANC can be used to improve hearing aids, and that will make the ear mold more comfortable and more signal strength is available especially in the noisy places.

ACKNOWLEDGMENT

Many people have contributed support, assistance, and ideas in accomplishing this paper. The author especially would like to acknowledge the contribution of Prof. W. Jenkins, Head of Electrical Engineering Department, Dr. Dean J. Krusienski, who was the direct supervisor of the project presented in the paper, and Mr. Siddharth Pal, graduate mentor, who has a significant help in the development of the project presented in this paper. This material is based upon work supported by the National Science Foundation under Grant No. EEC-0244030.

REFERENCES

[1] B. Widrow, J. Glover, JR., John McCool, John Kaunitz, Charles Williams, Robert Hearn, James Zeidler, Eugene Dong, JR., and Robert Goodlin, “Adaptive Noise Canceling: Principles and Applications”. Proceedings of the IEEE, VOL.63, NO.12 December 1975.

[2] S. Haykin, *Adaptive Filter Theory.* New Jersey: Prentice Hall (1996).

[3] W. Kenneth Jenkins, A. W. Hull, J. C. Strait, B. A. Schnaufer, X. Li, *Advanced Concepts in Adaptive Signal Processing.* Massachusetts: Kluwer Academic Publishers (1996).

PIEZOELECTRIC ENERGY HARVESTING USING ELEMENTS OF COUPLED MODE THEORY

Joseph P. Tadduni*, Yiming Liu[+], and Heath F. Hofmann[#]

Department of Electrical Engineering
The Pennsylvania State University
University Park, PA 16802

*Undergraduate student of
Department of Electrical and Computer Engineering
Duke University
Durham, NC 27708

ABSTRACT

As remote and embedded micro-scale systems have become ubiquitous, the problem of powering these systems remotely has become prevalent. The use of piezoelectric elements to harvest electrical energy from small mechanical vibrations has been proposed as a means by which these systems can power themselves. Previous studies have utilized controllers to maximize the power output of the piezoelectric element, but these controllers would need to consume some of the power that is harvested. This paper investigates the coupling of the electrical and mechanical modes according to coupled mode theory, and the circuitry that will maximize this coupling. A piezoelectric element is connected to an L-C circuit with the same resonant frequency as the vibrating structure. The transfer of energy between the electrical and mechanical modes is observed. The energy in the L-C circuit is compared to the energy absorbed by a resistor. Results are compared to a mathematical model of the system.

[#] Faculty Mentor
[+] Graduate Mentor

INTRODUCTION

The possibility of self-powered electronic devices and embedded sensors has spurred much research into energy harvesting, or the conversion of energy in the ambient environment into usable electrical energy. One common approach to the problem involves using piezoelectric transducers to harvest energy from ambient mechanical vibrations. Piezoelectric ceramics such as lead zirconate titanate (PZT) produce a voltage when a stress is applied, and conversely strain when a voltage is applied. The goal of energy harvesting is to eliminate the need for finite energy sources, such as electrochemical batteries, that eventually need to be replaced. An energy harvesting power supply should at least be able to extend the battery life, and ideally would make replacement unnecessary. Wireless, remote, and embedded sensors would especially benefit from energy harvesting, as these devices typically require low power levels, and it is expensive and inconvenient to replace their batteries. Piezoelectric energy harvesting has possible applications ranging from gathering energy in shoes while walking[1] to powering microelectromechanical systems (MEMS) devices.

Most previous investigations into power harvesting have involved a PZT device attached to or contained in a mechanical beam, which is shaken to simulate vibrations in the structure that the beam is attached to. Glynne-Jones *et al.*[2] attached a resistive load across the PZT to measure the available power it produced. Ottman *et al.*[3] used a rectifier circuit with an adaptive DC-DC step-down converter, which charged a battery at four times the rate of the rectifier alone. The advantage of this approach is that it can maximize power output over a wide range of excitation frequencies; however, the control circuitry consumes a significant portion of the available energy. An alternate method that minimizes power-consuming circuitry could possibly prove to be more efficient.

Another application in which PZT transducers are attached to vibrating structures is in shunt damping systems. In cases where vibration is undesirable, the PZT can be shunted by a series R-L circuit to damp out the vibrational energy. The R-L circuit is tuned to the resonance frequency or frequencies of the structure, creating a damped electrical resonance that dissipates the energy[4]. The operation of these systems depends on precise matching between the resonance frequencies of the mechanical structure and the electrical circuit. If the structure is excited by an impulse of energy, most of the energy will go into the first vibrational mode. If the shunted R-L circuit is tuned to the frequency of this mode, most of the vibrational energy can be damped out.

The aim of this project is to harvest energy using the general approach to shunt damping, i.e. using a resonant L-C circuit tuned to the resonant frequency of the structure. According to coupled mode theory[5], the energy in the system should oscillate between the mechanical and electrical modes as a result of the coupling provided by the piezoelectric material. This provides a means for the vibrational energy to be quickly and completely transferred into electrical energy.

When the energy is in the electrical mode, the current could be switched out of the inductor and used to charge a battery, or stored in another manner. This strategy could also be applied to vibrational damping systems and might help remove the mechanical energy from the vibrating structure more quickly. Mathematical and physical models were constructed to test this approach and to observe whether or not the system provides better coupling between the two modes than other approaches do.

EXPERIMENTAL DESCRIPTION

Mathematical Model

A mathematical model of the electromechanical system was developed using Matlab in order to study the basic dynamics of its behavior. While the mechanical portion of the system would actually be a vibrating cantilever beam, it was modeled as a spring-mass system with spring constant k and mass m, which is a continuum system. This system exhibits the same mathematic behavior as the beam's fundamental mode of vibration. The electrical mode consisted of the piezoelectric capacitance C shunted by an inductance L. For simplicity, the model assumed no mechanical or electrical damping.

The system is described by a state vector $\vec{s}$ of four variables,

$$\vec{s} = \begin{bmatrix} x \\ u \\ i \\ v \end{bmatrix} \tag{1}$$

where x and u are the displacement and velocity of the mass, and i and v are the current through and voltage across the inductor. The two equations governing the interaction between the electrical and mechanical modes are

$$D = \varepsilon E + d\sigma \tag{2}$$
$$e = s\sigma + dE \tag{3}$$

where D and E are the electric displacement and electric field, σ and e are the stress and strain, and d is the piezoelectric charge constant. Multiplying (2) by the area of the piezoelectric element and multiplying (3) by its length gives

$$q = Cv + df_p \tag{4}$$

$$x = \frac{f_p}{k} + dv \tag{5}$$

where q is the charge displaced across the piezoelectric element and f_p is the total force on the piezoelectric element.

The system behaves according to the state equation

$$\dot{\vec{s}} = A\vec{s} \tag{6}$$

where A is the state matrix. The state vector solution has the form

$$\vec{s}(t) = e^{At}\vec{s}(0) \tag{7}$$

where e^{At} is a matrix exponential. The solution can be computed if the state matrix A can be determined. The equation for $\dot{x}$ can be determined simply by noting that

$$\dot{x} = u\,. \tag{8}$$

The derivative of u can be found by considering the equation that describes the mechanical behavior. The force on the piezoelectric element, f_p, is the opposite of the force on the mass, f_m, and f_p can be solved for in (5), so

$$f_m = ma_m = -f_p = -kx + dkv\,. \tag{9}$$

Noting that $a_m = \dot{u}$, this equation yields

$$\dot{u} = -\frac{k}{m}x + \frac{dk}{m}v\,. \tag{10}$$

The equation describing the behavior of the electrical system determines the derivative of i,

$$\dot{i} = \frac{1}{L}v\,. \tag{11}$$

Lastly, $\dot{v}$ can be determined by once again using (5) to find an expression for f_p and substituting this into (4). Differentiating the result and solving for $\dot{v}$ gives

$$\dot{v} = \frac{1}{C - d^2k}i + \frac{-dk}{C - d^2k}u\,. \tag{12}$$

Equations (8), (10), (11) and (12) determine the state matrix,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{k}{m} & 0 & 0 & \frac{dk}{m} \\ 0 & 0 & 0 & \frac{1}{L} \\ 0 & \frac{-dk}{C-d^2k} & \frac{-1}{C-d^2k} & 0 \end{bmatrix} \tag{13}$$

which can be substituted into (7) to calculate the system's behavior.

In the simulation, the system was given an initial displacement, with the other initial conditions set to zero. The parameter values used are listed in Table I. The capacitance and inductance were chosen to match the values used in the physical experiment, but the others were chosen arbitrarily since they do not correspond to directly measurable characteristics of the experimental setup. The importance of the simulation was its general behavior, not the specific values it produced.

Table I. Parameter values used in simulation.

Parameter	Value
d	100 nC/N
L	502.6 H
C	126 nF
k	4.62×10^4 kg/s^2
m	2.93 kg
$x(0)$	11 μm

The total energy in the mechanical mode, E_m, and in the electrical mode, E_e, were calculated and plotted. The energy was calculated from the equations

$$E_m = \tfrac{1}{2}kx^2 + \tfrac{1}{2}mv^2 \tag{14}$$

$$E_e = \tfrac{1}{2}Li^2 + \tfrac{1}{2}Cv^2 . \tag{15}$$

Physical Experiment

The experimental setup consisted of a small piezoelectric plate attached to a vibrating structure. At first, the use of a tuning fork with PZT plates attached to its sides was attempted because of its high resonant frequency. However, because of the delicate nature of the tuning fork, the attached plates (8mm × 19mm × 1mm) overdamped the structure and prevented it from resonating. The structure used instead was an aluminum beam clamped tightly to a table, with properties listed in Table II. The piezoelectric element used was a Macro Fiber Composite (MFC) from Smart Material Corp., with properties listed in Table III. The MFC

consists of rectangular piezoceramic rods between layers of adhesive and electroded film.

Table II. Properties of aluminum beam.

Composition	6061 T6 Aluminum
Length	50.0 cm
Width	5.05 cm
Thickness	.635 cm
Resonant Frequency	20 Hz

Table III. Properties of piezoelectric element.[6]

Part Number	M 8528 P1
PZT Type	Navy Type II
Active Area	85 mm × 28 mm
Thickness	0.3 mm
Capacitance	7 nF
d_{33}	400 pC/N
Q_M	75

The piezoelectric element was connected to a shunt L-C circuit as shown in Figure I. The piezoelectric element can be modeled as a current source in parallel with the inherent capacitance of the element. Different L-C combinations with a resonant frequency of 20Hz were tested. Because of the large inductances required to achieve such a low resonance, a general impedance converter (GIC)[7] circuit was used to simulate an inductor, shown in Figure II. While this circuit would probably be impractical in an energy harvesting application, it is used here to demonstrate the overall concept being tested.

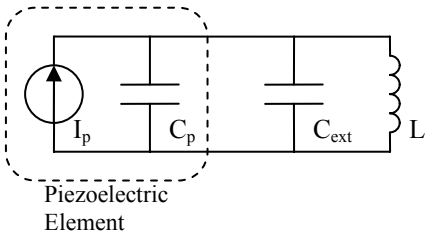


Figure I. Experimental circuit layout.

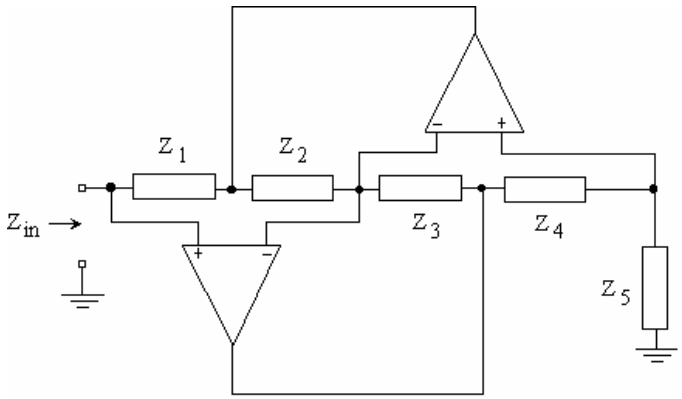


Figure II. General impedance converter.[7]

The impedance seen at the input terminals of the GIC is

$$Z_{in} = \frac{Z_1 Z_3 Z_5}{Z_2 Z_4}. \tag{16}$$

By making either Z_2 or Z_4 a capacitor and the rest resistors, the circuit will simulate an inductor with equivalent inductance given by

$$L_{eq} = \frac{R_1 R_3 R_5 C_4}{R_2}. \tag{17}$$

The virtual inductor was built using two LM741 op-amps, a 33nF capacitor, and varying resistances to simulate different inductance values.

The piezoelectric element was also placed in shunt with a resistor to compare the energy absorbed by the resistor to the peak energy seen in the L-C circuit. To achieve maximum power transfer, a resistance of 1MΩ was chosen to match the impedance of the element's 7nF internal capacitance at 20Hz.

RESULTS

Mathematical Model

The simulation demonstrates the theoretical behavior of the system. The energy starts out in the mechanical mode and is transferred to the electrical mode. It oscillates between the two modes, modulating the displacement, velocity, current and voltage in sinusoidal envelopes. Plots of the displacement and current are displayed in Figure III. If the resonant frequencies of the two modes are matched exactly, all the energy is transmitted from one mode to the other, shown in Figure IV(a). If the resonant frequencies are mismatched, even by a very small

amount, much less energy is transferred from the mechanical to the electrical system. Figure IV(b) shows the energy transfer for a mechanical resonance of 20Hz and an electrical resonance of 22Hz ($\Delta f_r = 2$Hz).

The energy conversion frequency f_{ec}, the frequency at which the energy is transferred back and forth between the two modes, is a function of the parameters used. The strongest dependence of f_{ec} is on the piezoelectric coupling coefficient: f_{ec} is directly proportional to d regardless of the other parameter values. Determining the other relationships was not as straightforward because the remaining parameters also affect the resonant frequency f_r. First, the parameters L and m were chosen to maintain f_r at a constant value while C and k were varied. It was found that f_{ec} is proportional to the square root of k and inversely proportional to the square root of C. Conversely, when C and k were chosen to maintain f_r while L and m were varied, f_{ec} was found to be proportional to the square root of L and to the square root of m.

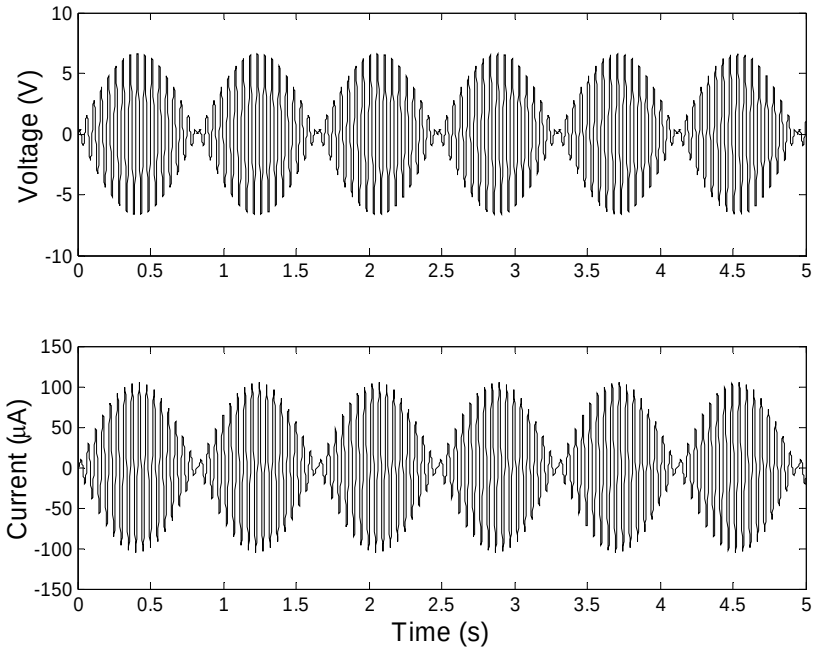


Figure III. Displacement and current in simulation.

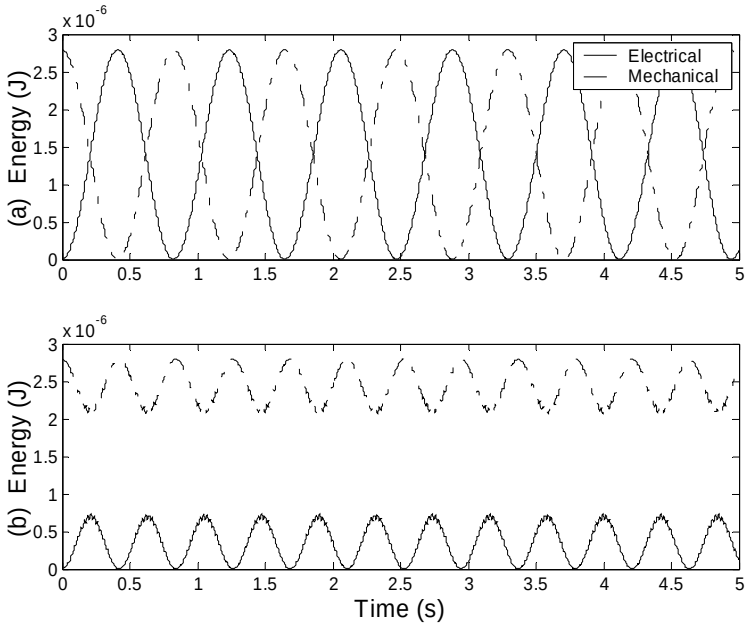


Figure IV. Energy conversion for (a) equal resonances and (b) $\Delta f_r = 2$Hz.

Experimental Results

The experimental circuit was tested with three different combinations of inductance and capacitance. The beam was struck with approximately the same force in each case. The peak values of current and voltage are displayed in Table IV along with the peak energy as calculated by $\frac{1}{2}CV^2$ and by $\frac{1}{2}LI^2$, and the energy conversion frequency. Figure V shows the voltage and current waveforms for the last case listed in the table.

Table IV. Energy transferred and energy conversion frequency for different L-C combinations.

C_{tot} (nF)	L (H)	V_{peak} (V)	I_{peak} (μA)	$\frac{1}{2}CV^2$ (μJ)	$\frac{1}{2}LI^2$ (μJ)	f_{ec}
192	330	4.3	111	1.78	2.03	0.91
159	398	5.6	102	2.49	2.07	1.16
126	503	6.1	106	2.34	2.83	1.67

The system did not display the same behavior when excited by vibrations of larger amplitudes. When the beam was struck any harder than just a light touch, the voltage and current showed a steady decay rather than the modulation seen in Figure V. After a few seconds, the waveforms became somewhat modulated but

still did not behave the same as for smaller excitations. The voltage and current waveforms for a larger vibration amplitude are shown in Figure VI.

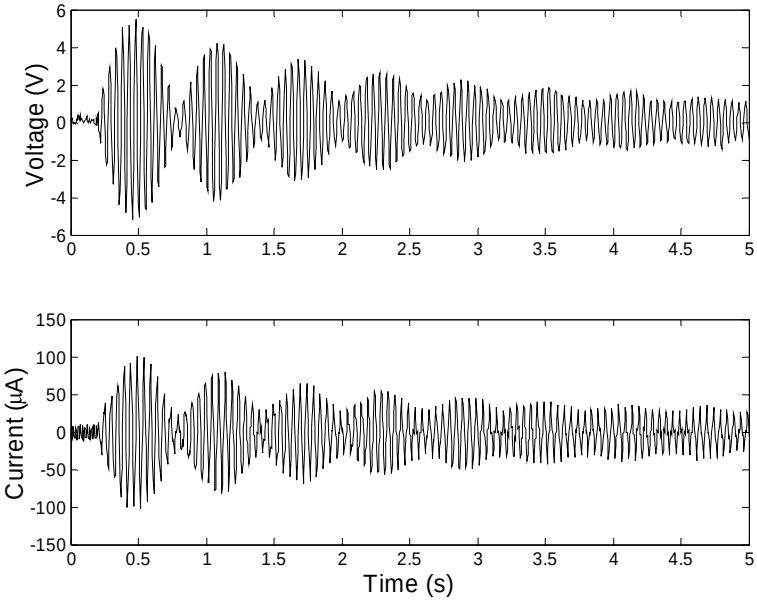


Figure V. Response to a small vibrational amplitude.

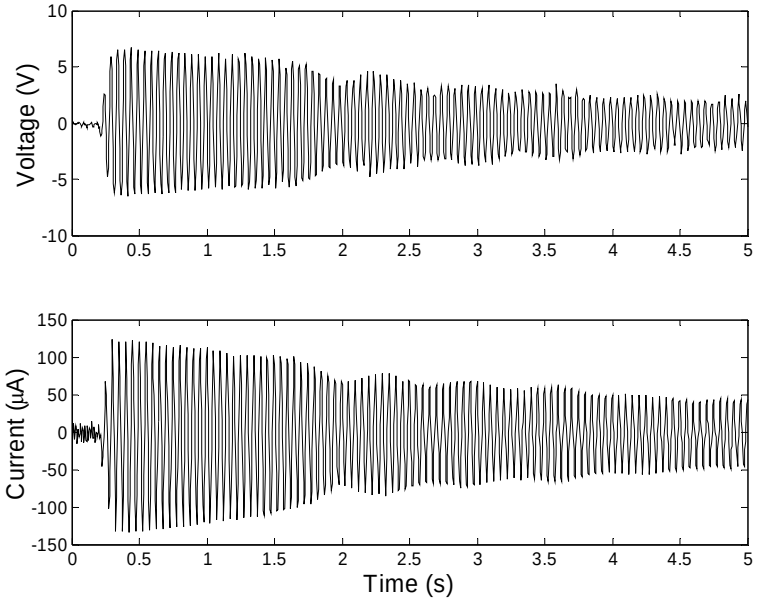


Figure VI. Response to a large vibrational amplitude.

The response of the piezoelectric element connected across the 1MΩ resistor was measured for a small vibrational amplitude similar to the ones used to collect the data in Table IV. The energy absorbed by the resistor was calculated by integrating the power absorbed over a five second period, after which most of the vibration had died out. The calculated value was 6.77μJ, nearly two and a half times more than the highest peak value observed in the L-C circuit.

DISCUSSION

The mathematical model demonstrates some essential characteristics of the electromechanical system. Most importantly, it reaffirms the idea presented by coupled mode theory that it should be possible, at least theoretically, to quickly convert all the energy from the mechanical mode to the electrical mode, provided that the two systems have the same resonant frequency and are linearly coupled. The observed relationships between the parameter values and the energy conversion frequency f_{ec} are significant in that they provide a means of controlling the speed at which energy is transferred from one mode to the other. This is important in energy harvesting applications because lower values of f_{ec} might result in significant loss of signal amplitude before the energy is finally converted to the electrical mode.

At low vibrational amplitudes, the experimental system exhibits the expected coupled mode behavior, similar to that of the model. One interesting trend in the data is that the levels of energy transferred rise, on average, as the external capacitance is reduced. This may be the result of better effective electro-mechanical coupling. Fleming *et al.*[4] showed that less external capacitance provides better effective coupling in shunt damping systems. Also, as the total capacitance is reduced, the energy conversion frequency increases, as predicted by the model.

It is not completely clear why the experimental system does not display coupled mode behavior at higher vibrational amplitudes. It may be due in part to distortion caused by the op-amps in the virtual inductor circuit. It may also be caused by too low a coupling factor, possibly because of the large external capacitance. There might also be a nonlinear saturation effect in the piezoelectric device at higher amplitudes that causes the observed behavior. It also seems that the mechanical vibrations never die out completely at any point, i.e. not all the energy is being converted. This was never measured directly, but the fact that the resistor absorbed more energy than was seen in the L-C circuit suggests that not all the mechanical energy is being converted. This again may be due to some sort of saturation of the piezoelectric device. It might also be the case that the system is especially sensitive to differences in the resonant frequencies, which would require a more finely tunable inductor than the one used. There may be vibrations in other modes besides the fundamental mode, which would not be coupled with the L-C circuit. While the mathematical model suggested that it should be

possible for all the energy to be transferred, the model may have overlooked some limiting factor in the mechanical system or the dynamics of the coupling that would prevent all the energy from being transferred. Since the model technically simulates a spring-mass type system, there may be something fundamentally different about the cantilever beam, which is a continuum system, that causes it to behave differently from what the spring-mass model predicts.

Further experimentation is necessary to evaluate the usefulness of the coupled mode approach to energy harvesting. It could also prove to be a useful alternative to current mechanical damping techniques. Instead of using resistors to dissipate the converted energy, a coupled mode damper would quickly convert the mechanical energy to electrical energy and then switch the current out of the inductor. In future experiments, it would be desirable to investigate a mechanical structure with a higher resonant frequency, which would not require such large inductances and might eliminate the need for a virtual inductor. This would also allow the use of a smaller external capacitance, which might provide better coupling. A useful improvement would be to have a regulated and reproducible means of striking or vibrating the beam, and a method for measuring its displacement to see what portion of the mechanical energy is being converted. These enhancements might allow for further evaluation of the usefulness of this approach to energy harvesting.

CONCLUSION

This research investigated the application of coupled mode theory to piezoelectric energy harvesting. A mathematical model suggested that it should be possible to quickly transfer all the mechanical energy to the electrical mode provided that the two modes have the same resonant frequency and have linear coupling. The model also showed the energy oscillating back and forth between the two modes. An experimental setup of an L-C circuit connected across a piezoelectric element displayed the coupled mode behavior predicted by the model for small amplitudes of vibration. For larger amplitudes, however, the energy did not oscillate between the two modes. Also, the energy transferred to the L-C circuit was significantly less than the energy absorbed by a resistor. The research provides encouraging results that reinforce the theory and suggest that this approach may be effective for energy harvesting or mechanical damping applications. Further work is necessary to determine if this method is actually capable of harvesting usable quantities of energy.

ACKNOWLEDGEMENTS

Special thanks to Dr. Hofmann and to Yiming Liu for all their help and patience, without which this project would not have been possible. This material is based upon work supported by the National Science Foundation under Grant No. EEC-0244030.

REFERENCES

[1] Nathan S. Schenck and Joseph A. Paradiso, "Energy scavenging with shoe-mounted piezoelectrics," *IEEE Micro*, **21** (3) 30–42 (2001).

[2] P. Glynne-Jones, S.P. Beeby and N.M. White, "Towards a piezoelectric vibration-powered microgenerator," *IEE Proceedings – Science, Measurement and Technology*, **148** (2) 68–72 (2001).

[3] Geffrey K. Ottman, Heath F. Hofmann and George A. Lesieutre, "Adaptive piezoelectric energy harvesting circuit for wireless remote power supply," *IEEE Transactions on Power Electronics*, **17** (5) 669–676 (2002).

[4] A.J. Fleming, S. Behrens and S.O.R. Moheimani, "Reducing the inductance requirements of piezoelectric shunt damping systems," *Smart Materials and Structures*, **12** (1) 57–64 (2003).

[5] Hermann A. Haus and Weiping Huang, "Coupled-mode theory," *Proceedings of the IEEE*, **79** (10) 1505–1518 (1991).

[6] Kevin Wedeward, "Negative-impedance converters and general impedance converters," New Mexico Institute of Mining and Technology, http://www.ee.nmt.edu/~wedeward/EE212L/SP04/lab06.html, (2003).

[7] "MFC," Smart Material Corp, http://www.smart-material.com/Smart-choice.php?from=MFC, (2004).

ULTRASONIC PIEZOELECTRIC HYPOCHLOROUS ACID HUMIDIFIER FOR DISINFECTION APPLICATIONS

A. Pezeshk*, Y. Gao‡, and K. Uchino#

Department of Electrical Engineering and
International Center for Actuators and Transducers
The Pennsylvania State University, University Park, PA 16802

*Undergraduate Student of
Department of Electrical Engineering
Michigan Technological University, Houghton, MI 49931

ABSTRACT

Hypochlorous Acid has proven to be a strong disinfectant with no side effects on humans. Coupled with the atomization of the acidic solution, much higher disinfection effects can be expected. The acid is not sold as a pure solution since it naturally disintegrates after a few hours.

Due to the safety and non-necessity of cleaning of the residuals from the disintegration of the acid, a device that could generate and atomize the acid would be ideal for disinfection of office and hospital buildings against viruses like SARS, Anthrax, etc.

In this study a corrosion resistant electrolytic cell was designed to produce hypochlorous acid from brine. Nafion® membranes were used to separate the anode and cathode compartments and produce hypochlorous acid and sodium hydroxide in the anode and cathode compartments respectively. Different designs were tested to optimize the cell and to ensure safety. An ultrasonic piezoelectric atomizer was utilized to generate micro-droplets of the diluted acid. The damping effects of various acid resistant coatings on the atomizer and effect of atomization on the pH of the acidic solution were analyzed. Modifications to this design can be made in order to produce portable disinfecting apparatus for disinfection of tools, etc.

Faculty Mentor
‡ Postdoctoral Mentor

INTRODUCTION

Facilities such as hospitals, industrial plants, agricultural and food processing plants, schools, and medical equipment are all subject to microbial contamination. Studies on hypochlorous acid solution show that it is very effective as disinfectant against bacteria, fungi, viruses and permanent forms. Due to its oxidizing potential, the hypochlorous acid transfers atomic oxygen in the form of a radical to the microorganism which destroys them[1].

There are many advantages to using hypochlorous acid for disinfection. Hypochlorous acid is nonirritant to skin and the environment, nontoxic to cells, and non-mutagenic. Hypochlorous acid is rated safe for transportation and storage, and does not require containment or ventilation. The cost of hypochlorous acid systems is significantly less than or comparable to other methods of disinfection, especially considering the risks inherent in the significant, sometimes even potentially catastrophic, disadvantages of the other methods[1].

Hypochlorous acid is not stable and breaks up into HCl and O_2 after a few hours; therefore it should be produced on site in an electrolytic cell from dilute brine. Due to the generation of HCl and HOCl in the anode and NaOH in the cathode it is necessary to use corrosion resistant materials for the cell and electrodes.

What is unique in this study is using an ultrasonic atomizer to produce a fine mist of the acidic solution and use the mist to disinfect buildings and devices. Ultrasonic atomizers use a piezoelectric disc to generate ultrasonic waves that cause capillary waves on the surface of the liquid which is being atomized. The frequency and amplitude of the vibration determine the size of droplets and the rate at which they are generated, respectively[2]. Figure 1 shows the formation of the droplets due to capillary waves of sufficient amplitude for atomization.

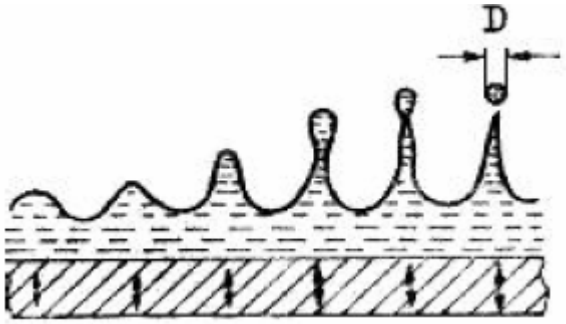


Figure 1: Droplet Generation by Capillary Waves[3]

The advantage of this type of humidification is that mist is generated without heating the liquid; therefore, the composition of the liquid does not change as a result of heating and the device consumes little amount of power. The smaller the disinfectant droplets are, the higher the disinfection power is. The droplets generated by ultrasonic humidifiers are smaller than 10μm in diameter. Thus this method is very effective for disinfection.

The mist can be sent into ventilation systems in order to safely disinfect buildings. In order to protect the vibrating disc a corrosion resistant alloy was used to cover and protect the disc.

Chemical Reactions

Production of hypochlorous acid is very similar to the chlor-alkali process used in the industry to produce caustic soda (NaOH) and chlorine. An electrolytic cell with an ion exchange membrane is used with brine as the electrolyte. The possible cathode reactions are:

$$Na^+ + e \rightarrow Na \qquad E^o = -2.71V$$

$$2H_2O + 2e \rightarrow H_2 + 2OH^- \qquad E^o = -0.83V$$

It is evident from the E^o values that water will be reduced in the cathode. The possible anode reactions are:

$$2Cl^- \rightarrow Cl_2 + 2e \qquad E^o = -1.36V$$

$$2H_2O \rightarrow O_2 + 4H^+ + 4e \qquad E^o = -1.23V$$

The oxidation potentials of water and chlorine are very close; so a combination of chlorine and oxygen gases might be expected in the anode; but the rate of oxidation of water is much slower than that of chlorine. Therefore in practice by using voltages higher than 1.36V mostly chlorine gas will be produced. Thus the net reaction happening in the cell is:

$$2NaCl + 2H_2O \rightarrow Cl_2 + H_2 \uparrow + NaOH_{(aq)}$$

Chlorine is highly soluble in water and dissolves in water according to the following reaction:

$$Cl_2 + H_2O \rightarrow HOCl + HCl$$

Figure 2 shows the percentage of dissolved Cl_2 in water as different compounds in terms of the pH of the solution. From this plot it is clear that the pH of the solution should be maintained near neutral pH conditions in order to maximize the percentage of HOCl present in the solution and maintain the safety of the mist.

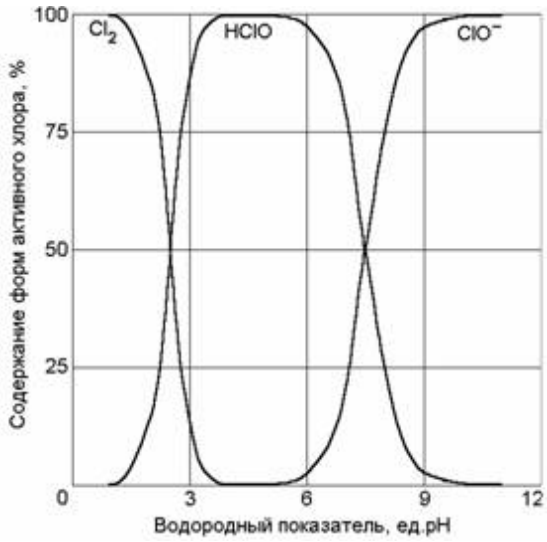


Figure 2: Percentage Composition of Dissolved Cl_2 Compounds vs. pH[4]

EXPERIMENTAL DESIGN

Figure 3 shows the general configuration of the electrolytic cell. The membrane in the cell is permeable to many cations and polar compounds but can almost completely reject anions and nonpolar species. Therefore Sodium ions will be trapped in the cathode compartment preventing them from reacting with the OCl^- ions to produce NaOCl which is not as strong a disinfectant as HOCl[5].

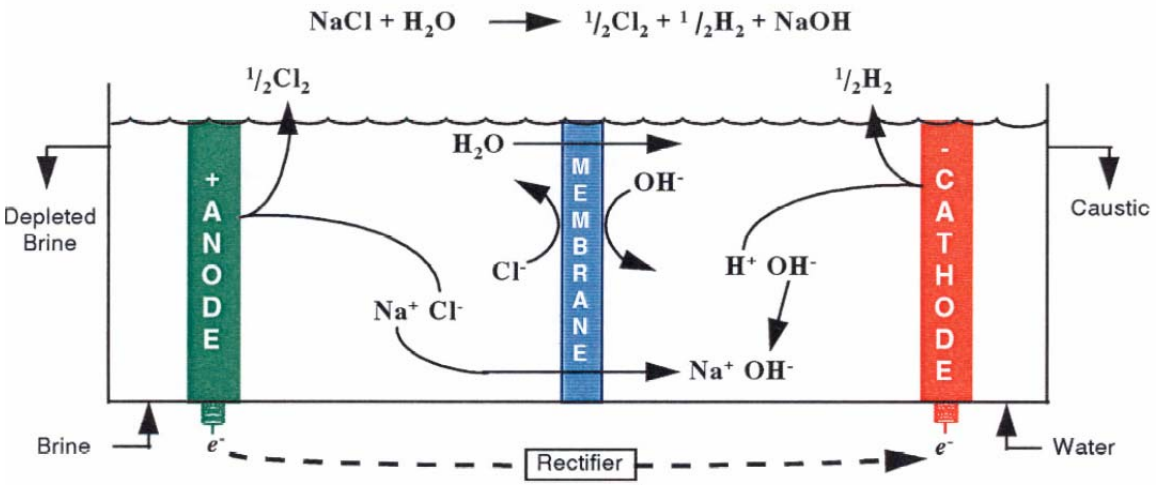


Figure 3: Generation of Chlorine and Caustic in a membrane electrolytic cell[6]

The cell container was made of PVC which has high corrosion resistance against all the chemicals present in the cell[7]. PVC sheets are also relatively cheap when compared to other corrosion resistant materials such as Teflon coated or corrosion resistant metals. The cell was divided into 2 compartments by a divider in the middle of the cell. The membranes were placed on windows on the divider. The membranes used in this experiment were Nafion 117 membranes. 7182D corrosion resistant sealant by Seals Eastern Inc. was used in order to prevent brine leakage between the two compartments. Figure 4 shows the divider design.

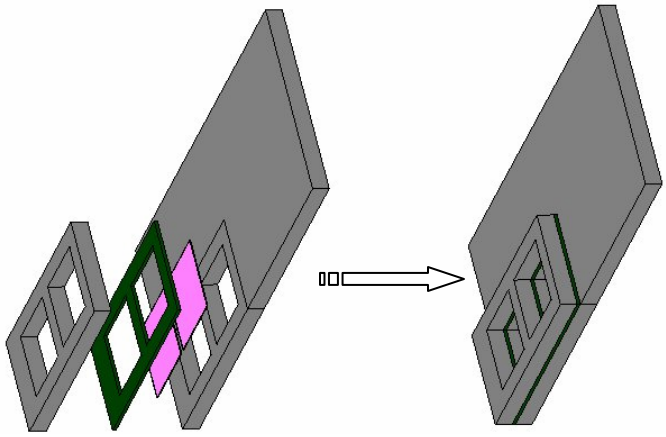


Figure 4: Divider layers from right to left: PVC, Nafion, Sealant and PVC Frame

Graphite electrodes were used due to their low price, availability, and acceptable corrosion resistance in the cell environment. To minimize the voltage needed for a given current the electrodes were placed 3cm apart from each other which was the minimum possible distance.

Multiple experiments were made using concentrations between 0.1% and 0.5% NaCl. In order to maximize the lifetime of the membranes and ensure product purity the brine solution was made from deionized water and pure salt.

The acid generated in the cell was transferred to the atomizing container. The acid was drained at the constant rate of 7ml/sec. So to maintain the concentration of the solution at 20mg/l the acid generation rate was adjusted by adjusting the current going through the cell. The current necessary to maintain the concentration of the solution was calculated by solving for $\frac{dm}{dt}$ in (1) and replacing it in (2) to find the current. The current was found to be 257mA.

$$C = \frac{m}{V} \Rightarrow \frac{dC}{dt} = \frac{V\frac{dm}{dt} - m\frac{dV}{dt}}{V^2} = 0 \quad \Rightarrow \frac{dm}{dt} = \frac{m\frac{dV}{dt}}{V} \tag{1}$$

Where C is the concentration of the liquid, m is mass of HOCl present in the solution and V is the overall volume.

$$m = \frac{Q}{F} \times A_{HOCl} = \frac{It}{F} \times A_{HOCl} \qquad \Rightarrow \frac{dm}{dt} = \frac{I}{F} \times A_{HOCl} \tag{2}$$

Where Q is the overall transferred charges, F is Faraday's constant (96485), t is time and A_{HOCl} is the molecular weight of HOCl.

It can be seen from (2) that the concentration of the solution is dependent on current and time. Therefore in order to be able to calculate the concentration at any time, it is best to use a constant current power supply to power the cell. The voltage required for 250mA was found to be 30V. The time necessary to reach the 20mg/l concentration before draining the cell can be calculated from (2).

A piezoelectric transducer running at 1.65MHz was used to atomize the solution. A water level sensor was used to ensure the presence of liquid in the container to protect the atomizer from burning. The mist generation rate is dependent on the amplitude of vibrations and the level of water above the atomizer. Figure 5 shows the mist generation rate in terms of the water level above the surface of a TDK atomizer. It can be seen from Figure 5 that the liquid level should be maintained at about 4cm to maximize the mist output. In order to maintain 4cm of liquid above the atomizer, a float sensor was used to prevent overfilling of the container.

The piezoelectric disc purchased for this experiment was a low power atomizer and could not atomize at the 7ml/sec rate. Therefore an extra container was built to store the extra acid that is generated after the acid level passes the 4cm level. This way the cell can be turned on long enough to fill both containers and then turned off. The storage container can supply the acid for the container with the atomizer for a few hours.

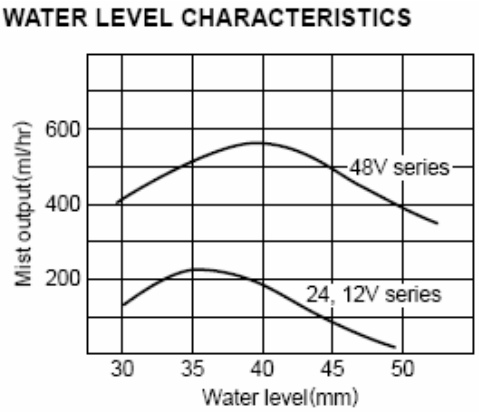


Figure 5: Water Level dependence of TDK Atomizers[9]

The vibrating disc of the atomizer was exposed to the acid and therefore the piezoelectric disc's original metal coating was removed and then covered with a 25μm thick Arnavar cobalt alloy strip which has very high corrosion resistant properties[8].

RESULTS

Figure 6 shows the device built according to the aforementioned design specifications. The top container is the electrolytic cell, the middle container is used for storage of the acid and the container with the float sensor is the atomizing unit. Figure 7 shows the mist being generated in the atomizing unit.

It was observed that the oxygen generated in the anode reacted with the graphite electrode. So after a certain duration of time the anode electrodes have to be replaced. The electrodes were then sputtered with Platinum to prevent the oxidation of graphite but experiments with the sputtered electrodes showed that they still oxidized. The oxidized Carbon could be readily seen peeling off the surface of the electrodes in small quantities after being placed in the electrolyte.

In the initial experiments only one electrode was used in the cathode and anode. The voltages required for a 0.1% solution to maintain the current at 60mA and 100mA were found to be 15V and 20V. When 4 electrodes were used at each side the voltage required for I=60mA and I=100mA dropped to 10.3Vand 15.1V, respectively. It was apparent that as the surface area of the electrodes increases, higher current can be generated from the same voltage level. In the industry, flat panel electrodes are used to maximize the surface area and minimize their distance[10].

Figure 6: Completed Device

Figure 7: Mist Generation

Figure 8 shows the pH of the cell when the products are not drained. It was observed that the rate of change of the cathodic and anodic pHs decreases as time passes. From (2) it is seen that chlorine generation rate is constant at a given current; therefore from the definition of pH it is clear that longer times are needed to lower the pH by 1 as pH of the solution decreases.

The voltage required to maintain the current drops as more products are generated. This is due to the increasing ion concentration in the cell which reduces the overall resistance of the liquid. Figure 9 shows the plot of the voltage required to maintain the 60mA current against elapsed time.

In order to generate the HOCl solution at the desired concentration, both the anode and cathode were drained. The cathode was drained at a slower rate to minimize the volume of the unwanted NaOH. The pH of the 20mg/l acid solution was 3.4; so some of the NaOH can be recycled to increase the pH of the solution. This will reduce the disinfecting efficiency of the acid due to generation of OCl^- ions. The advantages of adding NaOH to the solution are: decreasing the decomposition rate of HOCl, decreasing the corrosiveness of

the acid and ensuring the safety of the mist. Hence, overall it is desirable to increase the pH of the solution by adding some NaOH to the pure acid.

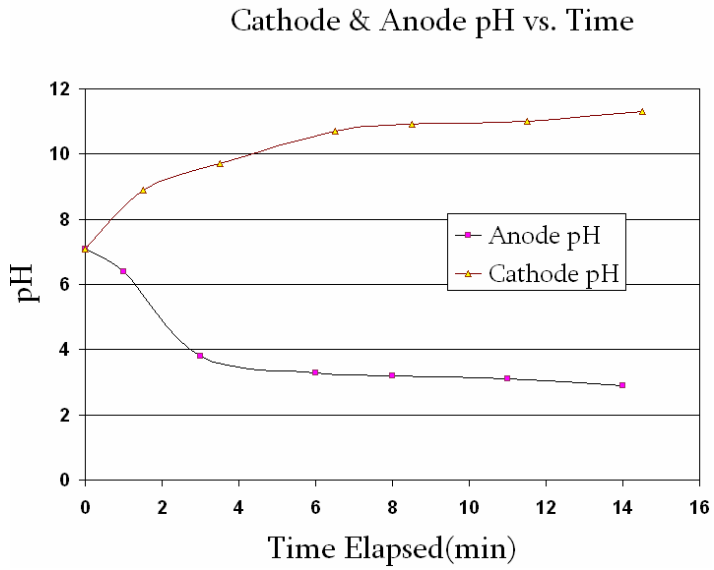


Figure 8: Cell Compartments' pH vs. Time

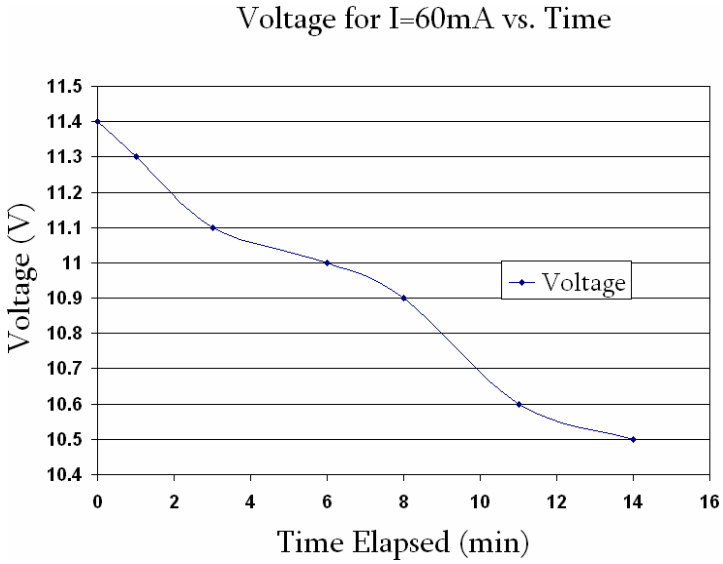


Figure 9: Cell Voltage vs. Time

The amount of HOCl present in the solution can be found by using free chlorine sensing kits and measuring the pH of the solution. Free chlorine is defined as the concentration of residual chlorine in water present as dissolved gas (Cl_2), hypochlorous acid (HOCl), and/or hypochlorite ion (OCl^-). Therefore knowing the amount of free chlorine present in the solution, the amount of

HOCl present can be calculated from Figure 2 by multiplying the percentage of HOCl at the pH of the solution by the amount of free chlorine. If the solution generated by the cell is not atomized immediately after generation, the aforementioned method should be used to ensure that enough free chlorine is present in the solution when the remaining solution is to be used.

The pH of the solution being atomized was also measured after being atomization. The pH of the pure acid was found to remain constant after 50 minutes of atomization. Figure 10 shows the graph of the measurements taken when NaOH was used to increase the pH of a 5mg/l acid from 4.6 to 5.3. The pH just changed by 0.1 after 40 minutes and remained constant for the next 30 minutes. Thus it seems reasonable to assume that the pH of the liquid is constant and therefore all the substances in the solution are present in the mist droplets.

The pH of the mist was measured by exposing pH papers to the mist. It was observed that the mist has the same pH as the solution. This indicates that the mist has the same chemical composition as the solution. More accurate measurements should be taken to verify our claim.

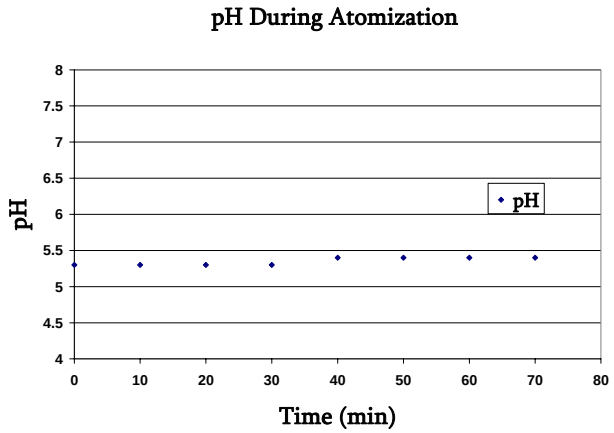


Figure 10: Acid Solution pH During Atomization with Added NaOH

Figure 11a shows the admittance of the original piezoelectric atomizer in terms of frequency when the atomizer is operating in open air. The resonance frequency of the disc in Figures 11a was 1.73MHz. It is seen that the driving circuit is driving the piezoelectric disc at a frequency close to the resonance frequency to maximize the power output. Figure 11b shows the admittance of the disc operating in open air after it was coated with the corrosion resistant coating. The resonance frequency of the coated disc was 1.74MHz which is

very close to the resonance frequency of the disc with the original coating. Thus the coating has no effect on the performance of the disc.

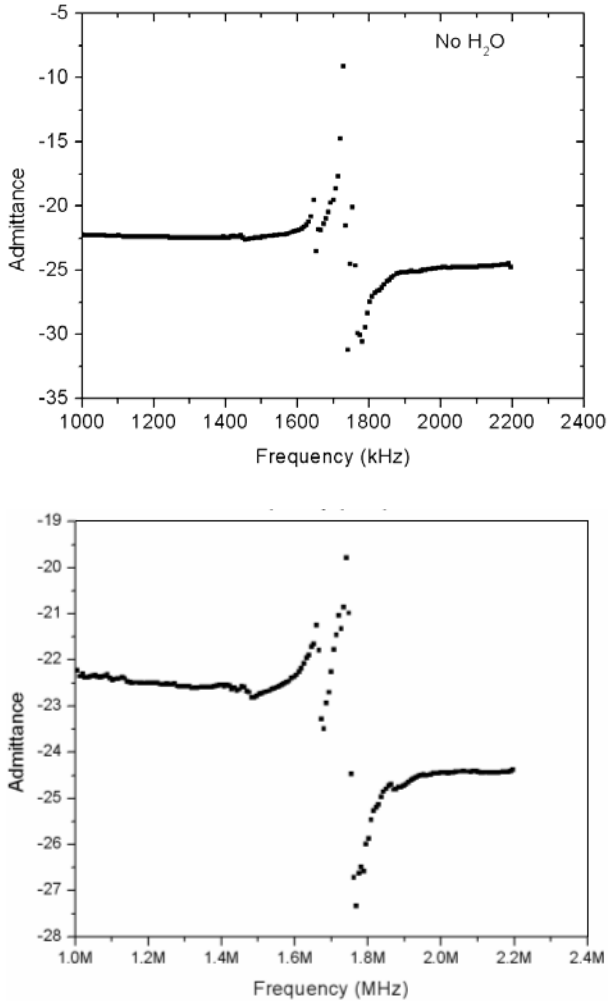


Figure 11: Admittance with a) Original Coating,
b) Admittance with Arnavar Coating

SUMMARY

A device that generates hypochlorous acid from brine in an electrolytic cell and atomizes the acidic solution for disinfection applications was successfully designed and built. This device can easily be adjusted to generate different concentrations of hypochlorous acid. Corrosion resistant materials were used to build the electrolytic cell and the other parts of the device to increase the lifetime.

The effects of atomization on the solution were analyzed by testing the pH of the solution. The value of pH was found to remain constant during

atomization. This indicates that the chemical properties of the solution and the mist are the same and all compounds dissolved in the solution are being atomized. Further tests need to be done to ensure that the mist has the desired disinfecting properties.

The effect of the corrosion resistant coating on the piezoelectric atomizer was also analyzed. The coating does not change the resonance frequency of the disc and has no other side effects on the operation of the disc.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. EEC-0244030.

I would like to thank everyone in ICAT, specially Mr. Seyit Ural and Mr. Hyeoung-Woo Kim for helping me get this project done on time and finding solutions to the problems I encountered during the 2 months. I would like to thank Mr. Ramanathan Ramna from chemical engineering department for providing me with the membranes and for his technical support. Thanks to Mr. Chris Jabco from MRL machine shop for making our professional-looking device. Finally, thanks to Arnold Magnetics and Seals Eastern Inc. for sending me free samples of their products.

REFERENCES

1 The Development of an Ultrasonic Hypochlorous Acidic Humidifier for Disinfecting Application, Yongkang Gao, Kenji Uchino, Private Publication

2 Lang RJ (1962) Ultrasonic atomization of liquids. J Acoust Soc Am 34:6–8

3 Picture Courtesy of Šarković D , Babović V (2003) Experiments on Ultrasonically Generated Aerosols: Results and Comments

4 Issues of chemical composition and operating properties of chlorine based inorganic liquid chemical germicides, V.M.Bakhir, B.I.Leonov, S.A.Panicheva, V.I.Prilutsky, N.Yu.Shomovskaya, The Russian Scientific & Research Institute for Medical Engineering (VNIIIMT MZ RF)

5 Gf and an Overview of Oxcide™: The Definitive Solution to Disinfection in Facility Water Distribution Systems & Equipment, Eric W Christensen, AIA, February 2003

6 Membrane cell figure courtesy of Dupont Nafion membrane product information sheet

7 http://www.harvel.com/downloads/chemical-resistance.pdf

8 http://www.arnoldmagnetics.com/products/rolled/pdf/strpfoil.pdf

9 http://www.tdk.com.hk/english/pdf/Piezo/Ultrasonic/eF441_NB.pdf

10 Industrial electrochemistry, D. Pletcher, Chapman and Hall, 1982

NONLINEAR BEHAVIOR OF MAGNETOELASTIC SENSORS: MODELING, SIMULATION, AND EXPERIMENT

Kong C. Tep* and Craig A. Grimes[#]

Department of Electrical Engineering
The Pennsylvania State University, University Park, PA 16802

*Undergraduate student of
Department of Electrical Engineering
Michigan State University
East Lansing, MI 48823

ABSTRACT

This work investigates the nonlinear behavior of the magnetoelastic sensors using numerical simulations with results compared to measured experimental data. Magnetoelastic sensors are generally modeled as a mass-spring-friction damped vibration system, which is a linear time invariant (LTI) system described by an impulse response in time domain, a transfer function in S-domain, and a frequency response in frequency domain. In practice, however, the sensors can be over-driven to nonlinear vibrations. There are a number of factors that contribute to the nonlinear vibrations. In this work the nonlinear behavior is simulated by modeling the spring constant as a function of displacement. The modeling and simulation are experimentally verified using a microcontroller-based instrument.

INTRODUCTION

Magnetoelastic sensors belong to the class of resonator sensors and are typically made of amorphous ferromagnetic ribbons cut into rectangular thick-film beams. Intensive research into magnetoelastic sensor technology, including sensor application, sensor characterization, and sensor instrumentation, has been carried out over the past few years in Grimes research group at Penn State University. Preliminary results have shown that magnetoelastic sensors offer an

---

[#] Faculty Mentor

attractive sensor platform for chemical sensing, biological sensing, and environmental monitoring not only due to their low cost and relative small size but also because of their wireless, passive, and remote query nature [1-16].

The sensors are generally modeled as a mass-spring-friction second order damped vibration system, described by a second order differential equation

$$m\frac{d^2y}{dt^2} + c\frac{dy}{dt} + ky = kx \tag{1}$$

where m is the mass, c the friction, and k the spring constant. Equation (1) describes a linear time invariant (LTI) system. By solving it we can obtain an impulse response $h(t)$ in time domain, a transfer function $H(s)$ in S domain, and a frequency response $H(j\omega)$ in frequency domain, given by

$$h(t) = \frac{\omega_n^2}{\omega_d}\exp(-\zeta\omega_n t)\sin(\omega_d t)u(t) \tag{2}$$

$$H(s) = \frac{{\omega_n}^2}{s^2 + 2\zeta\omega_n s + {\omega_n}^2} \tag{3}$$

$$H(j\omega) = H(s)\big|_{s=j\omega} = \frac{1}{1-(\omega/\omega_n)^2 + j2\zeta(\omega/\omega_n)} = |H(\omega)|\exp(-j\phi(\omega)) \tag{4}$$

where $\omega_n = \sqrt{\frac{k}{m}}$ is natural frequency of undamped oscillation, $\zeta = \frac{c}{2m\omega_n}$ is the damping ratio, $\omega_d = \omega_n\sqrt{1-\zeta^2}$ is the frequency of the damped oscillation, and $u(t)$ is the unit step function.

The characterization of a system is essentially the analysis of the system response to applied stimuli, or excitations. In time domain, the system response $y(t)$ to an excitation $x(t)$ is calculated using the impulse response $h(t)$ by

$$y(t) = x(t) * h(t) = \int_0^t A\sin(\omega(t-\tau))\cdot\frac{\omega_n^2}{\omega_d}\exp(-\zeta\omega_n\tau)\sin(\omega_d\tau)u(\tau)d\tau \tag{5}$$

In frequency domain, we have

$$Y(\omega) = H(j\omega)X(\omega) \tag{6}$$

Nonlinear behavior has been experimentally observed at relatively large excitations. As the excitation gets stronger and stronger, the spectrum starts to show multiple peaks. This work investigates nonlinear behavior using numerical simulations based on the mathematical model for spring constant k.

EXPERIMENTAL DESCRIPTION

MATLAB Simulations

The Grimes research group has developed a microcontroller-based magnetoelastic instrumentation system capable of both time domain analysis and frequency domain analysis using a threshold-crossing counting technique. In the threshold-crossing counting technique, a sensor is excited by a sine wave pulse and relevant information is extracted by counting the number of threshold-crossings of the transient oscillation [17]. The technique is numerically simulated using MATLAB in this work.

For a nonlinear system, we cannot obtain a close-form impulse response $h(t)$, and hence we cannot calculate the sensor response using Equation (5). Instead, we can numerically calculate the sensor response using Euler's method

$$y_n = y'_n \times \Delta t + y_{n-1} \tag{7}$$

$$y'_n = y''_n \times \Delta t + y'_{n-1} \tag{8}$$

where Δt is the time step, y' and y" are the first and second derivative, respectively. Using this method we can estimate the response of the sensors at any excitation amplitude based on the initial condition by iterative calculation process.

In the time-domain simulation, we first calculate the slope y" using Equation (1). Based on y", the first derivative y' and the displacement y can be calculated using the Euler method through Equation (7) and (8). The calculated result y is then stored into a vector [Y]. This iterative process continues until t reaches the predefined limit. Figure 1 shows the flowchart of the time-domain simulation. Frequency sweep is performed over a predefined frequency range. Basically, the sensor is subjected to different excitation frequency over the range. At every frequency, the time-domain simulation method is used to calculate the response. The vector [Y] that was obtained is used for the threshold counting technique. As shown in Figure 2, the number of threshold crossings is then stored into another vector [Z]. Similar to the time-domain method, another iterative process is implemented for the frequency sweeping. When the frequency f reaches the end of the frequency range, the Euler method stops generating vector [Y]. MATLAB then plots the number of threshold crossings in terms of the frequency.

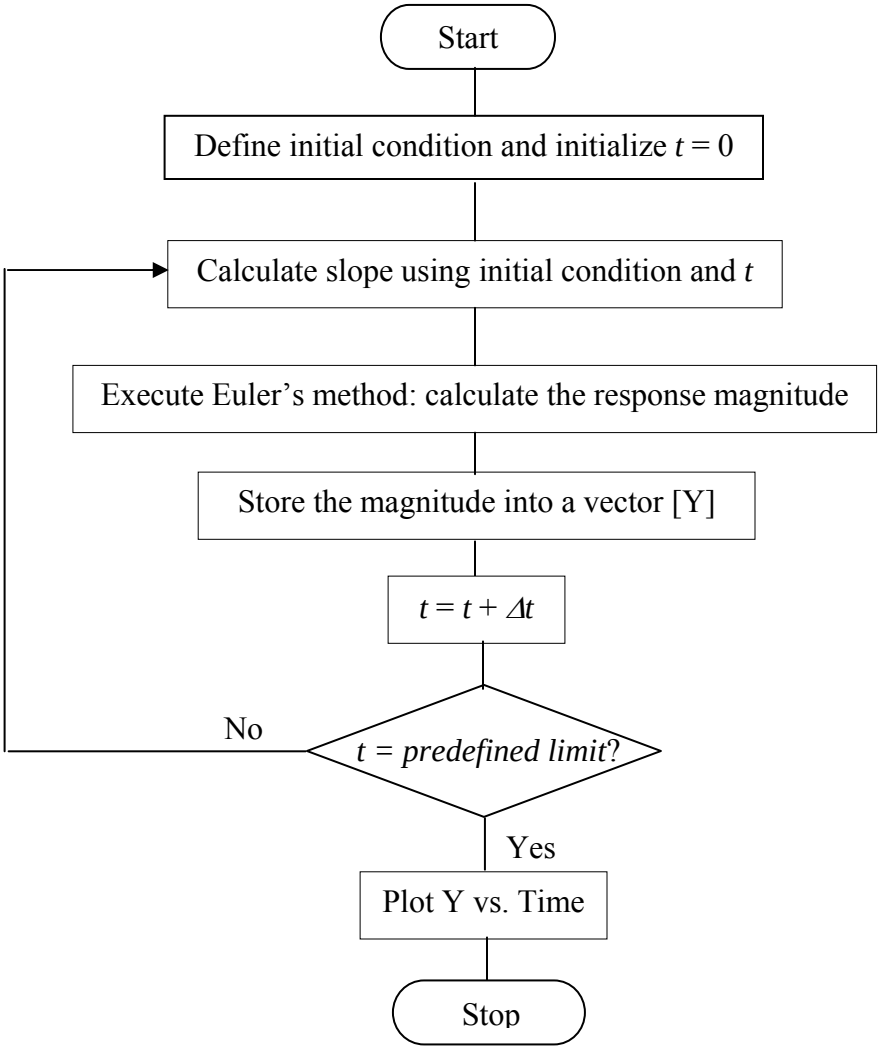


Figure 1: Flowchart of MATLAB simulation using the Euler method.

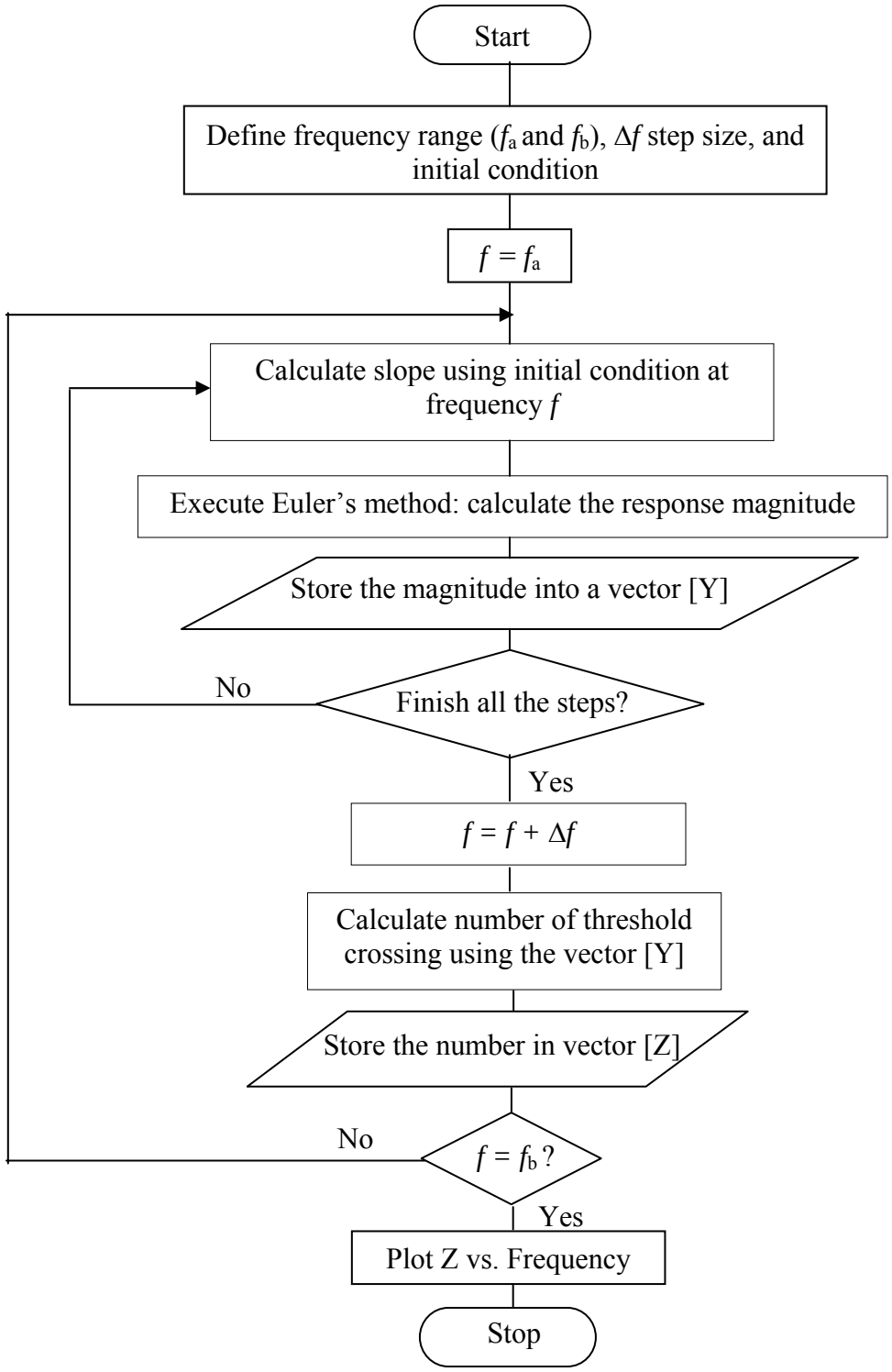


Figure 2: Flowchart of the frequency sweep using the threshold-crossing counting technique and the Euler method.

DATA AND RESULTS

Time-Domain Simulations

As mentioned earlier, Euler's method is used to calculate a sensor's response to an external excitation. Figure 3 shows the simulated sensor response to a sine wave pulse excitation at the resonance frequency. An experiment was done to verify the modeling and simulation using the microcontroller-based instrument. In the experiment, a 220 kHz sensor was excited at its resonance frequency and the response was captured by a digital oscilloscope, as shown in Figure 4. Comparing Figure 3 and Figure 4 we can tell that the simulation is accurate.

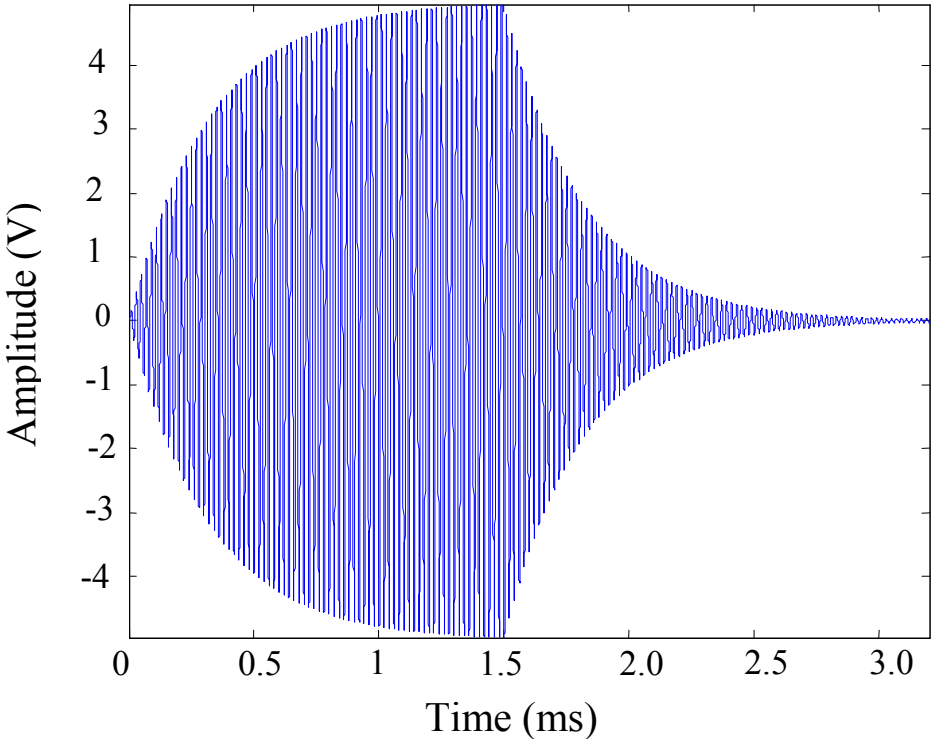


Figure 3: Sensor response to a sine wave excitation calculated using Euler's method.

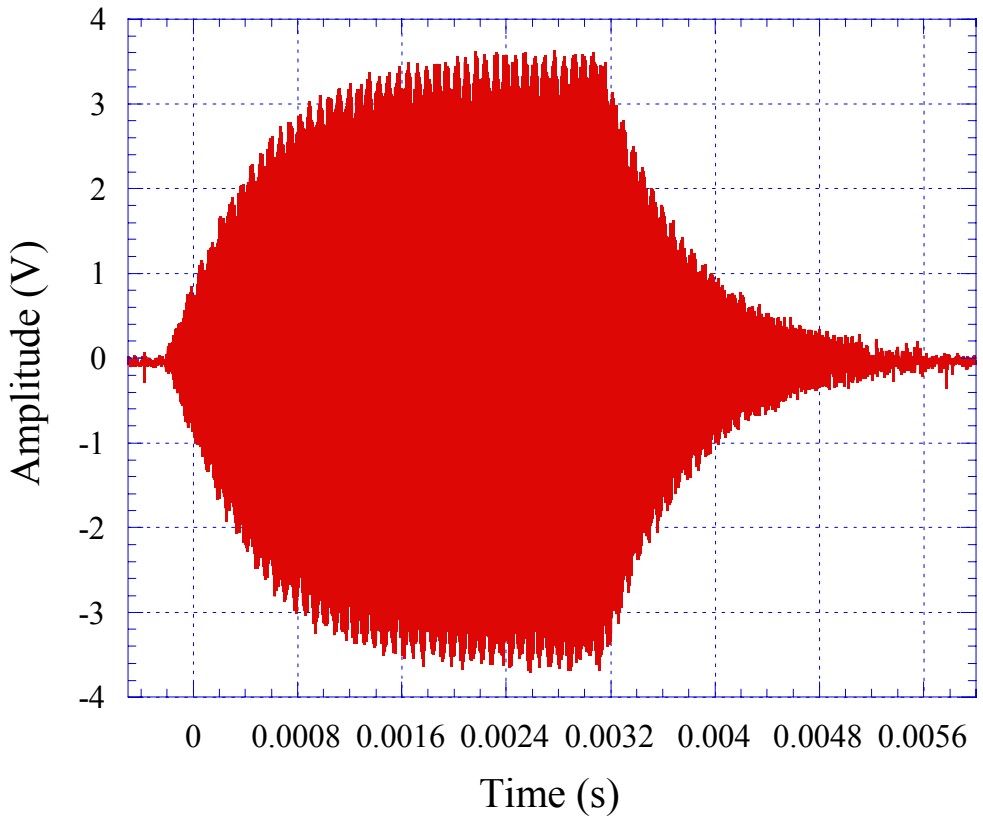


Figure 4**:** Oscilloscope-captured sensor responses to a sine wave pulse excitation.

Frequency-Domain Simulations

In the microcontroller-based magentoelastic sensor instrumentation system, a frequency response plot is obtained by performing a frequency sweep over a predefined frequency range normally centered at resonance frequency of the sensors. The simulations use the same technique for spectrum analysis. Figure 5 shows the simulated spectrum of a linear model. The spectrum has only a single peak approximately located at the resonance frequency. In real experiments, the peak is split when the excitation is large, as shown in Figure 6. In order to model the nonlinear behavior, the spring constant k in Equation (1) is modeled as a function of displacement by

$$k = a + by^2 - cy^4 + dy^6 - ey^8 \tag{9}$$

where a, b, c, d, and e are constants. With this k model, a similar spectrum is obtained using MATLAB simulations, as shown in Figure 7. The experiment was repeated with different excitation amplitudes and the different spectrums are shown in Figure 8, with the corresponding simulation results shown in Figure 9.

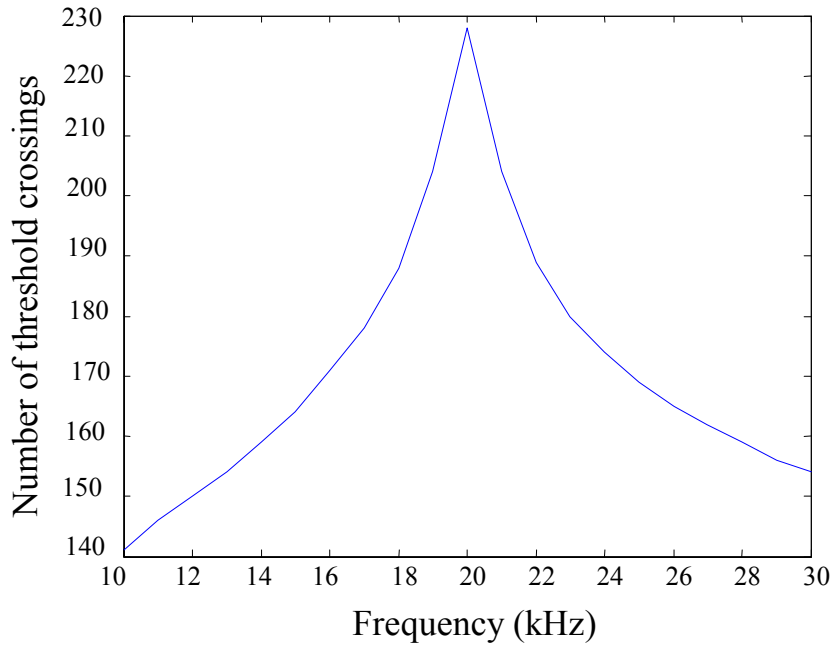


Figure 5: Simulated spectrum for a sensor using a linear model.

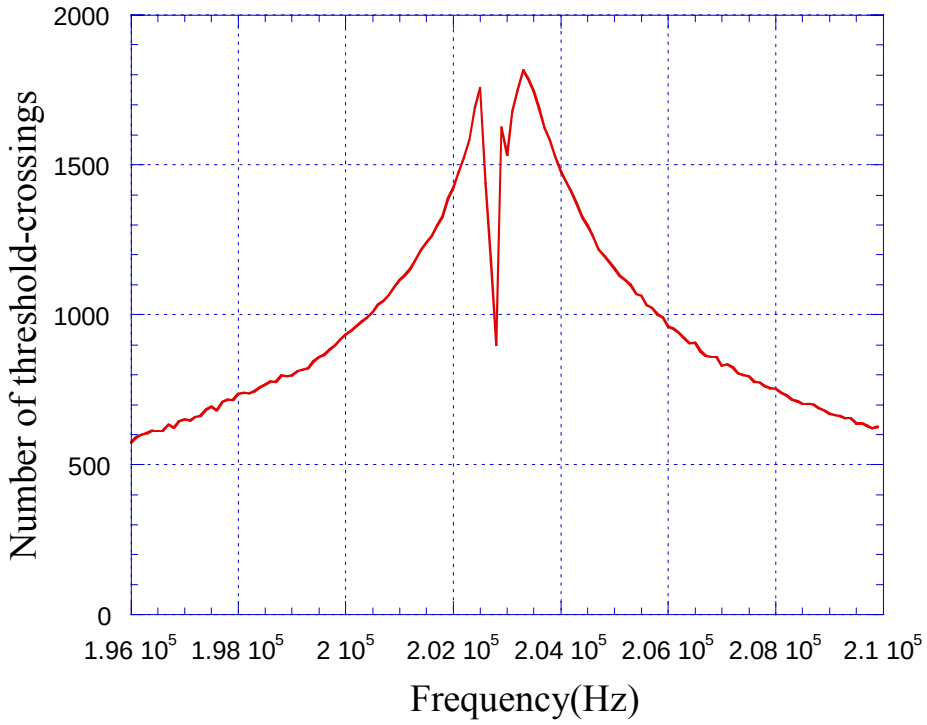


Figure 6: Measured spectrum of a sensor when the sensor is over-driven.

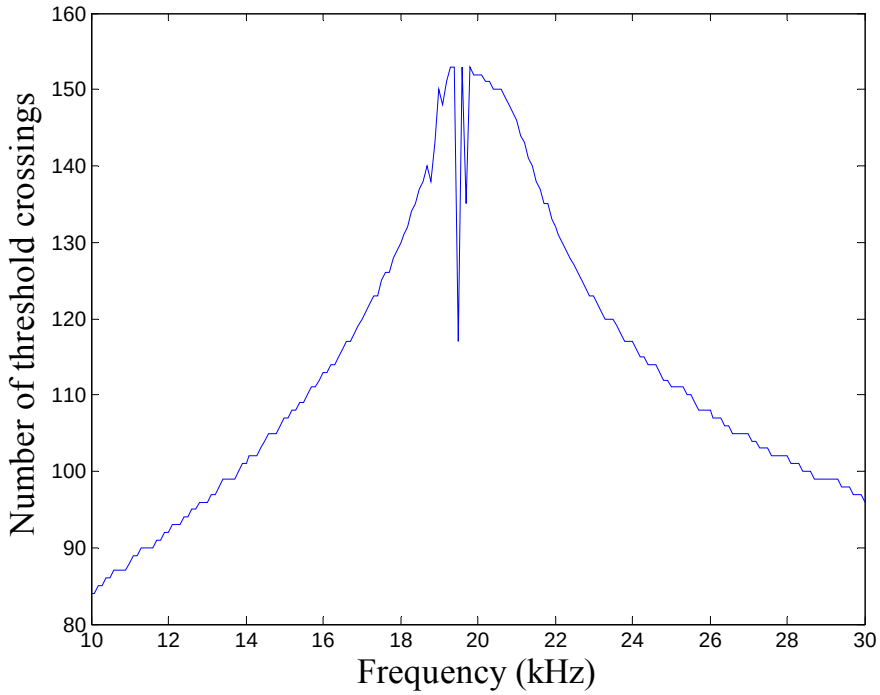


Figure 7: Simulated spectrum of a sensor using a nonlinear model.

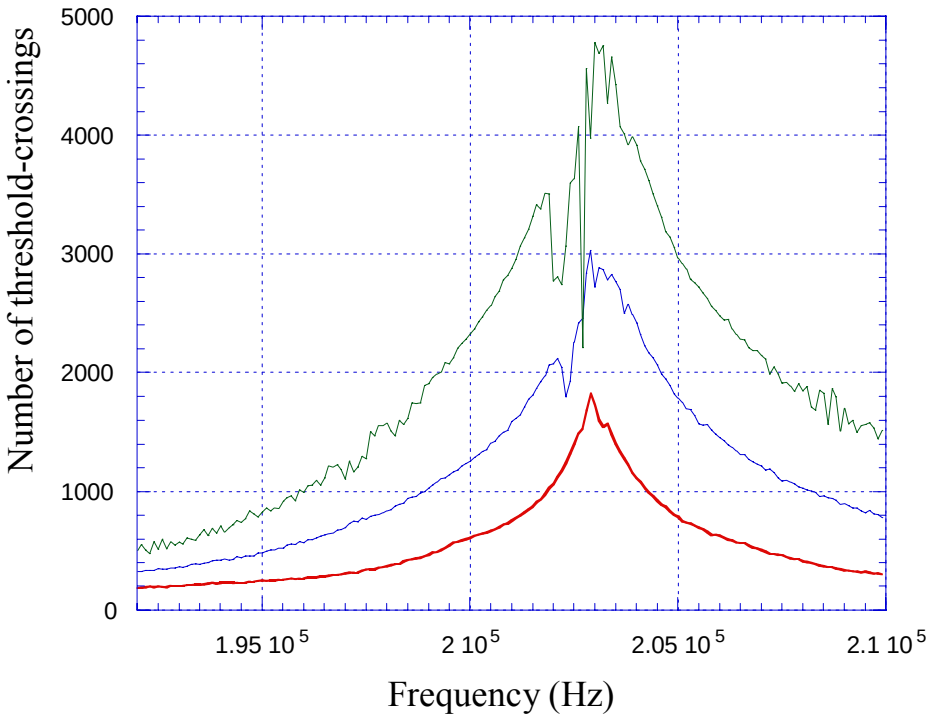


Figure 8**:** Measured sensor response to increasing excitation force**.**

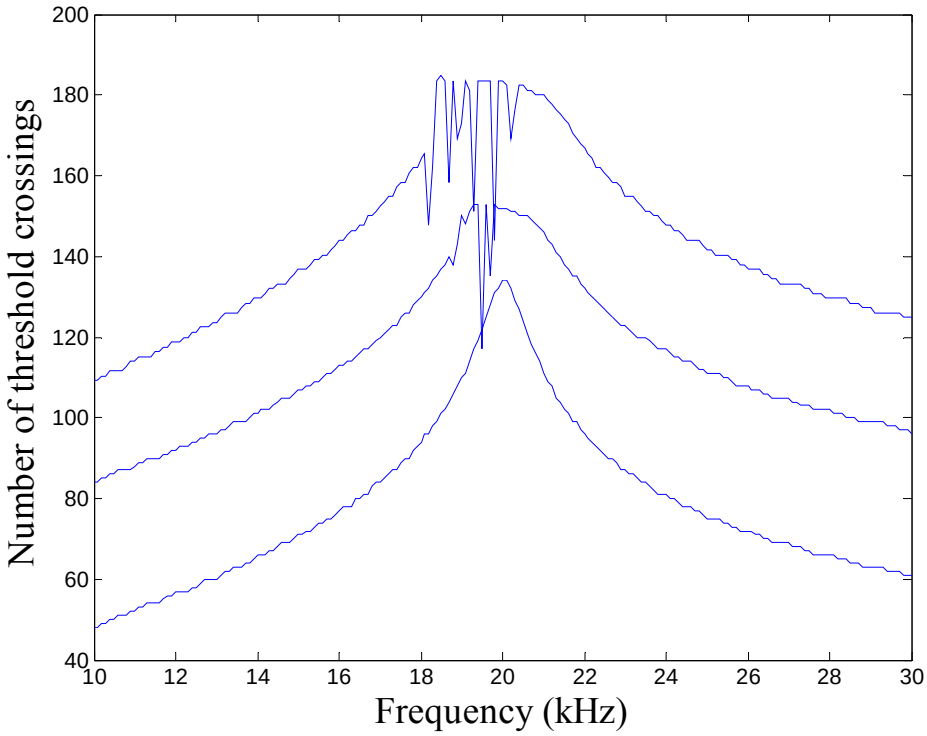


Figure 9**:** Simulated sensor response to increasing excitation force.

DISCUSSION

In this work, we consider the spring constant k is the major factor that causes the nonlinear resonance behavior. As shown in Equation (9), for small vibration amplitudes, the effect of displacement y on k is negligible and the vibration is almost linear. As the vibration amplitude increases, the effect of displacement y on k is no longer negligible and the sensor starts to experience nonlinear vibrations. From Figure 8 and Figure 9 we can see both experiments and simulations show the similar trend as the excitation increases. There are also other factors that contribute to the nonlinear resonance behavior, including friction and the distortion of excitation signal itself. The investigations of these other factors are left to the future work.

Euler's method proved to be very useful for sensor response calculation when close-form impulse response cannot be obtained. The accuracy of the method depends on the time step Δt. Higher accuracy requires longer simulation time.

SUMMARY

The nonlinear behavior of magnetoelastic sensors is investigated by comparing MATLAB theoretical simulations to the measured experimental results. The sensor response to an applied excitation is calculated using Euler's

method instead of the system impulse response in order to simulate the nonlinear behavior. Both time-domain and frequency-domain simulations are carried out and experimentally verified using a microcontroller-based magnetoelastic instrumentation system.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. EEC-0244030. The assistance of Dr. Zeng and Casey Mungle with this project is also gratefully acknowledged. Special thanks go to Dr. Grimes, Dr. Guo, and the Electrical Engineering Department of Penn State University for providing this research opportunity.

REFERENCES

[1] C.A. Grimes, K. G. Ong, K. Loiselle, P.G. Stoyanov, D. Kouzoudis, Y. Liu, C. Tong, and F. Tefiku, "Magnetoelastic Sensors For Remote Query Environmental Monitoring," *Journal of Smart Materials and Structures,* **8** 639-646 (1999)

[2] M. K. Jain, S. Schmidt, and C. A. Grimes, "Magneto-Acoustic Sensors for Measurement of Liquid Temperature, Viscosity and Density," *Applied Acoustics,* **62** 1001-1011 (2001)

[3] M. K. Jain, C. A. Grimes, "A Wireless Magnetoelastic Micro-Sensor Array for Simultaneous Measurement of Temperature and Pressure," *IEEE Transactions on Magnetics,* **37** (4) 2022-2024 (2001)

[4] S. Schmidt, C. A. Grimes, "Elastic Modulus Measurement of Thin Films Coated onto Magnetoelastic Ribbons," *IEEE Transaction on Magnetics,* **37** 2731-2733 (2001)

[5] A. Ersoz, J. C. Ball, C. A. Grimes, L. G. Bachas, "Magnetoelectrochemistry: Characterization of Electrochemically Deposited Polypyrrole by Magnetoelastic Transduction," *Anal. Chem.*, **74** 4050-4053 (2001)

[6] M. K. Jain, S. Schmidt, C. Mungle, K. Loiselle, C. A. Grimes, "Measurement of Temperature and Liquid Viscosity Using Wireless Magneto-Acoustic/Magneto-Optical Sensors," *IEEE Transactions on Magnetics*, **37** (4) 2767-2769 (2001)

[7] S. Schmidt and C. A. Grimes, "Characterization of Nano-dimensional Thin Film Elastic Moduli Using Magnetoelastic Sensors," *Sensors & Actuators A*, **94** 189-196, (2001)

[8] M. K. Jain and C. A. Grimes, "Effect of Surface Roughness on Liquid Property Measurements Using Mechanically Oscillating Sensors." *Sensors and Actuators A,* **100** 63-69 (2002)

[9] K. G. Ong, C. A. Grimes, "The Higher-Order Harmonics of a Magnetically-Soft Sensor: Application to Remote Query Temperature Measurement," *Applied Physics Letters*, vol. **80 (**20) 3856-3858 (2002)

[10] K. G. Ong, C. A. Grimes, "Tracking the Harmonic Response of Magnetically-Soft Sensors for Wireless Temperature, Stress, and Corrosive Monitoring." *Sensors and Actuators A,* **101** 49-61 (2002)
[11] C. A. Grimes, P. G. Stoyanov, D. Kouzoudis and K. G. Ong, "Remote Query Pressure Measurement Using Magnetoelastic Sensors," *Review of Scientific Instruments*, **70** (12) 4711-4714 (1999)
[12] Q. Y. Cai, C. A. Grimes, "A Wireless, Remote Query Ammonia Sensor," *Sensors and Actuators B,* **77** 614-619 (2001)
[13] P. G. Stoyanov, C. A. Grimes, "A Remote Query Magnetostrictive Viscosity Sensor," *Sensors and Actuators A* **, 80** 8 – 14 (2000)
[14] C. A. Grimes, C. S. Mungle, K. Zeng, M. K. Jain, W. R. Dreschel, M. Paulose, and K. G. Ong, Invited Paper: "Wireless Magnetoelastic Resonance Sensors: A Critical Review," *Sensors*, **2** 289-308 (2002)
[15] C. Ruan, K. Zeng, C. A. Grimes, "A Mass-sensortive pH Sensor Based on A Stimuli-responsive Polymer, Analytica Chimica ACTA 497" 123-131 (2003)
[16] C. Ruan, K. Zeng, O. K. Varghese, C. A. Grimes, "Magnetoelastic Immunosensors: Amplified Mass Immunosorbent Assay for Detection of *Escherichia coli* O157:H7" *Analytical Chemistry*, **75** 6494-6498 (2003)
[17] K. Zeng, *Magnetoelastic Sensors and Sensing Systems: Sensor Characterization and Instrumentation*, Ph.D.Thesis, EE, Penn State (2004)

CHARACTERIZATION OF THE DYNAMIC BEHAVIOR OF LOSS MECHANISMS IN THIN-FILM HIGH TEMPERATURE SUPERCONDUCTING RESONATORS

Himani Suhag* and Jeffrey L. Schiano[#]

Department of Electrical Engineering
The Pennsylvania State University, University Park, PA 16802

*Undergraduate student of
Electrical, Computer, & Systems Engineering
Rensselaer Polytechnic Institute
Troy, NY 12180

ABSTRACT

Research is underway to detect concealed explosives using quadrupole resonance (QR). In order to improve the signal-to-noise ratio (SNR) of QR measurements, we are developing QR probes based on high-temperature superconducting (HTS) resonators. During excitation of the QR response, it was observed that the Q-factor of the superconducting QR (SQR) probe is dependent on the excitation field. We review literature describing the effect of the excitation field on the resonator quality factor and present nonlinear models for describing the dynamic behavior of the resonator current during pulse excitation. These results will guide future development of SQR probes.

INTRODUCTION

Following September 11, there has been a dramatic increase in the interest to prevent terrorist attacks dealing with biological weapons and explosives. Currently, airports use X-ray machines to inspect baggage for contraband. Unfortunately, these machines cannot detect plastic explosives, which present an enormous threat to our nation's security. QR detection systems can detect threat quantities of explosives contained in sealed packages and are immune to false alarms caused by trace quantities. Unlike X-ray systems, QR systems are capable of detecting sheet explosives. The downside of this method is the low SNR

[#] Faculty Mentor

of QR measurements that results in unacceptably large detection times. The SNR is limited in part by the Quality (Q) –factor of the QR probe. It is known that the SNR is proportional to the square root of the Q-factor. [1] While normal metals coils have Q-factors on the order of 10^2, thin-film high-temperature superconductors (HTS) resonators have Q-factors on the order of 10^5 and therefore provide a significant enhancement in SNR. In order to design a QR detection system using HTS resonators, a mathematical model of the resonator is needed. Unlike normal metal resonators, the Q-factor of HTS resonators is dependent on the amplitude of the magnetic field seen by the resonator. [2-8] As a result, the HTS resonator cannot be adequately modeled by a series RLC circuit with a fixed resistance value. The objective of this research is twofold. First, using literature that quantifies loss mechanisms in HTS resonators, a nonlinear lumped parameter model that describes the dependence of resistive losses on magnetic field amplitude is presented. Second, using experimental data, a least-squares estimate of the model parameters is obtained.

Q-Factor Dependence on Magnetic Field

The unloaded Q-factor, Q_o, of a resonator is defined as the ratio of the maximum energy stored per cycle over the energy dissipated per cycle [9]

$$Q_o = 2\pi \frac{\text{max energy per cycle}}{\text{energy dissipated per cycle}} \tag{1}$$

Unloaded means there are no external losses electrically or magnetically coupled to the circuit. The losses within the resonator include resistive losses, dielectric losses, and radiation losses. [10] An equivalent definition of Q-factor, derived from the frequency response characteristics of the resonator, is equal to the center frequency divided by the 3 dB bandwidth. [11]

In a QR detection system, the HTS resonator is magnetically coupled to a low noise preamplifier. The magnetic coupling introduces external losses in the HTS resonator, thereby reducing its Q-factor to a value less than Q_o. Of particular interest is the case of critical coupling, where the losses coupled into the HTS resonator reduce Q_o by a factor of two

$$Q_L = \frac{1}{2} Q_o, \tag{2}$$

where Q_L is the loaded Q-factor at critical coupling.

Techniques available for measuring the Q-factor of a resonator can be categorized as either frequency-domain or time-domain. The frequency-domain technique used in this research consists of measuring the frequency response of the resonator at critical coupling by using a network analyzer to measure the S_{21} parameter. [12] In this method, the network analyzer slowly sweeps the excitation

frequency across the resonance peak and therefore is considered to be a continuous-wave (CW) excitation measurement technique.

The time-domain approach uses a RF pulse to excite the HTS resonator.[13] The frequency of the gated sinusoid matches the resonant frequency of the HTS resonator, while the width of the pulse is chosen so that the response of the HTS resonator achieves a steady-state value before the pulse is turned off. The exponential rise and decay of the resonator response is then used to determine the Q-factor of the resonator. In comparison to the CW excitation procedure, the transient approach requires significantly less time. For example, while the width of the RF pulse is typically on the order of milliseconds, an S_{21} measurement may take minutes. This large disparity in time is because the sweep rate of the network analyzer is limited so that the measurement bandwidth is less than the width of the resonator peak.

Figure 1 shows the unloaded Q-factor of an HTS resonator obtained using S_{21} and transient response measurements.[14] The vertical axis presents Q_o while the horizontal axis displays the excitation power (P_e) used to make the measurement. In these measurements, the excitation loop is critically coupled to the HTS resonator. There are two important features of these results. First, both the frequency-domain and the time-domain methods are in agreement, and second, Q_o is a function of P_e.

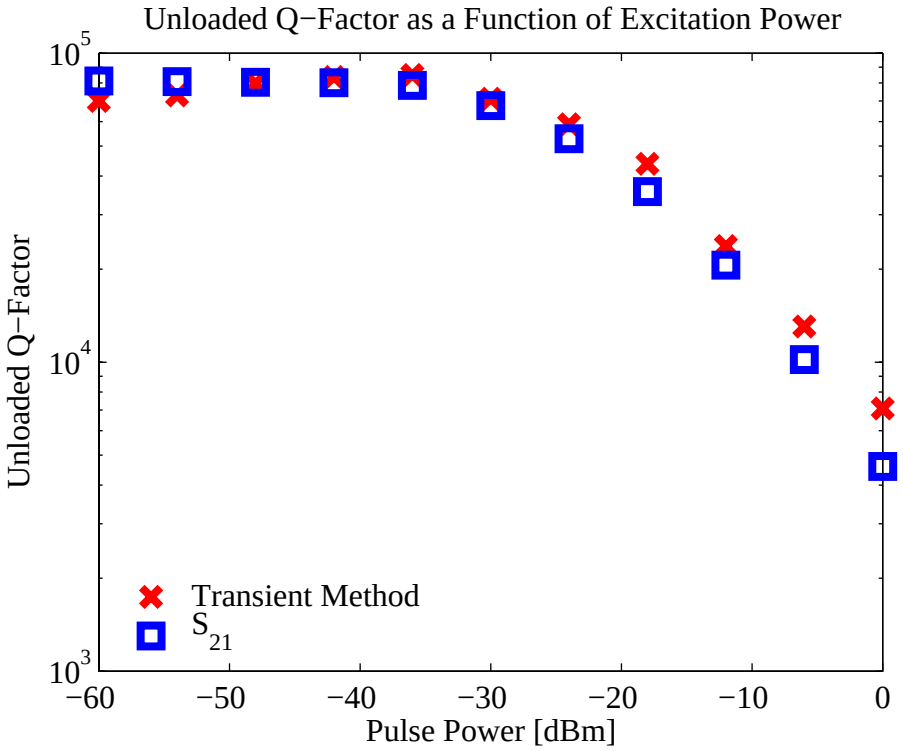


Figure 1: Graph of CW and pulsed measurements of Q-Factor.

Lumped Parameter Model

At frequencies of interest in QR detection systems, the wavelength is large compared to the dimensions of the HTS resonator. Therefore, the lumped parameter model shown in figure 2 is used to describe the resonator. The capacitance C and the inductance L are fixed by the geometry and permittivity of the HTS resonator. In a normal metal resonator, the loss mechanisms are represented by the resistance R, and Q_o, derived from equation (1), is

$$Q_o = \frac{w_n L}{R}, \tag{3}$$

where $w_n = \frac{1}{\sqrt{LC}}$ is the natural frequency of the resonator. However, for HTS resonators, it is a well known fact that R is dependent on the RF magnetic field observed by the resonator.

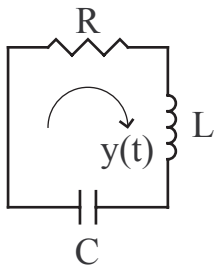


Figure 2: Lumped parameter model of the HTS resonator.

There has been a significant effort to quantify the effect of the RF field on the losses of the HTS resonator.[3,6-7,15-21] This work has led to several empirical expressions for representing R as a function of the RF magnetic field. Polynomial expansions of the form

$$R(H_{rf}) = R_o + c_1 H_{rf} + c_2 H_{rf}^2 \tag{4}$$

were presented by Nyugen [3] and Oates. [22] On the other hand, Jacob [21] and Wilker [8] used exponential representations of the form

$$R(H_{rf}) = R_o + c_1 e^{c_2 H_{rf}^{c_3}} \tag{5}$$

In these expressions, H_{rf} is known and R is indirectly determined by measuring the Q-factor [8] or current distribution in the HTS resonator. [23] The parameters R_o and c_i are obtained from a least-squares fit. As the loss mechanisms in HTS

resonators are not well understood, theoretical models do not provide a good fit to the experimental observations, and therefore are not considered in this paper.

In our experiment, the HTS resonator is magnetically coupled to an excitation source and RF preamplifier. For simplicity, this arrangement is modeled by the series RLC circuit shown in figure 3. Note that R is a function of the current y because y is proportional to H_{rf}. The ODE model describing the circuit shown in figure 3 is

$$\frac{d^2y}{dt^2} + \frac{R(y)}{L}dy + \frac{1}{LC}y = \frac{1}{L}f \ , \tag{6}$$

where the input excitation is represented by the voltage source f(t) and the output response is the current y(t). The input f(t) is a gated sinusoid whose frequency is set to the natural frequency of the HTS resonator. The response y(t) is a sinusoidal pulse with exponentially rising and falling edges. The width of the input is chosen sufficiently large so that y(t) reaches a sinusoidal steady-state response before the falling edge. As the Q-factor is large, the time constants associated with the exponential response are much larger than a single period of the sinusoid. For this reason, to simplify the estimation of the parameters c_i, we only consider the envelope, $y_e(t)$, of the response.

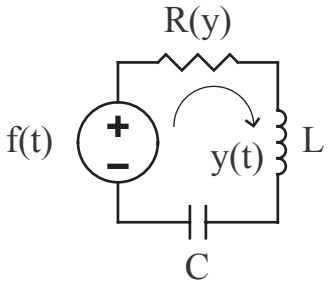


Figure 3: Incorporation of power dependent losses in the HTS resonator model.

A first-order ODE describing the envelope of the response is obtained by taking two facts into consideration. First, the exponential rise and fall of the envelope are set by the roots of the characteristic equation for the ODE in equation (6). Second, when this system is driven at its resonant frequency, the sinusoidal steady-state gain can be found from the frequency response function

$$H(jw) = \frac{Y(j\omega)}{F(j\omega)} = \frac{\frac{1}{L}(j\omega)}{(j\omega)^2 + \frac{R(y_{ss})}{L}(j\omega) + \frac{1}{LC}} \tag{7}$$

where y_{ss} is a constant representing the value of the sinusoidal steady-state current in the HTS resonator. At the natural frequency, the gain is

$$H(j\omega_n) = \frac{1}{R(y_{ss})} \tag{8}$$

The resulting first-order ODE for the envelope y_e of the response is

$$\frac{dy_e}{dt} + \frac{1}{\tau} y_e = kF_e \frac{1}{\tau R(y_e)}, \tag{9}$$

where F_e takes on the value of 1 or 0 to indicate the presence or absence of the excitation pulse, respectively, and k is a constant parameter to be estimated that takes into account the coupling between the HTS resonator and the excitation source. The time constant τ governs the exponential rise and fall of the envelope and is derived from the real part of the roots of the characteristic equation

$$\tau = \frac{2L}{R(y_e)} \tag{10}$$

In addition, the dc gain of equation (9) must match the sinusoidal steady-state gain of the second-order system at the natural frequency. For this reason, the forcing function includes the term $\frac{1}{R(y_e)}$. Equation (10) is written as

$$\frac{dy_e}{dt} = -\frac{R(y_e)}{2L} y_e + \frac{k}{2L} F_e \tag{11}$$

in order to estimate values for $\frac{R}{L}$, $\frac{c_i}{L}$ and $\frac{k}{2L}$.

EXPERIMENTAL DESCRIPTION

The aim of the experimental work is to determine whether or not equation (11) can describe the measured response of the HTS resonator to pulse excitation, using either the polynomial or exponential dependence of the resonator loss on the magnetic field given by equations (4) and (5) respectively. To accomplish this task, experimental measurements are used to obtain least square estimates of the model parameters.

Setup

Figure 4 shows the block diagram of the apparatus used to measure both S_{21} and transient data. The excitation loop is circular while the receiver loop has a figure-8 shape, thereby preventing these two coils from magnetically coupling,

but allowing coupling to occur with the HTS resonator. The excitation loop is critically coupled to the resonator so that it presents a 50 Ω load to the network analyzer. Critical coupling is obtained by varying the location of the excitation with respect to the HTS resonator. The figure-8 coil is placed several inches from the HTS resonator, so that it is weakly coupled to the resonator. The Hi-Z amplifier is used to further reduce the effect of the figure-8 coil on the losses introduced in the resonator amplifier.

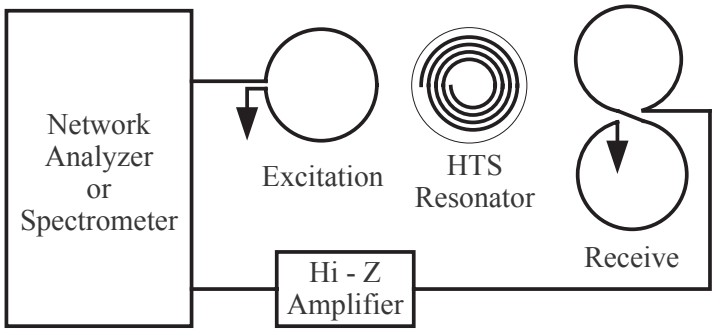


Figure 4: Apparatus used to measure S_{21} and transient data.

The experimental set up is shown in figures 5 and 6. In order to obtain critical coupling, the excitation loop is located inside the cryostat so that it is in close proximity of the resonator. The figure-8 coil is placed outside the cryostat.

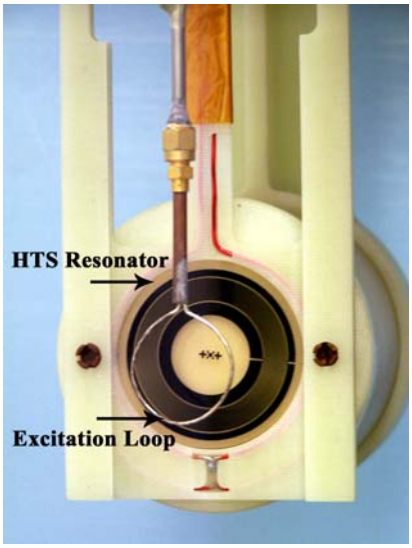


Figure 5: Inside of cryostat.

Figure 6: Outside of cryostat.

Procedure

1. Set a fixed value for the excitation level: ranging from 0 dBm to -60 dBm in -6 dBm steps.
2. At the chosen excitation level (P_e), physically position the circular excitation loop to obtain a 50 Ω input impedance (critically coupled).

Depending on the power level, it is observed that the physical location needed to achieve a 50 Ω match varies.

3. Measure S_{21} using the network analyzer.
4. Making sure not to change the positioning of the coils, measure the pulse response. The pulse width is adjusted so that the HTS response reaches a steady-state value. As the excitation power level decreases, the exponential time constant τ increases so that a larger RF pulse width is needed. In the pulsed measurements, the QR spectrometer generates the excitation pulse and records a voltage that is proportional to the envelope y_e of the resonator current.
5. Repeat procedure for the remaining excitation levels (P_e).

Processing Data

For the S_{21} measurements, the loaded Q-factor is reported by the network analyzer. The unloaded Q-factor, Q_o, is obtained by multiplying this result by two. The transient data is recorded by the spectrometer and is saved to a file to be read at a later time by MATLAB. The MATLAB functions lsqcurvefit() and ode45() are used to obtain a least-squares estimate (LSE) of the model's parameters. Three models are fit to the experimental data. The linear ODE model is

$$\frac{dy_e}{dt} = -c_1 y_e + kF_e \tag{12}$$

The polynomial representation of the resistance R leads to the model,

$$\frac{dy_e}{dt} = -(c_1 + c_2 y_e + c_3 y_e^2) y_e + kF_e, \tag{13}$$

while the exponential representation of the R produces a model

$$\frac{dy_e}{dt} = -(c_1 + c_2 e^{c_3 y_e^{c_4}}) y_e + kF_e \tag{14}$$

In equations (12) through (14), the inductance L is absorbed within the parameters to be estimated. The coefficient c_i and k are chosen to minimize the loss function

$$J = \int (y_e^{est} - y_e^{\exp})^2 dt, \tag{15}$$

where y_e^{est} is the estimated response and $y_e^{\exp}$ is the measured response.

RESULTS AND DISCUSSION

As a first step, data from the pulsed experiments at -60 dBm were used to estimate the unknown parameters. The second column of table I shows the resulting loss functions while figure 7 shows the measured and simulated responses. First, note that the exponential rise-time is smaller than the exponential fall-time for the measured response. Therefore, even at the -60 dBm excitation level, the HTS resonator does not represent a linear system. From table I, the exponential fit is slightly better than the other two models.

As a second step, estimates for the unknown parameters were determined from the pulsed experiments at 0 dBm. The third column of table I presents the loss functions while the measured and simulated responses are displayed in figure 8. Unlike the previous case, the polynomial model had the best fit for the high power level.

Table I: Normalized loss function J

Loss	P_e = -60 dBm	P_e = 0 dBm
Linear [$R(H_{rf}) = R_o$]	5.06 e-7	1.53 e-1
Polynomial [$R(H_{rf}) = R_o + c_1 H_{rf} + c_2 H_{rf}^2$]	5.17 e-7	8.08 e-2
Exponential [$R(H_{rf}) = R_o + c_1 e^{c_2 H_{rf}^{c_3}}$]	3.59 e-7	1.14 e-1

Figure 9 displays the best fit models for the 0 dBm power level using data accumulated from all excitation levels. It is evident that neither the polynomial nor the exponential models provide a good fit for the 0 dBm power level. In general, the higher the excitation power is, the larger the time constants are. For this power level, the polynomial model loss provides the better fit.

CONCLUSIONS

The results show that neither the exponential nor the polynomial models of the resistance R provide a good fit to the experimental data when used with the first-order ODE for the envelope. It was observed that for low power levels, the exponential model had the best fit. On the other hand, the polynomial model displayed the best curve fit for high power levels. A reason for the mismatch between the measured and estimated responses may lie in the derivation of the first-order ODE model. In particular, we replaced the steady-state resonator current y_{ss} with the instantaneous current y_e in the first-order ODE model. Future

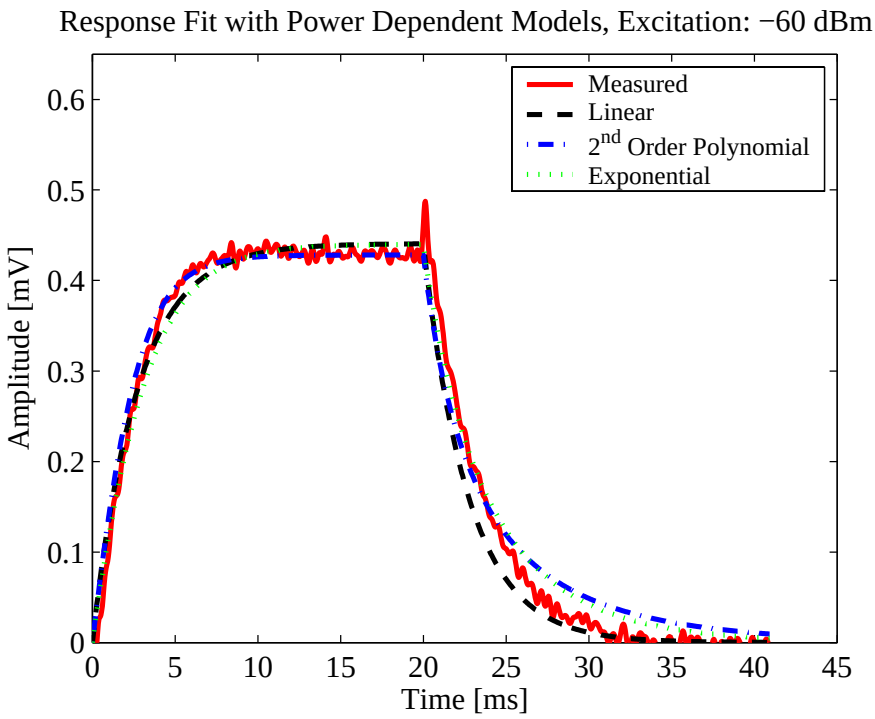


Figure 7: Best fit at an excitation power of -60 dBm.

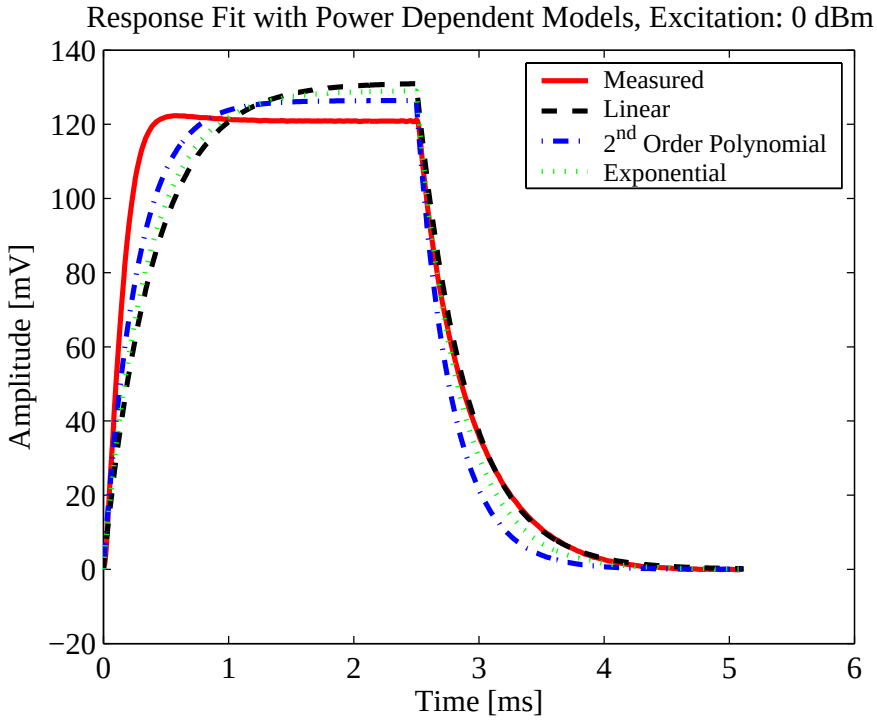


Figure 8: Best fit at an excitation power of 0 dBm.

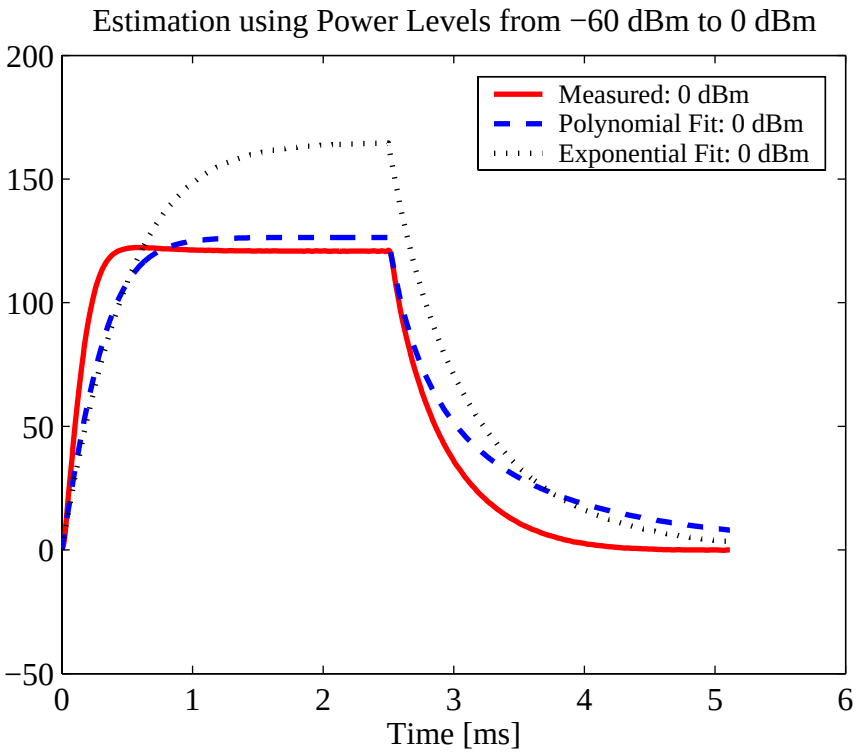


Figure 9: Polynomial and Exponential best fits for 0 dBm.

work would include verifying the validity of this substitution, as well as exploring the possibility that the resonator loss R may directly be a function of F_e in addition to y_e.

ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to Dr. Guo and Dr. Jenkins; without your effort and support this prestigious program would not be possible. I would like to thank my mentor, Dr. Schiano, for allowing me to take part in and learn from his fascinating research. Thank you to the following people for their continuous technical support in making this project possible: Dan Swain, Mike Pusateri, the scientists and engineers at DuPont, and the students in Dr. Schiano's laboratory. I have gained a vast amount of knowledge from this research experience and thank all of you for making it a wonderful learning experience. Thank you to Linda Becker, Amanda Skrabut, and Bo Wang for making all the field trips and industry visits possible and enjoyable. Best wishes to all my fellow EE REU students in all that they choose to pursue; you have made this a most pleasant experience. Last but not least, thank you to my family for supporting me through all of my trials and tribulations.

This material is based upon work supported by the National Science Foundation under Grant No. EEC-0244030 and E. I. du Pont de Nemours and Co.

REFERENCES

[1] Abragam, A., "Macroscopic Aspects of Nuclear Magnetism," pp. 82-83 in *Principles of Nuclear Magnetism*, Clarendon Press, Oxford, 1961.

[2] Shen, Z.-Y., Wilker, C., Pang, P., Holstein, W. L., Face, D. and Kountz, D. J., "High T_c Superconductor-Sapphire Microwave Resonator with Extremely High Q-Values up to 90 K," *IEEE Transactions on Microwave Theory and Techniques*, **40** (12) 2424-2432 (1992).

[3] Nguyen, P. P., Oates, D. E., Dresselhaus, G. and Dresselhaus, M. S., "Nonlinear surface impedance for $YBa_2Cu_3O_{7-x}$ thin films: Measurements and a coupled-grain model," *Physical Review B*, **48** (9) 6400-6412 (1993).

[4] Klein, N., Tellmann, N., Dahne, U., Scholen, A. and Schulz, H., "YBCO Shielded $LaAlO_3$ Dielectric Resonators for Stable Oscillators," *IEEE Transactions on Applied Superconductivity*, **5** (2) 2663-2666 (1995).

[5] Chaloupka, H., Jeck, M., Gurzinski, B. and Kolesov, S., "Superconducting planar disk resonators and filters with high power handling capability," *Electronics Letters*, **32** (18) 1735-1737 (1996).

[6] Hein, M. A., Aminov, B. A., Baumfalk, A., Chaloupka, H. J., Hill, F., Kaiser, T., Kolesov, S., Muller, G. and Piel, H., "High-power high-Q YBaCuO disk resonator filter," *Institute of Physics Conference Series No. 158*, The Netherlands, 1997.

[7] Booth, J. C., Vale, L. R., Ono, R. H. and Claassen, J. H., "Power-Dependent Impedance of High Temperature Superconductor Thin Films: Relation to Harmonic Generation," *Journal of Superconductivity*, **14** (1) 65-72 (2001).

[8] Wilker, C., Shen, Z.-Y., McKenna, S. P., Carter III, C. F. and Brenner, M. S., "Power Handling of High Temperature Superconductors Measured in an HTS/Sapphire Resonator: A Power Reference Film," *Journal of Superconductivity*, **16** (5) 881-887 (2003).

[9] Jackson, J. D., "Wave Guides and Resonant Cavities," pp. 357 in *Classical Electrodynamics*, 2nd ed., John Wiley & Sons, New York, 1975.

[10] Taber, R. C., "A parallel plate resonator technique for microwave loss measurements on superconductors," *Review of Scientific Instruments*, **61** (8) 2200-2206 (1990).

[11] Nilsson, J. W. and Riedel, S. A., "Introduction to Frequency Selective Circuits," pp. 719-722 in *Electric Circuits*, 6th ed., Prentice Hall, New Jersey, 2001.

[12] Witte, R. A., "Network Analyzers," pp. 214-224 in *Spectrum and Network Measurements*, Prentice Hall, New Jersey, 1991.

[13] Montgomery, C. G., ed., "The Measurement of Wavelength," pp. 340-342 in *Technique of Microwave Measurements*, volume 1, Dover Publications, New York, 1966.

[14] Pusateri, M. A., Schiano, J. L, Wilker, C., McCambridge, J. D., Alvarez, R. L. and Laubacher, D. B., "Measuring the Quality-Factor of High-Temperature Superconducting Probes for Quadrupole Resonance Detection Systems." To be submitted to the *Journal of Magnetic Resonance*.

[15] Diete, W., Getta, M., Hein, M., Kaiser, T., Muller, G., Piel, H. and Schlick, H., "Surface Resistance and Nonlinear Dynamic Microwave Losses of Epitaxial HTS Films," *IEEE Transactions on Applied Superconductivity*, **7** (2) 1236-1239 (1997).

[16] Habib, Y. M., Lehner, C. J., Oates, D. E., Vale, L. R., Ono, R. H., Dresselhaus, G. and Dresselhaus, M. S., "Measurements and modeling of the microwave impedance in high-T_c grain-boundary Josephson junctions: Fluxon generation and rf Josephson-vortex dynamics," *Physical Review B*, **57** (21) 13833-13844 (1998).

[17] Kharel, A. P., Velichko, A. V., Powell, J. R., Porch, A., Lancaster, M. J. and Humphreys, R. G., "Unusual features in the nonlinear microwave surface impedance of Y-Ba-Cu-O thin films," *Physical Review B*, **58** (17) 11189-11192 (1998).

[18] Halbritter, J., "Materials science and surface impedance *Z(T,f,H)* of Nb and YBCO and their quantitative modeling by the leakage current of weak links," *Superconductor Science and Technology*, **12**, 883-886 (1999).

[19] Gaganidze, E., Schwab, R., Halbritter, J., Heidinger, R., Aidam, R. and Schneider, R., "Power Handling Capabilities of Y-Ba-Cu-O Wafers and Patterned Microstrip Resonators," *IEEE Transactions on Applied Superconductivity*, **11** (1) 2808-2811 (2001).

[20] Hao, L., Gallop, J. C., Purnell, A. J. and Cohen, L. F., "Intermodulation Measurements on High Temperature Superconducting Thin Films," *Journal of Superconductivity*, **14** (1) 29-33 (2001).

[21] Jacob, M. V., Mazierska, J. and Srivastava, G. P., "Modeling of Nonlinear Surface Impedance of High-T_c Superconductors Using an Exponential Vortex Penetration Model," *Journal of Superconductivity*, **12** (2) 377-383 (1999).

[22] Oates, D. E., Anderson, A. C., Sheen, D. M. and Ali, S. M., "Stripline Resonator Measurements of Z_s Versus H_{rf} in $YBa_2Cu_3O_{7-x}$ Thin Films," *IEEE Transactions on Microwave Theory and Techniques*, **39** (9) 1522-1529 (1991).

[23] Culbertson, J. C., Newman, H. S. and Wilker, C, "Optical probe of microwave current distributions in high temperature superconducting transmission lines," *Journal of Applied Physics*, **84** (5) 2768-2787 (1998).

NSF EE REU PENN STATE Annual Research Journal Vol. II (2004)

STEADY-STATE MODEL FOR THE CALCULATION OF THE ELECTRIC FIELD OF A SPRITE

Tyler Sullivan*, Ningyu Liu[+], and Victor P. Pasko[#]

Department of Electrical Engineering
The Pennsylvania State University
University Park, PA 16802

*Undergraduate student of
Department of Electrical Engineering
Department of Computer Science and Engineering
The Pennsylvania State University
University Park, PA 16802

ABSTRACT

A numerical steady-state model is constructed to compute the electric field of the atmosphere during the occurrence of a sprite given optical emission data of molecular nitrogen (N_2) and ionic nitrogen (N_2^+). The optical emission data will be obtained from the ROCSAT-2 satellite launched by Taiwan's National Space Program Office and currently orbiting the earth recording data about upper atmospheric events [1]. The model uses intensity data from the satellite and applies the data to a relation between intensity and density of certain species along with a relation describing continuity of the density of a certain species with respect to time. The steady-state model is validated using a data set from a previous numerical study of sprites [2].

[#] Faculty Mentor
[+] Graduate Mentor

1. INTRODUCTION

The goal of this paper is to develop a model to study the steady-state electric fields and densities of excited states of N_2 and N_2^+ when sprites occur. Sprites and the ROCSAT-2 satellite are discussed first to provide background information and motivation. Next, the steady-state model is described in detail. Finally, the model will be used with actual data sets from *Liu and Pasko* [2] to test for accuracy and correctness.

1.1 Sprites and Mechanisms of their Production

Sprites are transient luminous events occurring at altitudes of ~40–90 km between the top of a thunderstorm and the ionosphere. Sprites typically occur in clusters measuring ~40 km across with a single sprite having a lateral dimension of ~5–10 km, and only last for several milliseconds [3]. Figure 1 illustrates a red sprite (center right) along with other transient luminous events.

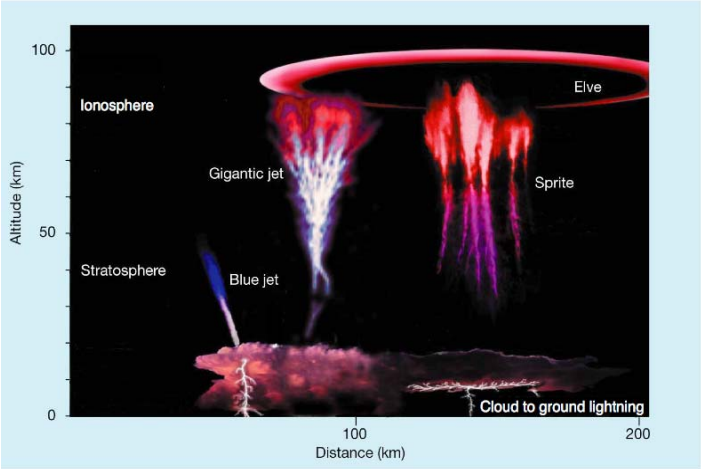


Figure 1: Sprites and other transient luminous events [4, 5].
Reprinted from *Pasko* [5].

Emissions of different bands of N_2 and N_2^+ contribute to the different colors of sprites. The red color of sprites is due to the first positive band of N_2 ($1PN_2$) emissions at the higher altitudes of the sprite's discharge. The other bands of N_2 that contribute to the spectra of sprites to be studied in this paper are the second positive band of N_2 ($2PN_2$) and the first negative band of N_2^+ ($1NN_2^+$). The $1NN_2^+$ and $2PN_2$ bands emit a blue light that dominates the optical emissions at lower altitudes of the sprite [1]. The red bodies of sprites occur from ~50 – 90 km and the blue tendrils of sprites occur near 20 km (the cloud top) while in-between is a color transition phase [3].

Sprites typically follow a positive cloud to ground discharge (+CG) and they follow a negative cloud to ground discharge (-CG) in rare cases [6]. There are several hypothesized mechanisms for the occurrence of sprites. The mechanism

considered for this paper is the quasi-electrostatic (QE) heating model described by *Pasko et al.* [3]. The mechanism starts by having large amounts of charge build up in a cloud. The cloud has negative charge at the bottom and positive charge at the top of the cloud. Also, induced charges between the ionosphere and the cloud's top build opposite to the charges seen in the cloud. The induced charges and cloud charges create a weak electric field between the ionosphere and the cloud. A lightning discharge between the cloud and ground then rapidly occurs removing one of the charges in the cloud (typically the positive charge). A large QE field exists between the cloud and upper atmosphere for a short duration of time based on the relaxation time at each altitude.

The large QE field now provides energy to ambient electrons and generates ionization changes in the atmosphere. The QE field generated after a lightning strike is larger than the breakdown field at high altitudes (~75 km) in the atmosphere resulting in a discharge that leads to the optical emissions comprising sprites. The discharge between the cloud top and the ionosphere causes sprites typically lasting for 1 – 10 ms [3].

Other mechanisms have been proposed to explain why sprites occur. One of these mechanisms includes electromagnetic pulses (EMP). Electromagnetic pulses from lightning strokes may also initiate sprites and are considered an important factor to include in addition to QE fields alone [7]. Another mechanism proposed is the runaway air breakdown mechanism by *Yukhimuk et al.* [8] that considers electric field breakdown not accounted for in the QE model described above. This mechanism results from the ionization effects in the upper atmosphere initiated by energetic electrons generated by cosmic rays passing through and interacting with the upper atmosphere. These ions account for changes in atmospheric conductivity, electron distribution and electric field strength. These mechanisms, among others, may contribute to the initiation of sprites.

1.2 The ROCSAT-2 Satellite

The ROCSAT-2 is the first satellite to perform measurements on upper atmospheric luminous events [1]. The previous techniques used to record sprites included: photography, video cameras and photometers from the ground and airplanes, space shuttles orbiting the earth and, most recently, balloons [6]. The benefit from making sprite measurements from space is that the intensity / emissions data is not severely attenuated by the lower atmosphere as is the case for ground based measurements. The satellite is equipped with several imagers and sensors to record data pertaining to the emissions of lightning-induced phenomena as part of the Imager of Sprites and Upper Atmospheric Lightning (ISUAL) experiment. The ISUAL experiment is one of many uses of the satellite. The sensor packages pertaining to the ISAUL experiment are described below [1].

There are three main packages of sensors for the ISUAL experiment on the satellite with a controlling circuit to help trigger when certain sensors would record data on a CG lightning stroke. The first package of sensors on the satellite

is the sprite imager that records digital images of sprites at a higher temporal resolution necessary due to the short duration of sprites (~5 – 10 ms). The second package of sensors onboard the satellite is two array photometers that record the development of sprites emissions vertically. The last package on the satellite is the spectrophotometer. The spectrophotometer has six channels each tuned to focus on a specific wavelength range that correspond to emissions of a band of N_2 or N_2^+. One channel is designated to detect ultraviolet light which may provide a signature of sprites. Two channels are set to detect emission wavelengths of 337.0 nm and 391.4 nm corresponding to the presence of the second positive band of N_2 and the first negative band of N_2^+. Another channel covers a range of wavelengths which should encompass most of the emissions from the first positive band of N_2. The other channels in the spectrophotometer are not used in the model. Each channel has a photocathode and window that affect the number of particles that pass through the window to the filter for detection [1].

2. DESCRIPTION OF THE STEADY-STATE MODEL

The steady-state model implemented in MATLAB takes intensity data at an altitude as input parameters. Interpolation models were designed to handle outputs for each channel of the photometer. The first step in the model is to correct for the filter effects of the spectrophotometer. The filter functions for the transmission curves of each channel in the spectrophotometer are from *Harris* [9]. Quantum efficiency curves for the different photocathode and window materials used are from *Heetderks* [10]. Corrections due to the quantum efficiency of the windows and transmission rates of the channel are computed and applied to the intensity data. Next, partial intensities are used from *Vallance Jones* [11] to compute the amount of light the photometers detect of a certain species based on the wavelengths of the filter. Tables containing emission data from *Vallance Jones* [11] of different species are used to compute the total intensity a channel of the photometer could detect based on the wavelengths the filter could detect. The corrected intensities are used to compute the density of the species using the equation [2]:

$$I_k = 10^{-6} \int_{L'} A_k n_k d\ell \qquad (1)$$

An assumption is made so the integral above is taken at a point so that (1) becomes [2]:

$$n_k = 10^6 \frac{I_k}{A_k} \qquad (2)$$

where the intensity, I_k is given in Rayleighs (R) and the radiation transition rate for species k, A_k is needed to compute the density of excited particles for species

k, n_k [cm^{-3}]. The A_k transition rates for the different species are: 1.7 x 10^5 s^{-1} for $1PN_2$, 2 x 10^7 s^{-1} for $2PN_2$ and 1.4 x 10^7 s^{-1} for $1NN_2^+$. Since only multiplication operations have been used, the order of applying correcting factors for the filter and quantum efficiency does not matter.

The model then uses the steady-state solution of the continuity equation at a specific altitude [2]:

$$\frac{\partial n_k}{\partial t} = -\frac{n_k}{\tau_k} + \sum_m n_m A_m + \nu_k n_e \qquad (3)$$

where τ_k is the lifetime of the state k, the summation of $n_m A_m$ accounts for cascading effects while the $\nu_k n_e$ term is the source term that accounts for the presence of the electric field [2]. The only species considered that has cascading effects is $1PN_2$. Cascading accounts for the population of excited states that result in energy transitions, which may involve radiation, from higher excited states [11]. The term ν_k is the excitation coefficient for different species computed as a function of the altitude and electric field at the altitude. Figure 2 contains a plot of the excitation coefficients used by *Liu and Pasko* [2]. The term n_e is the electron density in the air. This term is the same for any species. The lifetime of a species, τ_k, is computed at a given altitude using the formula [2]:

$$\tau_k = \left[A_k + \alpha_1 \cdot [N_2] + \alpha_2 \cdot [O_2]\right]^{-1} \qquad (4)$$

where A_k is the radiation transition rates for a species k and the values of A_k used are mentioned above. $[N_2]$ and $[O_2]$ are the densities of N_2 and O_2 present in the atmosphere. α_1 and α_2 are the quenching rates of the species k. Quenching accounts for the excitation energy lost in collisions with other molecules. At higher altitudes this effect only matters in excited particles when the species studied collide with N_2 and O_2 [11]. The quenching coefficients used for the $1PN_2$ species are α_1 = 1 x 10^{-11} cm^3s^{-1} and for the $2PN_2$ species α_2 = 3 x 10^{-10} cm^3s^{-1} and for the $1NN_2^+$ species α_1 = 4 x 10^{-10} cm^3s^{-1}. Quenching coefficients not listed are zero [11].

The steady-state solution of the continuity equation (3) is in general:

$$n_k = \tau_k \left(\sum_m n_m A_m + \nu_k n_e \right) \qquad (5)$$

where the only species to include the cascading effects, the term, is the $1PN_2$ species where the density of the $2PN_2$ state and A_k term for the $2PN_2$ state are needed. Equation (5) derives a ratio between ν_k terms for the $2PN_2$ and $1NN_2^+$

species. The ratio of ν_k terms is then used to compute the electric field. Figure 3 shows the ratio of ν_k terms used to determine the electric field.

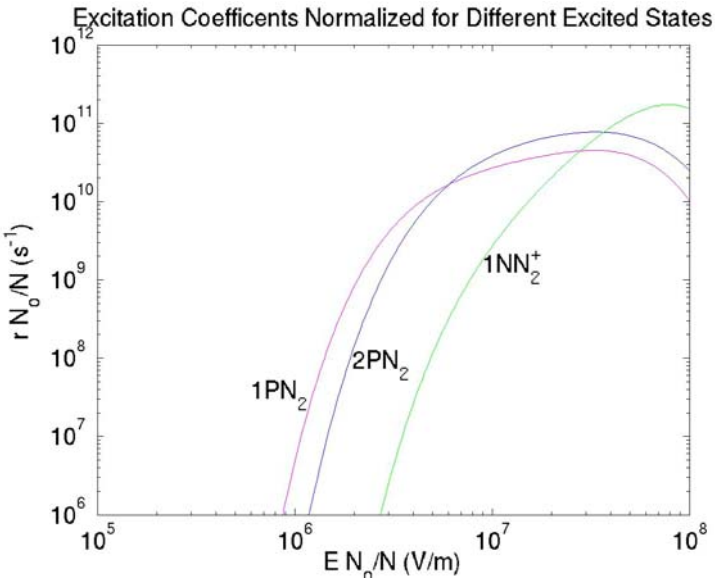


Figure 2 (above): Plot of the excitation coefficients used for the $1PN_2$, $2PN_2$ and $1NN_2^+$ species from *Liu and Pasko* [2]. The electric field and excitation coefficients used were normalized or do not account for the density of air at a given altitude. N_o is the density of air at ground (0 km) while N is the density of air at a given altitude.

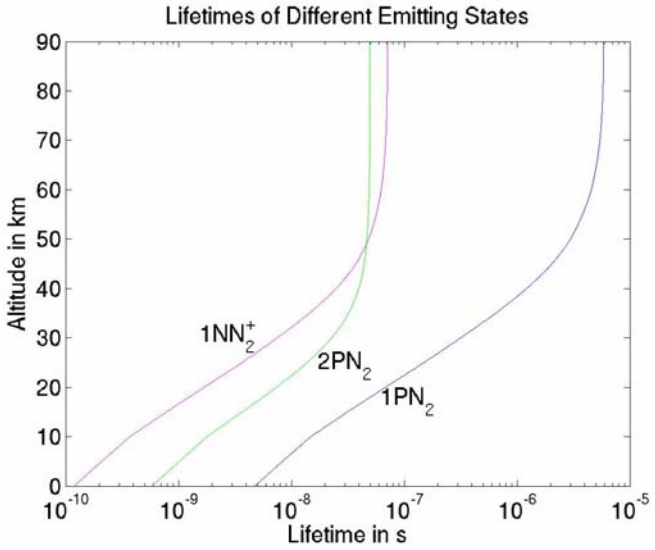


Figure 3: Plot of the lifetimes of each species for different altitudes using equation (4).

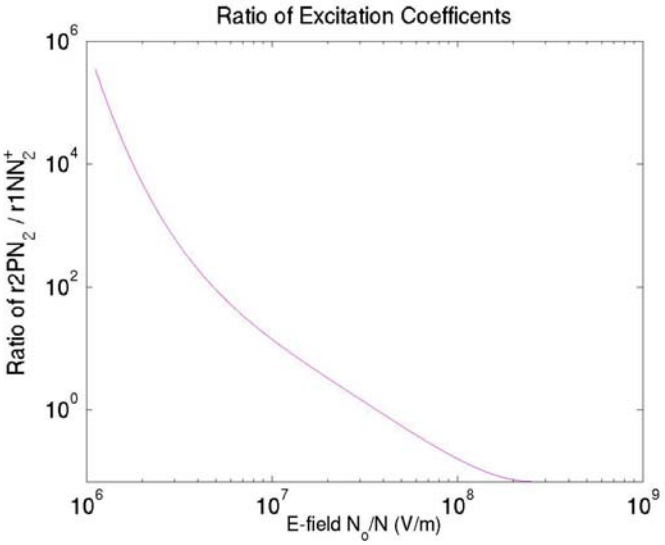


Figure 4: Plot of the ratio of excitation coefficients. The ratio of excitation coefficients used for this graph was $2PN_2$ / $1NN_2^+$. The electric field used for input was normalized or do not account for the density of air at a given altitude. N_o is the density of air at ground (0 km) while N is the density of air at a given altitude.

3. TEST RESULTS

Data from *Liu and Pasko* [2] was used to test the steady-state model by supplying data about the electric field and density of species at an altitude of 70 km. The steady-state model was used to compute the electric field given the density information. The electric field from *Liu and Pasko* [2] was compared against the one found using the steady-state solution developed. The data used for this test covered a region in space that was 200 m in height starting at 70 km and did not extend in the radial direction horizontally. The electric fields used in both the steady-state model and the provided data set did include the atmospheric effects. The input for the steady-state model is given in Figure 5.

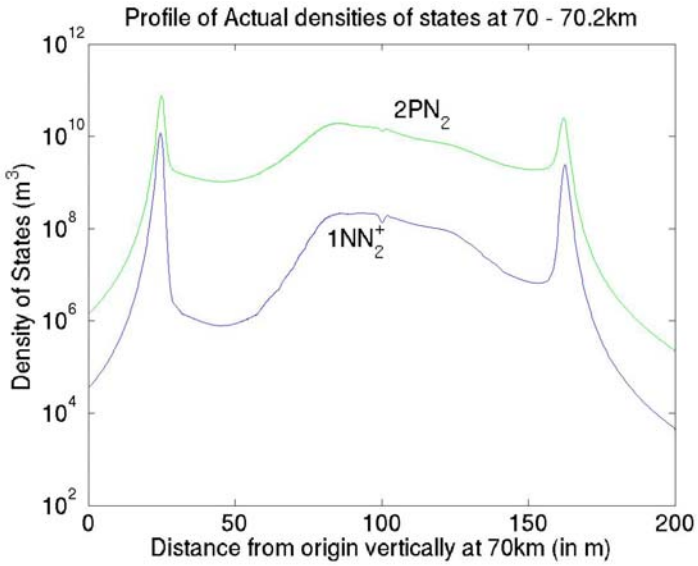


Figure 5: Plots of the input density curves used by the steady-state model from *Liu and Pasko* [2].

Intensity curves were not used as input since the steady-state model was implemented using the assumption that the intensity was determined at a point to avoid integration. The electric field results were plotted together to draw comparisons between the different models used. A plot of the percent error between the curves was constructed as well for the analysis. The percent error was determined by comparing the steady-state model electric field against numerical results from *Liu and Pasko* [2].

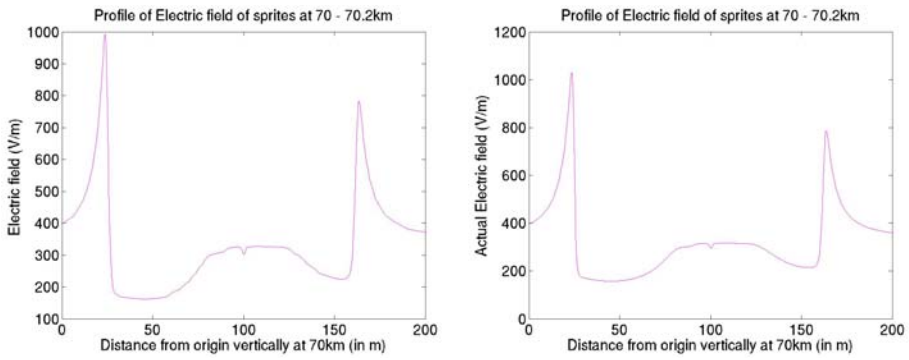


Figure 6: (Left) The computed electric field (in V/m) using the steady-state model. (Right) The electric fields (in V/m) used the data provided from *Liu and Pasko,* [2]. Both plots above show the electric fields at an altitude of 70 km.

The electric field from the steady-state model (left) possessed all of the same features as the data provided by *Liu and Pasko* [2]. The model was able to exhibit the peaks and valleys of the expected result accurately. The differences between the curves were noticeable when the percent error was computed. The percent error curve and a graph with the two fields above super-imposed on each other are shown in Figure 7:

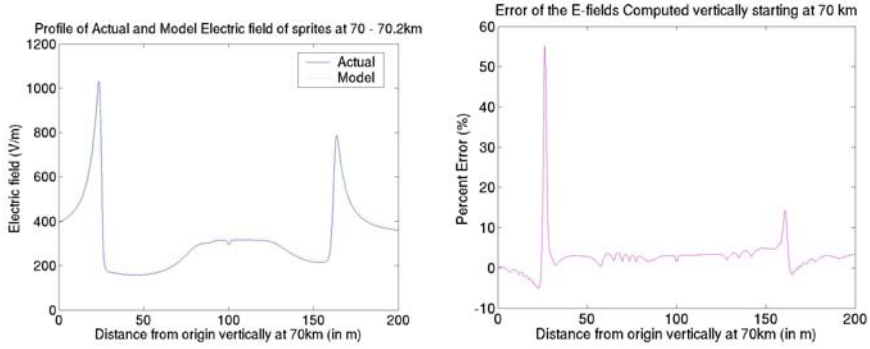


Figure 7: (Left) Plots of both E-fields (in V/m) at 70 km using the plots in Figure 6 above. (Right) The percent error between the electric field plots.

The percent error for the model was greatest when the electric field was beginning to rapidly increase or finishing a quick decrease. The percent error curve also showed that the significant errors were not at the peaks of the curves or when the electric field was changing rapidly. The plots below demonstrate the gap between the model and actual results.

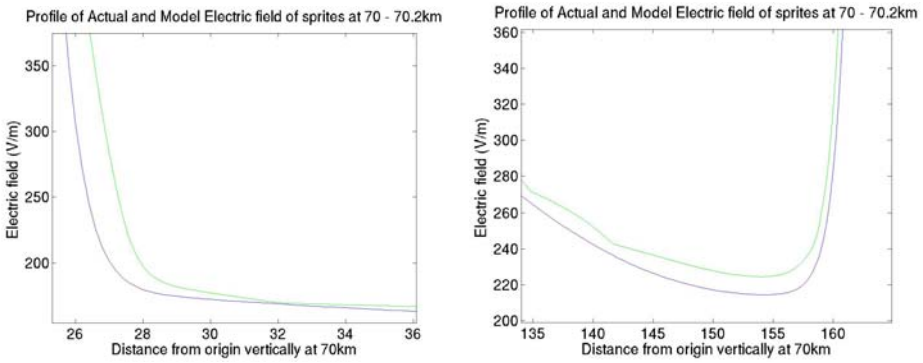


Figure 8: (Left): The plot of the difference of the actual and model electric field emphasizing the difference at about 26 m. The percent error at this point was 55%. (Right) The plots of the actual and model electric field curves are shown to emphasize the gap between the curves at 161 m. The percent error at this point is about 14%.

The gaps in Figure 8 occurred after and before sharp change in the electric field with respect to the altitude. The reason for these gaps was that the steady-state model assumed that change in particle density was zero resulting in information being lost. The electric fields used for comparison from *Liu and Pasko,* [2] were computed knowing past electron densities, densities of excited states and electric fields when computing the electric field at a given time.

4. CONCLUSION

The steady-state model developed computes quickly an approximate electric field which can be compared to electric fields computed with more sophisticated methods. The model was able to simplify computations by making assumptions ignoring the geometry of the space considered and assuming the density of excited states was constant with respect to time (steady-state assumption). The elimination of both spatial and temporal variations results in a problem that is simple to compute. The avoidance of integrals and differential equations provide a quick steady-state model for fast numerical approximations. The steady-state model demonstrates that a reasonable answer within 5% can be computed without knowing the past electric field or knowing about past excitation state densities.

Acknowledgements

Special thanks to Dr. Victor Pasko and Ningyu Liu for their assistance in completing this paper. This paper is based upon work supported by the National Science Foundation under Grant No. EEC-0244030 and by NASA Grant No. NAG5-11832 to The Pennsylvania State University.

References

[1] Chern, J. L., R. R. Hsu, H. T. Su, S. B. Mende, H. Fukunishi, Y. Takahashi and L. C. Lee, "Global survey of upper atmospheric transient luminous events on the ROCSAT-2 satellite," *Journal of Atmospheric and Solar-Terrestrial Physics*, 65, 647 – 659, 2003.

[2] Liu, Ningyu and Victor P. Pasko, "Effects of photoionization on propagation and branching of positive and negative streamers in sprites," *Journal of Geophysical Research*, 109, doi:10.1029/2003JA010064, 2004.

[3] Pasko, V. P., U. S. Inan, T. F. Bell and Y. N. Taranenko, "Sprites Produced by quasi-electrostatic heating and ionization," *Journal of Geophysical Research*, 102 (A3), 4529 – 4561, 1997.

[4] Lyons, Walter A., Thomas E. Nelson, Russell A. Armstrong, Victor P. Pasko and Mark A. Stanley, "Upward Electrical Discharges from Thunderstorm Tops," *American Meteorological Society*, 445 – 454, April 2003.

[5] Pasko, Victor P., "Electric Jets," *Nature*, 423, 927 – 930, 26 June 2003.

[6] Bering, E. A., J. R. Benbrook, L. Bhusal, J. A. Garrett, A. M. Paredes, E. M. Wescott, D. R. Moudry, D. D. Sentman, H. C. Stenback-Nielsen and W. A. Lyons, "Observation of transient luminous events (TLEs) associated with negative cloud to ground (-CG) lightning strokes," *Geophysical Research Letters*, 31, doi:10.1029/2003GL018659, 2004.

[7] Adachi, T., H. Fukunishi, Y. Takahashi and M. Sato, "Roles of the EMP and QE fields in the generation of columniform sprites," *Geophysical Research Letters*, 31, doi:10.1029/2003GL019081, 2004.

[8] Yukhimuk, V., R. A. Roussel-Dupré, E. M. D. Symbalisty and Y. Taranenko, "Optical characteristics of Red Sprites produced by runaway air breakdown," *Journal of Geophysical Research*, 103 (D10), 11473 – 11482, 1998.

[9] Harris, Stewart, "ISUAL Spectrophotometer Science Performance Test Report," Doc. 8998-W7 rev. B. 11 April 2003,

[10] Heetderks, H., "ISUAL Spectrophotometer," 7 June 2000, http://sprg.ssl.berkeley.edu/sprite/ago96/isual/SDR_Package/

[11] Vallance Jones, A., Optical Emissions from Aurora, p. 80 – 176, *Aurora*, First Edition, B. M. McCormac Editior, D. Reidel Publishing Company; Dordrecht, Holland; 1974.

PLASMA-ASSISTED OXIDATION FOR SURFACE PASSIVATION OF SILICON NANOWIRES

Daniel J. Black[1,*], Jim Mattzela[2,+], Tsung-ta Ho[3,+], Yanfeng Wang[3,+], K-K Lew[4,+] Joan Redwing[4,#], and Theresa S. Mayer[3,#]

[2]Intercollege Program in Material Science and Engineering
[3]Department of Electrical Engineering
[4]Department of Material Science and Engineering
The Pennsylvania State University
University Park, PA 16802

*Undergraduate Student of
[1]Department of Electrical and Computer Engineering
University of Utah, Salt Lake City, UT 84112

ABSTRACT

A thin layer of silicon dioxide (SiO_2) (10-20nm) was grown on planar surfaces of silicon in the (100), (111), and (110) directions to determine the role of orientation in oxidation of silicon nanowires. Both wet oxidation and plasma-assisted oxidation were performed. Ellipsometry was used to measure oxide thickness and capacitance-voltage measurements were used to characterize the dielectric layer. Based on the orientation dependent SiO2 growth rate, plasma-assisted oxidation is shown to closely resemble a high temperature thermal oxidation. These results suggest that the plasma oxidation method can be used to controllably grow high-quality oxide that will likely reduce interface states by passivating the nanowire surface.

[#] Faculty Mentor
[+] Graduate Mentor

INTRODUCTION

Nanowire field-effect transistors (NWFETs) are promising for a number of reasons. As indicated by Yang, et al[1], NWFETs and FinFETs can be developed with very high electron and hole mobility as well as extremely low leakage currents. In addition, the geometry of these devices allows a gate that covers much more surface area, improving the overall speed and control of the device.

However, this large surface area in comparison to the total volume of small-diameter silicon nanowires introduces more interface states, which can overwhelm the electrical response of NWFETs. Back-gated FET structures such as that presented by Cui, et al[2] and shown in Figure I typically have a very hysteretic nature with no surface passivation. As shown in Figure II, the voltage threshold changes dramatically depending on the direction of the voltage sweep. This is most likely explained by the high number of interface states present which work to trap charge between the silicon surface and the low-quality native oxide.

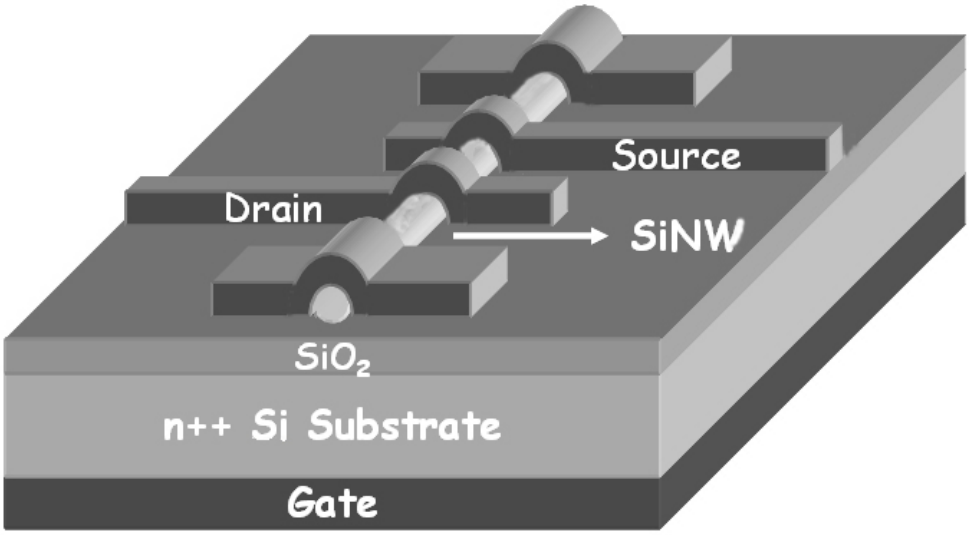


Figure I. Back-gated NWFET test structure

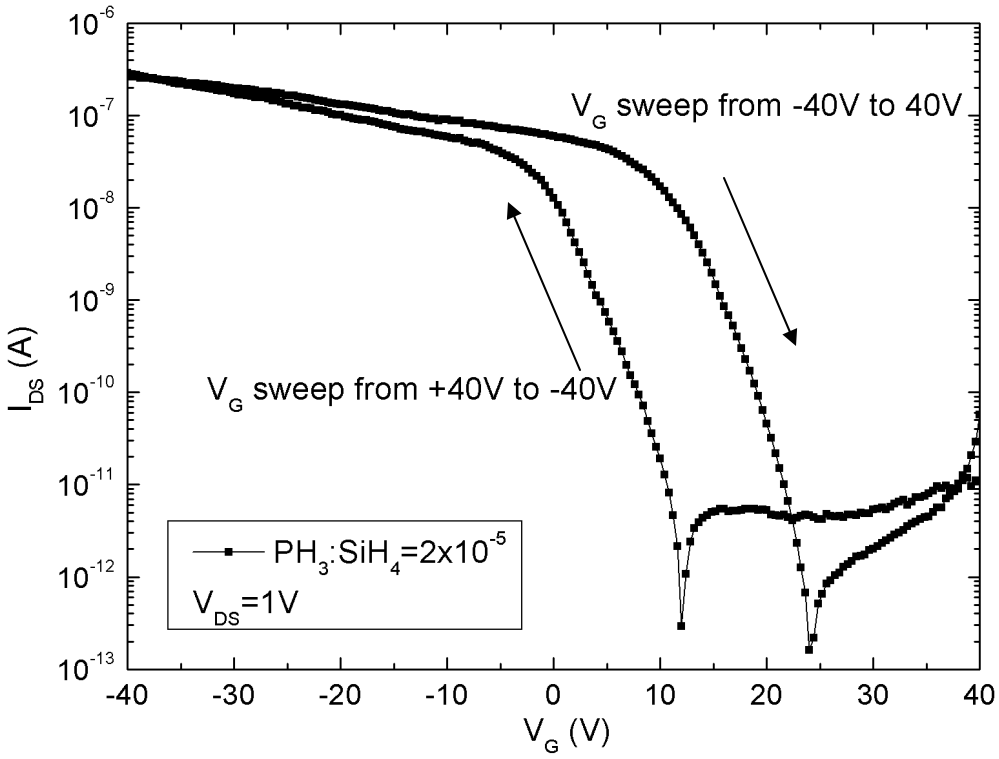


Figure II. Hysteresis of silicon NWFET devices

It is necessary, therefore, to develop a method for passivating the surface of the nanowires, minimizing surface states. The most obvious and practical method of passivating the surface of the nanowires is to grow a high quality thermal oxide as is commonly done in standard CMOS processing. This will also be beneficial in the development of a top-gated NWFET structure.

However, depending on the orientation of nanowire growth, it is possible that several different surface orientations may exist around the circumference of the nanowire. It has been shown that oxidation proceeds at different rates depending on surface orientation[3], which could lead to a non-uniform oxide layer. This effect is exaggerated for oxidations performed at lower temperatures[4].

In this paper, wet (H_2O vapor) and plasma-assisted oxidation experiments are reported. These experiments were performed on planar silicon samples of different orientations. It was found that plasma-assisted oxidation grows similar thicknesses of oxide regardless of crystal orientation, more closely resembling a higher temperature thermal oxidation. The physical characteristics of the results were measured by ellipsometry, while the electrical properties were measured by capacitance-voltage (C-V) measurements.

EXPERIMENTAL DESCRIPTION

Planar samples of (100), (111), and (110) silicon were prepared by standard CMOS wafer cleaning techniques (RCA1, RCA2, HF dip). The (100) samples were n-type and the (111) and (110) were p-type, each with a resistivity in the range of 1-10 Ωcm. One sample of each orientation was placed in the wet oxidation furnace for 30 minutes at 750 ºC.

Plasma-assisted oxidations were performed using a Trion Orion Inductively Coupled Plasma (ICP) Enhanced Chemical Vapor Deposition system. Plasma-assisted oxides were grown on samples of each orientation for 5, 7.5, 10, and 15 minutes. All plasma-assisted oxidations were performed at 300 mTorr and 200 ºC with an oxygen flow of 75 sccm, an ICP power of 300 Watts and a low frequency chuck power of 50 Watts. One additional sample was oxidized for 10 minutes and then annealed *in-situ* for 45 minutes in argon (no plasma) at 300 ºC in the plasma chamber. Oxide thickness and the effective index of refraction were determined using ellipsometry.

Photolithography was then performed on the samples to define contact vias to the silicon surface. Aluminum was deposited by thermal evaporation. Capacitors of 100 μm x 100 μm were then patterned and etched. Capacitance-voltage measurements were made at 10 kHz and 1 MHz on several capacitors on each sample.

RESULTS AND DISCUSSION

Table I shows the ellipsometry results of the wet oxidation samples. Note that the oxide thickness is highly dependant on crystal orientation, as expected[3].

Table I. Wet oxidation results

Crystal orientation	Thickness (Å)
(100)	68.72
(111)	92.13
(110)	146.3

Conversely, the ellipsometer showed almost identical oxide thicknesses for the different orientations in plasma-assisted oxidations. These results are summarized in Figure III. This seems to indicate that plasma-assisted oxidation more closely emulates high temperature thermal oxidation, in which orientation has a smaller effect on the oxide thickness[4].

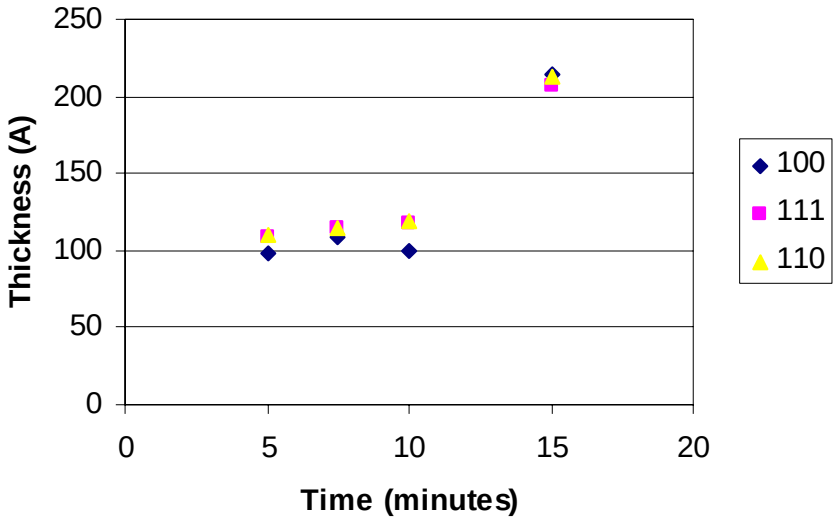


Figure III. Oxide Thickness, Plasma-Assisted Oxidation

Interestingly, the plasma-assisted oxidation seems to yield similar thickness for all samples from 5-10 minutes, at ~100 Å. The oxide thickness doubles for the 15 minute plasma-assisted samples. In future studies, intermediate data points should be collected to determine an exact growth mechanism. However, a preliminary understanding of this phenomenon can be gleaned from the refractive index of each sample, which was determined from ellipsometry measurements. These data are shown in Figure IV.

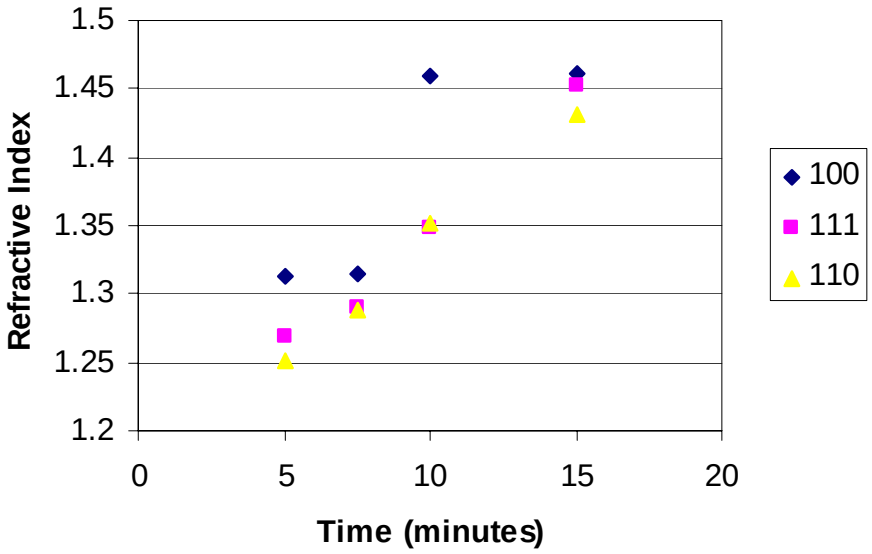


Figure IV. Refractive Index of Oxide Plasma Assisted Oxidation

It is apparent from this figure that as time increases, the refractive index also increases until it reaches the expected refractive index of silicon dioxide, 1.459. This indicates that as oxidation begins, the 5 minute 100 Å oxide is oxygen rich compared to silicon dioxide. As time progresses the density of the oxide increases until it is closer to stoichiometric silicon dioxide. At that point, additional oxidation increases the thickness of the oxide without reducing the density of the oxide. This effect appears to occur first for the (100) orientation, which has fewer surface atoms per unit area to react with the oxygen compared to the other crystal orientations.

C-V measurements were also performed on each sample and the flatband voltage given by Equation (1) was determined using the graphical technique outlined by Schroder[4].

$$V_{FB} = \phi_{MS} - \frac{Q_f - \lambda Q_m - \lambda Q_{ot} - Q_{it}(\phi_S)}{C_{ox}} \qquad (1)$$

The above equation contains several terms for different types of charge. However, if we consider all these together, we can define the total contribution to flatband voltage due to charge Q/C_{ox} as the difference between V_{FB} and Φ_{MS}, the metal-semiconductor work function.

$$\frac{Q}{C_{ox}} = \phi_{MS} - V_{FB} \qquad (2)$$

Φ_{MS} was calculated based on dopant concentrations of $4x10^{15}$ for the n-type (100) samples and 10^{16} for the p-type (111 and 110) samples. Figure V shows the difference between the measured flatband voltage and the calculated metal-semiconductor work function.

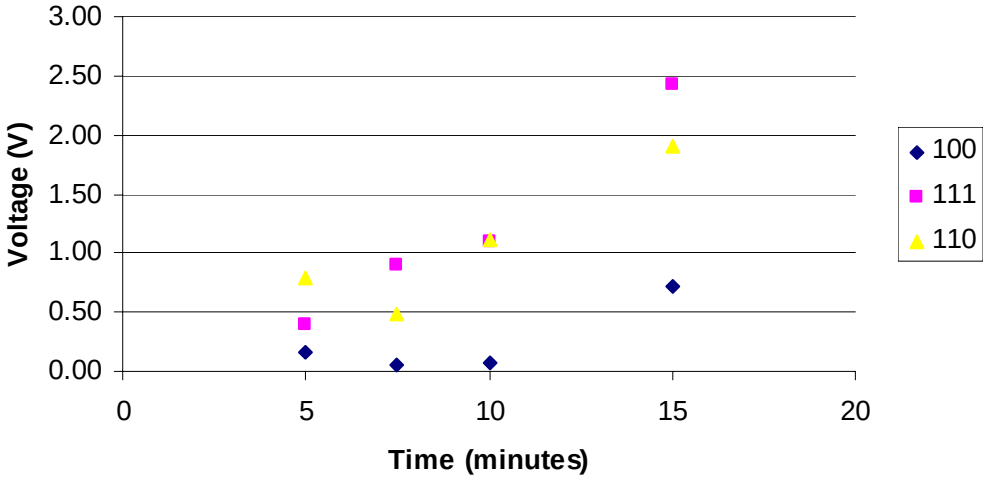


Figure V. Oxide Charge Contribution to Flatband Voltage

There is a general trend toward greater charge with increasing time. As the oxide continues to grow, more charge traps are introduced into the oxide, increasing the magnitude of the flatband voltage.

It should be noted that because the (100) orientation is n-type, it has a smaller numeric discrepancy between the flatband voltage and the metal-semiconductor work function (see Figure IV). However, Φ_{MS} is about 0.05 V for the n-type, compared to about 0.7 V for p-type, so the results are quite comparable based on a percentage comparison.

Table II compares the charge of the 10 minute samples to the annealed samples. Note that in the (111) and (110) cases, annealing greatly reduces the trapped charge. This indicates a plasma-assisted oxidation followed by annealing may be a good solution for minimizing interface states and improve the performance of future NWFETs.

Table II. Effects of annealing on trapped charge

Crystal orientation	Without anneal (V)	With anneal (V)
(100)	0.07	0.29
(111)	1.09	0.43
(110)	1.11	0.46

CONCLUSION

It has been shown that plasma-assisted oxidation resembles a high temperature thermal oxidation. This oxidation results in a consistent oxide thickness for all three crystal orientations studied. It was shown that a longer oxidation time increases the density and refractive index of the oxide, but also introduces increased trapped charge. This charge can be reduced by annealing to form a high quality thin oxide that should be suitable for nanowire devices.

In future work, the effects of annealing on different crystal orientations should be further studied, as well as oxidation and annealing of the actual nanowires.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. EEC-0244030.

REFERENCES

[1] F. L. Yang, D. H. Lee, H. Y. Chen, C. Y. Chang, "5nm-Gate Nanowire FinFET," *IEEE 2004 Symposium on VLSI Technology Digest of Technical Papers*, p.196, 2004.

[2] Y. Cui, Z. Zhong, D. Wang, W. U. Wang, C. M. Lieber, "High Performance Silicon Nanowire Field Effect Transistors," *Nano Letters,* vol. 3, no. 2, p. 149, 2003.

[3] D. B. Kao, J. P. McVittie, W. D. Nix, and K. C. Saraswat, "Two-Dimensional Thermal Oxidation of Silicon—1. Experiments," *IEEE Transactions on Electron Devices,* vol. ED-34, p. 1008, 1987.

[4] E. H. Nicollian and J. R. Brew, *MOS Physics and Technology*, New York: Wiley, p. 681-689, 1982.

[5] D. K. Schroder, *Semiconductor Material and Device Characterization,* 2nd ed., New York: Wiley, p. 350, 1998.

EFFECTS OF CARBON NANOTUBES ON MEMS STRUCTURES

Rakesh Reddy* and Srinivas A. Tadigadapa[#]

Department of Electrical Engineering
The Pennsylvania State University
University Park, PA 16802

*Undergraduate student of
Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh, PA 15213

ABSTRACT

The focus of this study is to determine the dissipation mechanisms and their effect on the resonance characteristics of micro mechanical structures in air, vacuum, and when coated nano-materials. Understanding of the viscoelastic dissipation behavior in various ambients and with different coatings is important to the functioning of the resonant micromechanical structures for various sensing applications.

The work will begin with the automation of data acquisition using LabVIEW® software. Labview® is a graphical development environment that allows for the control of and data acquisition from various instruments. Using LabVIEW®, measurement instruments will be automated to acquire real time data of the resonance of the micro-mechanical structures in order to observe the drift in the frequency and quality factor in various ambients. Agilent 4294A precision impedance analyzer will be automated by LabVIEW and be used to measure the characteristics of the structures. Automation will include real-time tracking of the resonance frequency as well as fitting the data to Lorentzian curve using least square technique for the numerical evaluation of the resonance frequency and Q-factor.

Micromachined structures to be measured under various environments include micro-mechanical resonating structures such as bridges and cantilevers which have been fabricated by graduate students in the group.

[#] Faculty Mentor

INTRODUCTION

The field of micro-electro mechanical systems (MEMS) shows great promise. As a relatively young field, there are still many properties to be studied. The aim of this experiment is to observe how nanomaterials affect the characteristics of micro-electro mechanical devices. Specifically, we will be looking at micro-machined cantilevers and bridges and how the deposition of carbon nanotubes and the changes on the resonance frequency and Q-factor. The resonance characteristics of bridges and cantilevers have many applications including microscopy, chemical detection and bio-sensing.

The devices used were made in EE597 Pennsylvania State University under Prof. Tadigadapa. The fabrication process consisted of a boron etch stop and anodic bonding. A p-type Silicon wafer was doped with a high concentration of Boron. Contact photolithography process was used to first define the gap between the structures and common electrode and then to define actual structures. Reactive ion etching (RIE) was used to create a ~1μm etch for the gaps and a 6-7μm etch for the structures. Using a lift-off process, Cr and Au layer were deposited and patterned on a borosilicate glass wafer. The metal deposit layer defines the electrodes for the devices. After the devices were placed in a Piranha clean solution, they went through an anodic bonding process to fuse the two wafers together. The devices were diced and went through an Ethylene Diamine Pyrocatechol (EDP) etch process. The EDP etch process does not etch Silicon that is heavily doped with Boron. As a result, all of the undoped Silicon is removed, leaving the freestanding structures. Finally, devices were dried using critical point drying in order to reduce the chances of stiction[1]. An image of a die after packaging is shown in Figure 1 and the dimensions of the devices are shown in the Table 1.

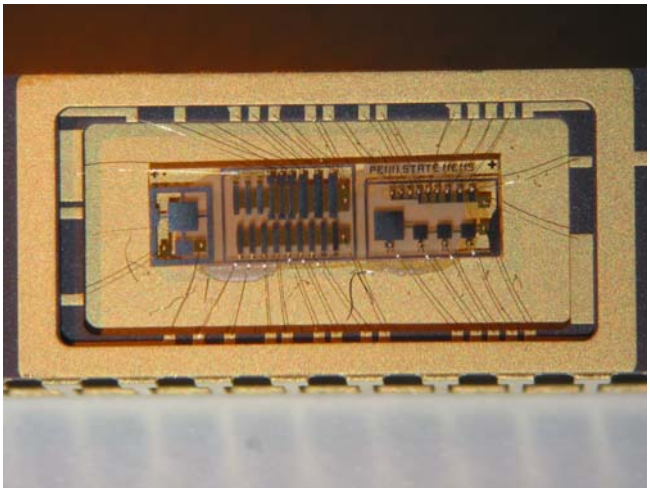


Figure 1. Picture of packaged device.

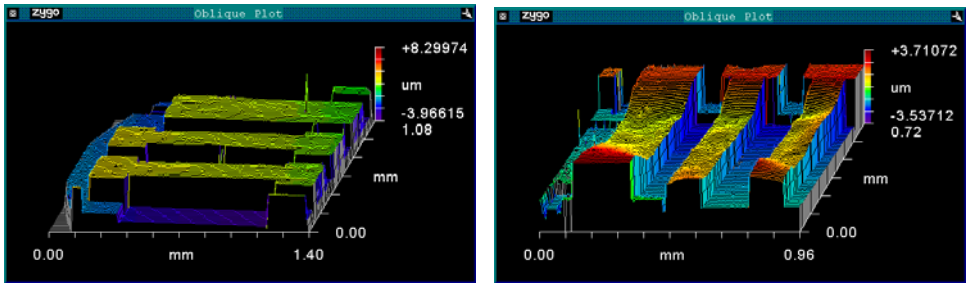
Table 1. Nominal Dimensions of Bridges and Cantilevers

Structure	Type	Thickness	Gap	Width(mm)	Length(mm)
Cantilever	A	3 mm	1 mm	100	150
	B			100	300
	C			200	300
Bridge	A			200	600
	B			120	1000
	C			300	1000

Once a device is chosen, the resonance characteristics will be measured. Carbon nanotubes suspended in N-Methyl Pyrroldone were deposited and the resonance characteristics were measured again. The resonance characteristics of these devices will be measured over extended periods of time in order to observe subtle changes that may occur as a result of the carbon nanotubes. My primary job was to automate the data acquisition process with the LabVIEW programming language to allow for the measurement of the devices for these long periods. We believe that changes in the resonance characteristics and Q-factor will be observed after the carbon nanotubes are deposited.

EXPERIMENTAL DESCRIPTION

The first step this experiment was to determine which devices to use. To do this surface profiles of the structures were taken using a Zygo NewView 100 white-light interference microscope. Several things were looked at in determining whether to further test a device on a die. Some devices shared the same electrode resulting in unwanted interference in resonance measurements. Devices were analyzed in order to find cantilevers with less than 0.1μm bend and bridges that are flat to 0.5 μm. Cantilevers often suffered from being bent up or down while bridges suffered from buckling. Cross sections of the Zygo measurements could be taken using MetroPro software in order to analyze the quality of the device. A surface profile of a set of flat bridges is shown in Figure 2a. Bent cantilevers and buckled bridges were undesirable, as the resonance frequencies of these devices are much poorer than flat devices. Stiction is a state caused by Van Der Waal forces, causing devices to attach to the common electrode during the drying of the device. The result of stiction in these devices is an electrical short and poor to no observable resonance characteristics. Devices suffering from stiction were discarded from further testing. The surface profile of a bridge suffering from stiction is shown in Figure 2b.



a) (b)

Figure 2: a) Shows the surface profile of flat bridges, b) Shows the surface profile suffering from stiction.

After the die's were analyzed, several of the devices that fit the requirements mentioned above were prepared for further testing. The devices were mounted on a Dual In-line Package (DIP) using super glue. A wire bonder was used to attach electrodes on the die to pins of the DIP. The connections from the die to the DIP were mapped out under a microscope.

Packaged devices were tested using the Agilent 4294A High Precision Impedance Analyzer to look at the resonance characteristics of these devices. Devices were placed in a vacuum chamber for measurements because resonance characteristics were indiscernible at around 1 Atmosphere. The device was placed in a chamber and connected to a vacuum pump with a pressure of 50 μTorr. The chamber containing the device was an Aluminum cylinder. This construction acts as a Faraday cage in minimizing disturbance from electrical interference. Precaution is still needed, as small changes in the environment cause noticeable changes in measurements. The series capacitance (Cs) and series resistance (Rs) were measured for devices to find the resonance characteristics. This model was chosen since the separation between the common electrode and device acts like a capacitor. The Rs curve was analyzed to find the resonance peaks and determine the Q-factors of the devices.

After the devices were analyzed with the impedance analyzer, we chose to the bridge with strongest resonance. The dimensions for this device were 200μm wide and 600μm. The resonance characteristics for the device were measured overnight. Once the measurement was completed, MicroFab Technologie's MJ-AB jetting device was used to deposit carbon nanotubes suspended in a solution. The jetting device allows for drop sizes as small as 20μm. The device operates using actuation by PZT actuators bonded to a glass capillary. As voltage is applied to the PZT, the cross-sectional area is increased or decreased, changing the pressure on the liquid[2]. This effect results in drop formation of the fluid in the device. The micro dropper was controlled using JetDrive III software and a mirco-dropper controller also from MicroFab Technologies. After helping setup the equipment for this part of the experiment, a graduate student deposited the carbon nanotubes on the chosen device.

The changes in the resonance characteristics by applying carbon nanotubes to the MEMS structures, may be very subtle. As a result, structures were measured for long periods of time with the impedance analyzer. The impedance analyzer does not have the ability to perform data acquisition and storage for long periods of time. It would be a mundane take to have someone manually transfer data with a diskette every 5 minutes for periods of up to 12 hours. The solution to this problem was to automate the impedance analyzer with a computer. The impedance analyzer is controllable through a General Purpose Information Bus (GPIB). GPIB is a protocol, which many industrial devices support allowing for the control of multiple devices using one controller. This allows for a PC equipped with a GPIB interface to send commands to the impedance analyzer, and collect and manipulate data for extended periods of time without a user present.

The automation of the device over a GPIB was done using the LabVIEW graphical programming language. LabVIEW is a platform independent, graphical programming language by National Instruments based on the general purpose programming language, *G*. It is based on the premise of virtual instruments, which turn real instruments into software controllable instruments[3]. Instruments with various interfaces like RS-232 and RS-485, can be controlled and automated with a computer running a LabVIEW application. Depending on the hardware configuration of the system and the communication protocol being used, it is possible to automate multiple devices in one application. Unlike text based languages, LabVIEW programs are created via block diagrams called Virtual Instruments (VIs), composed of icons connected by wires. Diagrams are then compiled into executable code. Another distinction between LabVIEW and other languages is LabVIEW runs independent loops in parallel automatically. On the other hand, procedural code requires the implementation of complex protocols in order to achieve this. This improves the efficiency of the program by allowing data to be processed while instruments are busy instead of sitting idle. With its extensive library, many tasks can be accomplished without knowing the underlying details that are necessary had a language such as C been used. Specifically, functions contained in the LabVIEW library take care of GPIB protocol, allowing the user to simply specify the command to be sent.

The first step was to establish communication between the device and computer. This was accomplished using NI Communicator and NI Spy software which, were supplied with the PCMIA GPIB card. These two programs were employed to resolve hardware and driver conflicts, such as resource sharing on the computer and allows the user to type dynamically type in commands to determine their behavior.

The necessary commands involved initial settings for the device, starting a sweep to acquire data, retrieving data from the machine, and tracking the peak. Some of these commands are contained in the IEEE 488.2 protocol: device identification, triggering and enabling status registers. All other commands are specific to the Agilent 4294A Impedance Analyzer. Three VIs were created in

LabVIEW to accomplish our goals. The main VI, Agilent4294AController.vi, addresses a bulk of the work including initialization, triggering sweeps, reading data from the device and controlling the other two VIs. Data parsing was done in the Agilent4294ADataParser.vi to format the data to be read in Excel. The last vi, Agilent4294APeakTracker.vi was used compensate for any drift or shifts in the peak that occurred.

Before the application is run, several settings need to be made on the impedance analyzer: setting the maximum of 801 points per sweep, bandwidth and point averaging to reduce noise in the data, choosing to measure the series capacitance and series resistance, finding the resonance frequency and setting the span to 1kHz. The file path for data to be recorded and the value for how many sweeps are to occur between recording sweeps are then supplied by the user in LabVIEW. The latter setting is further explained in the discussion of peak tracking. Finally, the impedance analyzer's trigger source must be set to GPIB. The connection between the machines is established once the Agilent4294Acontroller.vi is started. The first steps of the VI are to adjust settings that can only be changed remotely. These settings include the data transmission format from the impedance analyzer and setting register bits on the impedance analyzer to enable service requests (SRQ). SRQ's are interrupts generated by the impedance analyzer to determine the state of the machine. SRQ's were used to ensure that commands would not be sent while the machine was busy during a sweep. Initializing SRQ's was accomplished by setting control bits in the registry as shown in Figure 3[4]. The second bit of the Instrument Event Status Enable Register was set to 1 to enable the Manual Trigger Waiting bit. The machine sets this bit to 1 when a sweep triggered through the GPIB is completed. The third bit of the Service Request Enable register is also set to 1 to flag any

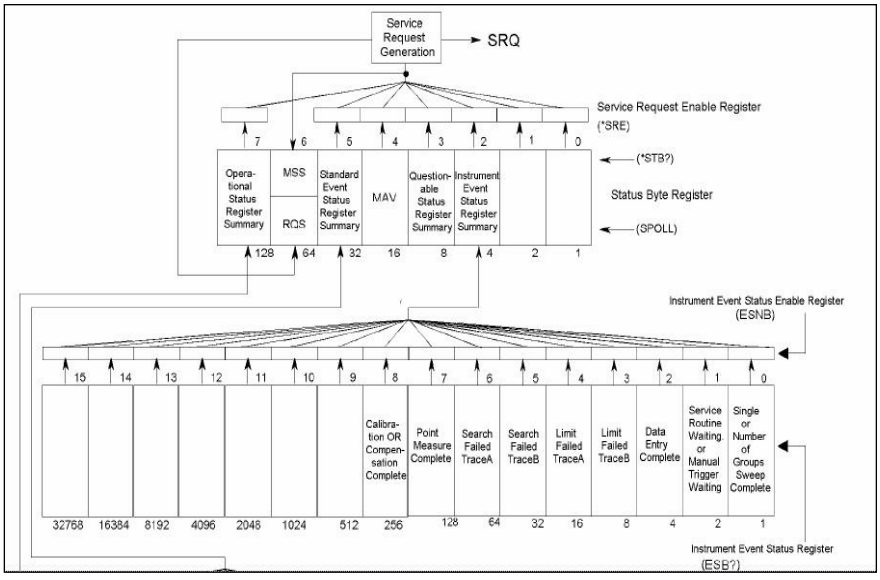


Figure 3. Agilent 4294A Event Status and Status Byte

changes that are enabled in the Instrument Event Status Register; in this case the Manual Trigger Waiting bit. By enabling the bits on these 2 registers, a SRQ is generated after a sweep has been completed. Finally, these Status registers are cleared in order to ensure that the prior status of these registers does not cause any false interrupts.

Using SRQ's makes the controller waits to send further commands until the sweep is completed. Depending on the bandwidth and point averaging settings, sweeps can take up to a minute. Commands sent during this time often abort the current sweep, resulting in no data being acquired. The solution of using interrupts is a more elegant solution rather than a constant delay or polling status of the device. By waiting for an interrupt to be generated, we take advantage of LabVIEW's built in ability of parallel execution of data. Polling the device would eliminate the ability to use the built in parallel processing while a constant delay would not account for measurement settings. This would change the timing that would result in aborted or inefficient sweeps.

Once the initial settings are made, the automated data acquisition begins by triggering a sweep. The system waits for the SRQ signifying that the sweep is complete. Once an SRQ is generated the machine is ready for more commands and the status registers are cleared for the next interrupt. The VI instructs the machine to transmit the sweep parameters (x-values), the series capacitance data, and the series resistance data. Data is transmitted with an 8-bit header that consists of "#6", the number of bytes being transmitted followed by the real and complex values of each point. Each real and complex number is transmitted as an IEEE 32-bit floating-point number. This data is stored in the temporary files SWPRM, the x-values, TRACA, the series capacitance values from trace A, and TRACB, the series resistance values from trace B. The block diagram for Agilent4294AController.vi is shown in Figure 4.

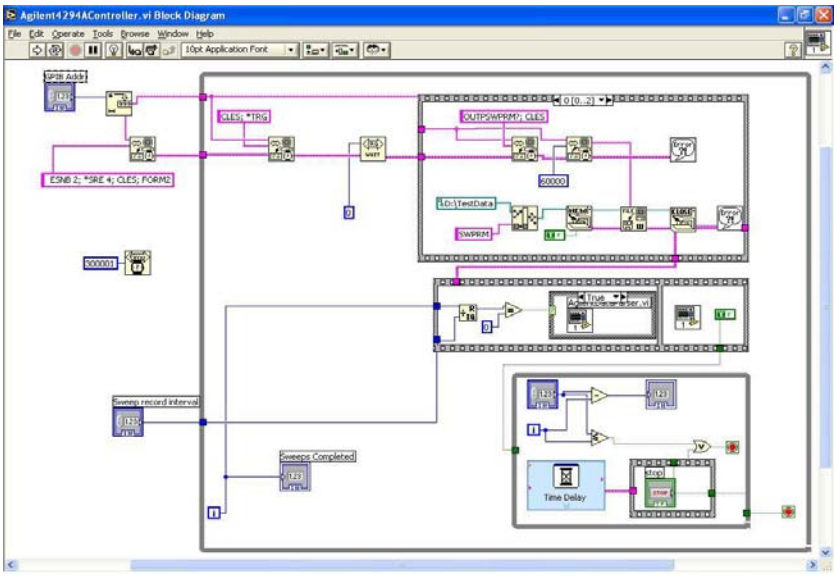


Figure 4. Agilent4294AController.vi block diagram.

After the data has been transferred and stored in temporary files, the Agilent4294ADataParser.vi is called. The information from the temporary files is opened using the Read From Sgl File.vi which reads the raw data stream into an array format. Imaginary values transmitted at this time since there are no imaginary components when measuring the series resistance or capacitance. The 3 values are appended to form a single array. This data is time stamped and appended to the rest of the data that has been recorded in the file that was initially specified by the user (figure 5).

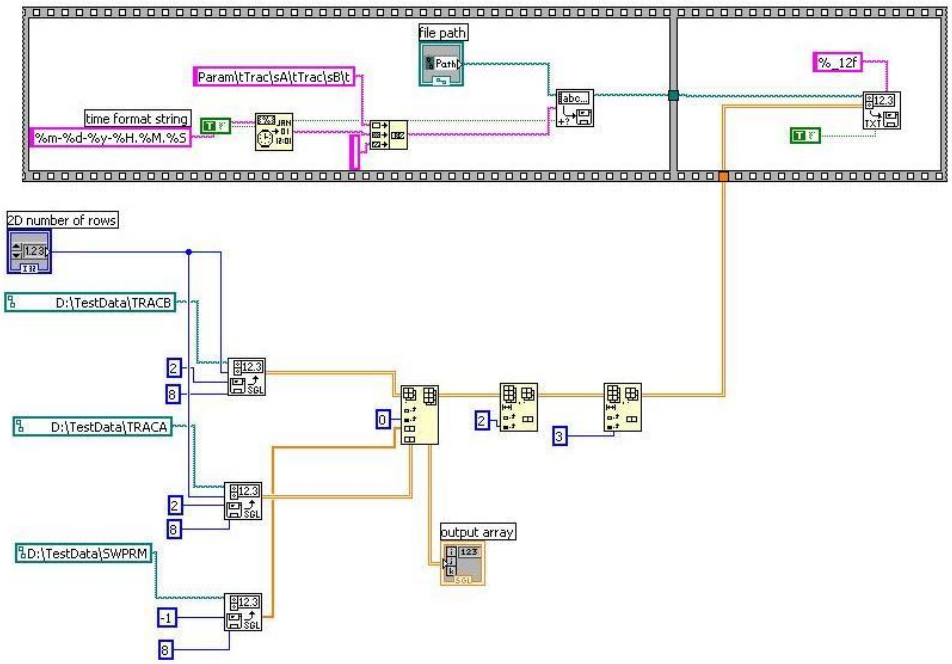


Figure 5. Agilent4294ADataParser.vi Block Diagram

The final step in an iteration is to monitor the peak. In some cases the resonance frequency would drift out of the initial 1KHz span. In order to address this issue, the Agilent4294APeakTracker.vi was created. This VI operates by querying the machine of the maximum series resistance value. This response is parsed and a command is sent to set the center value of the sweep to the x-value for the maximum series resistance from the last sweep (Figure 6). This did not ensure that the peak of the next sweep would be perfectly centered but it helped track the peak so it would remain within the 1kHz span. The next issue arising from tracking the peak was when significant drift occurred. However, the peak tracking is limited. If between sweeps the peak is out of the 1kHz range, the centering value, would be the next highest value which is likely to be noise. The magnitude of this problem depends on the rate of the drift and how often a sweep is made. If the rate of drift is around a 100Hz a minute, and sweeps are done

every 5 minutes, the peak would drift 500Hz by the next sweep and would be lost (assuming the peak was centered leaving 500Hz on either side of the center). There were several solutions discussed in solving this issue. One was to expand the span of 1kHz. By increasing the span, we measure a larger range that is limited by the maximum of 801 points, resulting in a loss in data resolution. Moreover, the more useful data is around the peak. At spans larger than 1kHz, much of the data lies at the extremes and is of little use. Another option was to do a quick scan over a fairly large area, find the peak and do another sweep with a smaller span around the peak. This was much more complex and could cause many other issues that would have to be taken into consideration. The option that was chosen was to continuously perform sweeps, each taking approximately 1-minute. The obvious problem to this is the vast amount of data storage that would be necessary. With each sweep being approximately 30KB and occurring every minute, over a 12 hour period the resulting file would be 21 MB file and would be an excessive amount of data to work with. Instead, an option was given to the user before the program is run to choose how often to record sweeps. The user could choose to record one out of every 5 sweeps resulting in about a 5-minute delay. The difference with the original solution is during that 5 minutes, sweeps are still occurring and the peak is still being tracked but only every fifth sweep is being recorded. This was the simplest solution, permitting us to keep a fairly high resolution while still being able to track the peak.

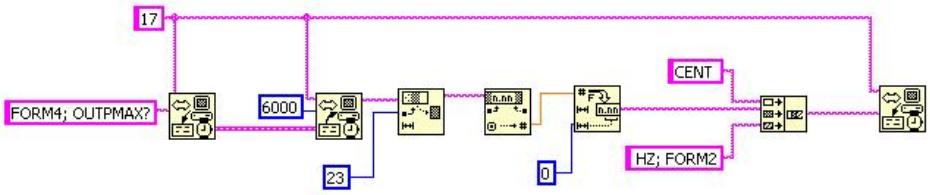
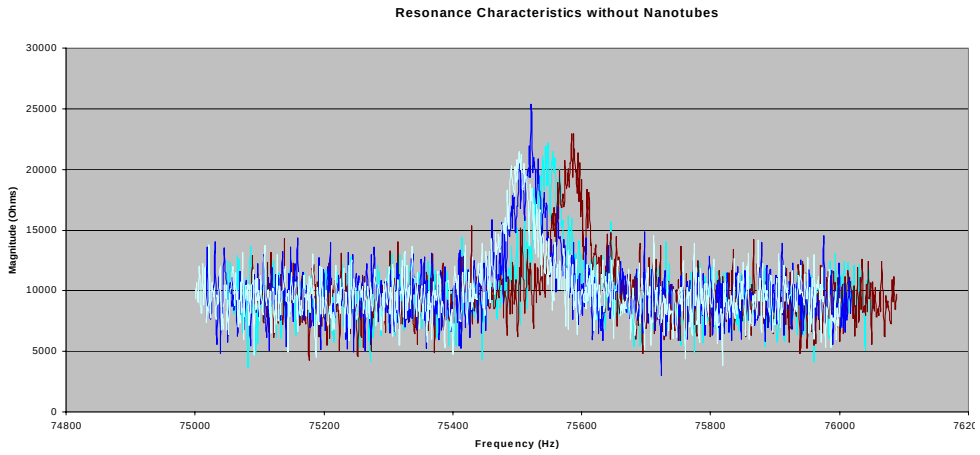
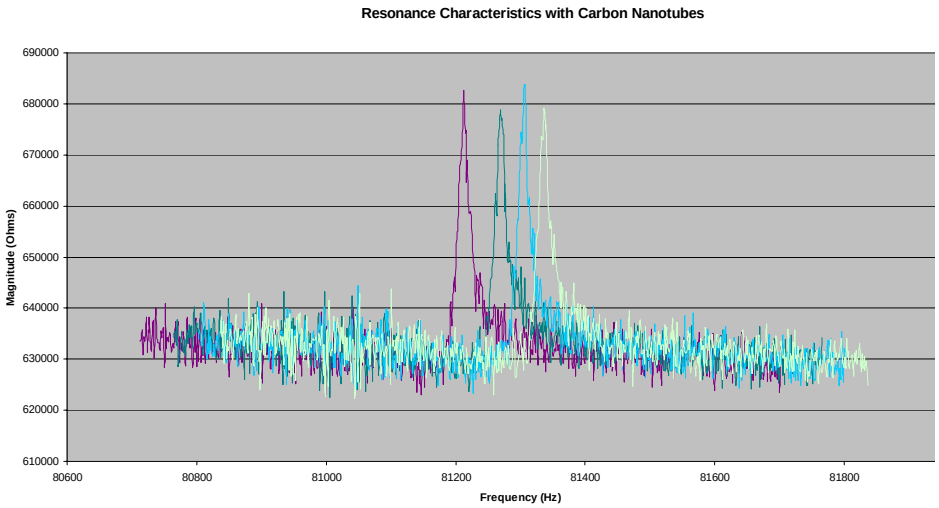


Figure 6. Agilent4294APeakTracker.vi block diagram

RESULTS & DISCUSSION



Graph 1: Plot of resonance characteristics for micro structure prior to carbon nanotube deposition



Graph 2: Plot of resonance characteristics for micro structure with carbon nanotubes deposited

Table 2. Q-factor with and without nanotubes

	Without Nanotubes	With Nanotubes
1	1508.72	3715.20
2	1485.00	3732.89
3	1330.27	3853.14
4	1384.85	3682.84
Avg	1427.21	3746.02

Graphs 1 and 2 show measurements of the resonance characteristics. The data displayed is an hourly sampling in the last 4 hours of measurement. Testing was done overnight for more than 12 hours. The resonance peaks of the device with carbon nanotubes was much sharper. The theoretical equation for a bridge is[5]:

$$f_i = \frac{\lambda_i^2}{2\pi L^2}\left(\frac{E'I}{m}\right)^{1/2}$$

where m is mass per unit length, I is the area moment of inertia, L is the length of the beam, $\lambda_1 = 4.730$ and E' is the plane strain Young's modulus. Based on the dimensions from Table 1, with *I* equal to:

$$I = \frac{h^3 * w}{12} = \frac{(3*10^{-6})^3 * (2*10^{-4})}{12} = 7.74*10^{-11}\ m^4$$

and *m* equal to

$$m = \rho * h * w = 2331 * (3*10^{-6}) * (2*10^{-4}) = 1.4*10^{-6}\ kg/m$$

where h is the thickness of the structure, w is the width and ρ is the density we find the resonance frequency as:

$$f_0 = \frac{4.73^2}{2\pi(6*10^{-4})^2}\left(\frac{172*10^9*7.74*10^{-11}}{1.4*10^{-6}}\right)^{1/2} = 73.58\ kHz$$

The theoretical resonance frequency is close to that of the device with not carbon nanotubes but not the device with nanotubes. The Young's modulus of carbon nanotubes is over 1 TPa. It is possible that this large Young's modulus of the carbon nanotubes has an effect in increasing the Young's modulus of the bridge. From the equation above, an increase in the Young's modulus, E', would increase the resonance frequency as seen in the graphs. We all so see a drastic change in the Q-factor from Table 2. The Q-factor was determined by fitting the data to a Lorentzian curve. The equation for the Q-factor is:

$$Q = \frac{f_0}{\Delta f}$$

with f_0 being the resonance frequency and Δf is the width at the half power point. An increase in the resonance frequency will cause an increase in the Q-factor. The magnitude of this change does not correlate to the average Q-factor being more than twice that of the initial test. The effect of the nanotubes is more than just increasing the resonance frequency.

The device after being coated with the nanotubes also continues to show drift even after being under test for a long period of time. On the other hand, the

resonance frequency of the device before carbon nanotubes shows little drift after the overnight measurements. Gas desorption could be playing an effect.

A drastic change in the magnitude of the device is also seen. The device's magnitude originally ranged from 5-25 kΩ. This range became 630-680 kΩ after the nanotubes were deposited. An explanation for this observation has yet to be made.

CONCLUSION

It is difficult to come to a definitive conclusion on the effects of the carbon nanotubes on the bridge, as these are preliminary tests. There are few if any experiments or theories relating to this experiment. However, the changes observed are enough to motivate further studies in this topic. The automation measurement equipment will greatly help in later experiments, as it is apparent that there are changes even after testing over night. The experiment should be repeated in order to verify that the results from the tests were not due to human error. The system used to deposit nanotubes presented difficulties in uniformly depositing the material on the devices. One possibility is to have the carbon nanotubes deposited in a different manner such as using critical point drying. Applying various pressures in the vacuum and observing the effects may also help in better understanding the effect of nanotubes on the resonance characteristics.

ACKNOWLEDGMENTS

I would like to thank Prof. Tadigadapa for giving me the opportunity to work with him. I greatly appreciate the exposure to a field that I knew little about. After my experience here, I intend on going to Graduate School, and MEMS technology will be one of my main areas of interest. I would also like to thank my graduate mentor Abhijat Goyal for his help and Prasoon for his literally backbreaking work in coating the devices with nanotubes. Outside of work, I would like to thank Plush, Lauren and Amanda for entertaining me on the weekends. And last Lindsay Holliday for fixing my horrible grammar. Also the NSF Grant that funded my research.

REFERENCES

[1] Pierron, Oliver. "EE 597F: MEMS Device Technology Spring 2003 Laboratory Report."

[2] MicroFab Technologies, Inc. *MJ-AB User's Manual.* Plano, TX, 19?.

[3] Johnson, Gary. *LabVIEW Graphical Programming: Practical Applications in Instrumentation and Control.* McGraw-Hill Inc., New York City, NY, 1997.

[4] Agilent Technologies, *Agilent 4294A High Precision Impedance Analyzer Programming Manual*. Part No. 04294-90061. Japan 2003.

[5] R. Blevins, "Formula for natural frequency and mode shape", Krieger ed., 1995 p.108.

ANALYSIS OF $Ba_{0.5}Sr_{0.5}TiO_3$(BST):MgO BULK CERAMICS COMPOSED OF NANOSIZED PARTICLES FOR THE PYROELECTRIC IR SENSORS

James R. Berninghausen*, Shashnk Agrawal[+], and A. S. Bhalla[#]

Department of Electrical Engineering and Materials Research Institute
Pennsylvania State University, University Park, PA 16802

*Undergraduate student of:
Department of Electrical Engineering and Computer Science
University of Portland, Portland, OR 97203

ABSTRACT

Fine structures of pyroelectric ceramics are important in making high density sensing arrays. In the current approach, the pyroelectric and dielectric properties of $Ba_{0.5}Sr_{0.5}TiO_3$ (BST) and MgO combined in multiple weight ratios in bulk ceramic form using micro and nanosized powders were studied. The basis for comparison is figure of merits used for point detectors. Measurements were conducted over a temperature range of 100-300K and at frequencies of 1, 10, 100, and 1000 kHz. The effort is to produce a higher figure of merit by decreasing the dielectric constant and loss tangent faster than the pyroelectric coefficient. Because BST has a high dielectric constant, MgO, which has a very low dielectric constant, is combined with it in an effort to reduce the overall dielectric constant. Results show that the dielectric constant and loss tangent can be tailored by the weight percentage of MgO in the sample to produce a figure of merit, which is an order of magnitude greater than that of that of pure BST.

[#] Faculty Mentor
[+] Graduate Mentor

INTRODUCTION

Pyroelectric materials are polar materials that respond electrically to changes in temperature. The material senses changes in heat through the infrared radiation emitted by bodies at temperatures above absolute zero provided the material has internal noise levels smaller than the incoming heat flux. The pyroelectric, or infrared (IR), sensors have a broad range of applications including fire and burglar alarms, night vision cameras, vidicon, and medical diagnostic instruments. The main sensing elements in these devices are ferroelectric or pyroelectric materials. Monitoring or tailoring the properties of pyroelectric materials can enhance the quality of sensing. On the other hand, for IR imaging and vision applications, very fine (or high density) array designs are desirable in which case the fine structures of ceramic pyroelectric sensing elements become an important issue. With this view the use of nanosize grains of the pyroelectric material for sensing element preparation also become an important factor.

At the heart of these materials is the dipole. Each electric dipole of a pyroelectric material can have a polarization that changes with temperature. As the charge on the ends of the dipole cannot change, the distance between the ends must fluctuate to cause a change in the dipole moment and therefore a change in polarization. The polarization we measure from the material can change in a couple different ways. The dipole can either change its distance or its orientation (pyroelectricity is a vector quantity). Since we attach electrodes at opposite ends, if a dipole tilts to one side and is no longer perpendicular to the electrodes, the polarization measured will be less. The internal structure and properties of these materials determine how the dipole will change with temperature and therefore how it will act as a pyroelectric sensor.

There are a wide variety of applications for pyroelectric material and there are often different properties that are more important than others for these different applications. In that sense, one type of material could be the best for one application while another could far surpass that material in a different application. Figures of merit are used to compare materials by taking into account their important properties and weighting those properties to best reflect the needs of the application. Bhalla and Guo[1] define several figures of merit for different applications.

$$F_D = \frac{p}{C_p \rho \sqrt{\varepsilon_0 k \tan \delta}} \qquad \text{pyroelectric point detector} \qquad (1)$$

$$F_i = \frac{p}{C_p \rho} \qquad \text{fast pulse detector} \qquad (2)$$

$$F_V = \frac{p}{C_p \rho \varepsilon_0 \kappa} \qquad \text{large area detector} \qquad (3)$$

$$F_{Vid} = \frac{p\gamma}{C_p \rho \varepsilon_0 \kappa} \qquad \text{vidicon} \qquad (4)$$

where, p is the pyroelectric coefficient at operating temperature, C_p is the specific heat, ρ is the density, ε_0 is the absolute permittivity in a vacuum, κ is the dielectric constant, γ is the reciprocal thermal diffusivity, and $tan\delta$ is the dielectric loss tangent. In this study we will limit ourselves to two simple figures of merit. The first is the most simple and is typically used for a single, relatively large sensor. The second is intended for a small point detector where we must take into account the dielectric loss because noise becomes a larger problem. The two figures of merit under consideration are:

$$F = \frac{p}{k} \qquad (5)$$

$$F = \frac{p}{\sqrt{k * \tan\delta}} \qquad (6)$$

Since the specific heat of a given material is a constant quantity, we signified the characteristics of a material in a simple form (equations 5 and 6).

Studies have shown that MgO doping has the greatest effect on reducing the loss tangents in bulk ceramics while also lowering the dielectric constant [2]. Both effects are desirable when these materials are used for pyroelectric detection.

(Ba,Sr)TiO_3 is one of the promising pyroelectric sensing materials and has been used in many of the state of the art IR sensing devices. Though the material is a promising one, it does have a large dielectric permittivity. In order to dilute the effects, composites of BST:MgO are studied in this study. It is also important that the two phases do not react during the sample preparation process and the resultant composite maintains the desired micro-nano structures of the two phases.

Conventional sintering, heating the ceramic in an oven, was not used in this study. Because MgO tends to absorb water, the goal was to achieve as low porosity as possible and therefore the highest possible density. This would prevent, or at least reduce, any water absorption that may affect the material. For this reason, microwave sintering was used as it produces material with a high density and also an improved microstructure and dielectric properties [3]. This method also requires a much shorter time to sinter which helps maintain the chemical stoichiometry of the two individual constituents.

EXPERIMENTAL DESCRIPTION

The bulk ceramics were composed of $Ba_{0.5}Sr_{0.5}TiO_3$ (BST) and MgO with particle sizes around 50nm for the BST and 100nm for the MgO. Different

weight percentage ratios of BST and MgO were synthesized and tested. Powders were ball milled for 24 hours, pelleted, and sintered in a 6kW microwave oven at 2.45GHz. After polishing, electrodes were deposited through gold sputtering and silver wires were attached.

All measurements were done with an automated dielectric and pyroelectric measurement system. This system consists of an LCR meter (HP-4284, Hewlett-Packard Inc.), closed-cycle helium cooled cryostat (HC-2, APD Cryogenics Inc.), temperature controller (model 330, Lakeshore Inc.), vacuum pump (Pascal 05, Alcatel), high voltage source (model 610D, Trek) and DMM (HP-3478A, Hewlett-Packard Inc.) for pyroelectric measurements, and a desktop computer. The sample is isolated in the vacuum chamber from any external influences. When measuring dielectric properties, the LCR meter is connected to the sample and to the computer to measure capacitance and dielectric loss. For pyroelectric measurements an electric field is applied to the sample with the high voltage source to electrically pole the sample while cooling the sample from 300K to 100K. After polling, the voltage source is disconnected and the DMM is connected for measuring the pyroelectric voltage. See Figure 1.

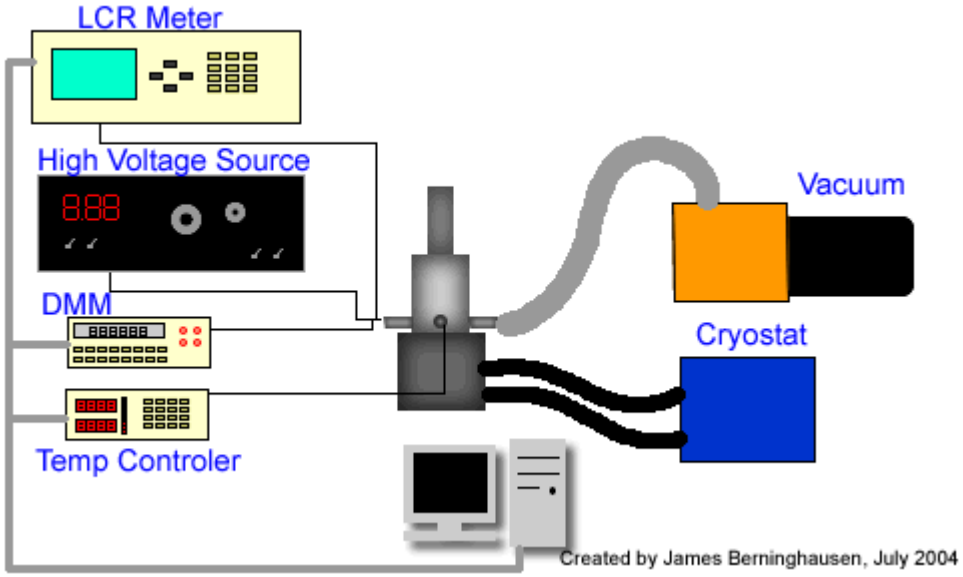


Figure 1. Experimental setup.

A sample is placed in the vacuum chamber on thermally conductive grease to prevent undesired movement and shorts from occurring. The silver wires are connected to the electrodes in the vacuum chamber and a vacuum is created. Once the vacuum is established, the dielectric or pyroelectric measurement can be started.

Dielectric measurements are conducted between 100K and 300K. The dielectric properties are first tested as the sample is cooled down to 100K. The

measurements are then repeated while the sample is heated back to 300K. The rate is kept constant at 2K/min from 300K to 100K and back to 300K.

Pyroelectric measurements are preceded with an applied electric field of 20kV/cm. The sample is cooled to 100K with electric field applied and the test is performed as it is heated. A second test is performed by poling the sample with the same field while cooling from 300K to 200K. The test then proceeds as the sample is slowly heated back up to 300K. For all pyroelectric measurements the rate of heating is about 1.67 degrees per minute (100 degrees per hour).

RESULTS AND DISCUSSION

As the ratio of BST to MgO is increased, the dielectric constant (k), the dielectric loss (tanδ), and the pyroelectric coefficient (p) all increase. This is to be expected as the material approaches the behavior of pure BST. The rates of these individual increases are dissimilar enough that there is a noticeable difference in the figures of merit that define good sensing materials.

All compositions showed a single pyroelectric and dielectric peak. A sample of pure BST 60 was tested for purposes of comparison and three distinct pyroelectric peaks were observed signifying three phase transitions. The first two phase transitions in BST are identified by these changes in polarization because the polar vector changes its orientation. In the rhombohedral phase the dipole is oriented along the body diagonal (111). It orients along a face diagonal (110) in the orthogonal phase. In the tetragonal phase the orientation is parallel to the sides (100) and the dipole gradually decreases its effective distance as it approaches the cubic region. It appears that the tetragonal-cubic phase transition is suppressed by the addition of the MgO. This is perhaps due to the non-uniform internal molecular stresses caused by the addition of MgO that has instead made the orthorhombic-tetragonal phase change more pronounced. The dielectric curve is considerably broader than that of pure BST, which is to be expected from a composite. The polarization curve also behaves as expected in a composite as having a more diffuse characteristic rather than a first order characteristic.

Table 1. Pyroelectric Properties of $(Ba_{0.5}Sr_{0.5})TiO_3$:MgO composites

Ratio	Pyroelectric Coefficient ($\mu C/m^2K$)	Pyroelectric (p)Peak Temp. (°C)	Dielectric Constant (κ)	Loss ($\tan\delta$)	$\frac{p}{\kappa}$	$\frac{p}{\sqrt{\kappa \tan\delta}}$
40/60	69.026	-14.45	88.3709	0.0024	0.781	149.883
50/50	182.516	-16.15	162.1332	0.0019	1.126	328.843
60/40	366.76	-12.65	368.7367	0.0019	0.995	438.174
70/30	633.371	3.05	573.236	0.0033	1.105	460.505
80/20	764.633	-17.85	1025.8477	0.0017	0.745	579.011
100/0*	513	1.25	7773.13	0.0124	0.066	52.253

*$(Ba_{0.6}Sr_{0.4})TiO_3$

In general the maximum dielectric value occurred between –40°C and –60°C. It has been shown that adding MgO to BST causes a shift in this dielectric peak from around 0°C, at the tetragonal-cubic transition, to approximately -40°C and also broadens the peak[4]. This broad dielectric curve lacks noticeable phase transitions which is consistent with other BST:MgO nano composite research[5]. This, however, does not mean the phase transition itself occurs at a lower temperature. The slightly anomalous data recorded for the 70/30 sample requires more study as the loss and the pyroelectric peak temperature are both slightly higher than would be expected based on the other samples. This could be due to a poor sample or problems encountered during the poling of this specific sample. Another issue to consider for this sample is a phenomenon referred to in this paper later as the first run effect.

For purposes of comparison, a pure sample of BST was tested. Due to resource limitations only a 60/40 ratio of BST was available. Figures 8 and 13 show the behavior of this sample. The three distinct phase transitions are especially visible in the pyroelectric curve. Figure 2 shows the dielectric curve of a pure BST 50/50 sample prepared with similar nano-sized particles.

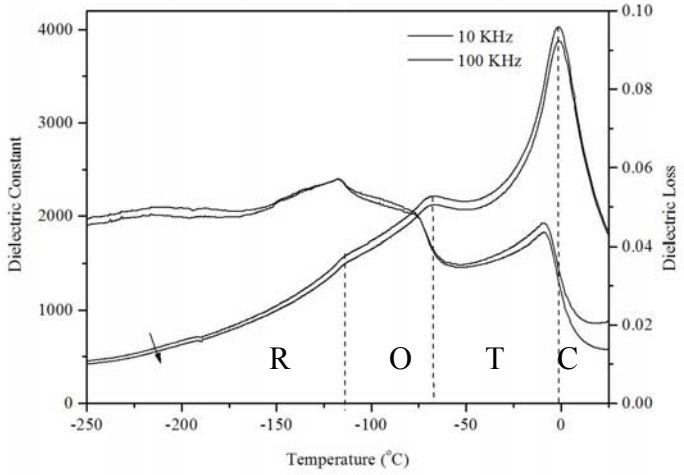


Figure 2. Temperature dependence of dielectric constant and tanδ of nano particle size $Ba_{0.5}Sr_{0.5}TiO_3$.[5] (R-Rhombohedral, O-Orthorhombic, T-Tetragonal, C-Cubic)

The following figures show the dielectric properties of each sample tested in the heating cycle. Samples were tested at four different frequencies: 1kHz, 10kHz, 100kHz, and 1MHz. The results show a decrease in the dielectric constant as more MgO is added. In all samples, the loss is significantly less than pure BST. Both of these reductions can be attributed to the low dielectric constant (~8) and the low loss (~0.0001) of MgO. Similar results were obtained in the cooling cycle with slightly higher dielectric peaks shifted to the left slightly.

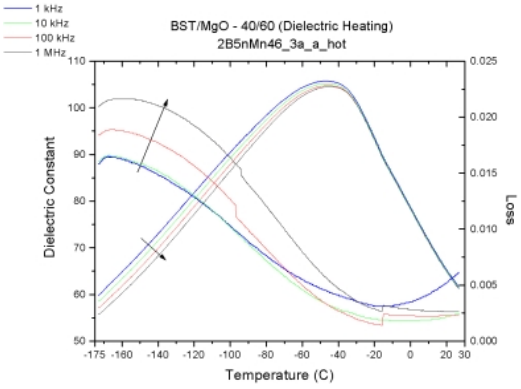


Figure 3. **40/60 Dielectric**

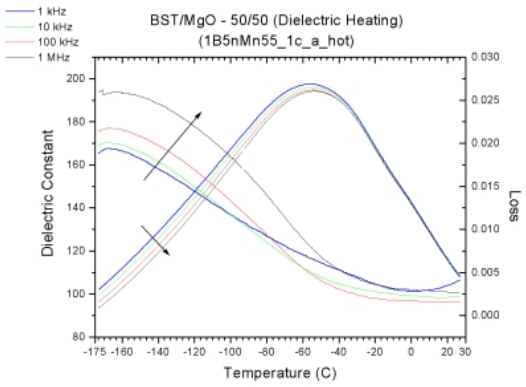


Figure 4. **50/50 Dielectric**

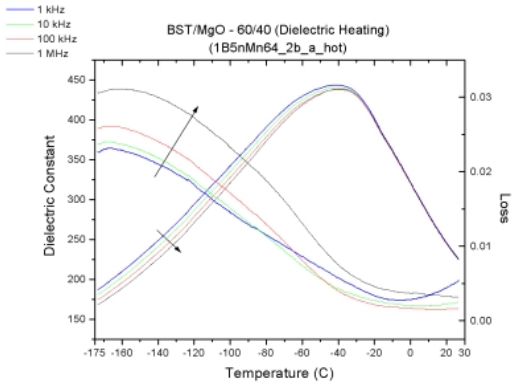


Figure 5. 60/40 Dielectric

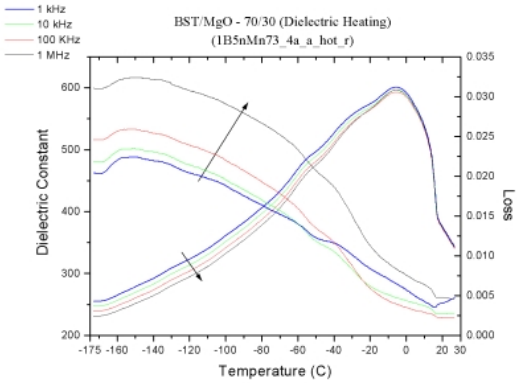


Figure 6. 70/30 Dielectric

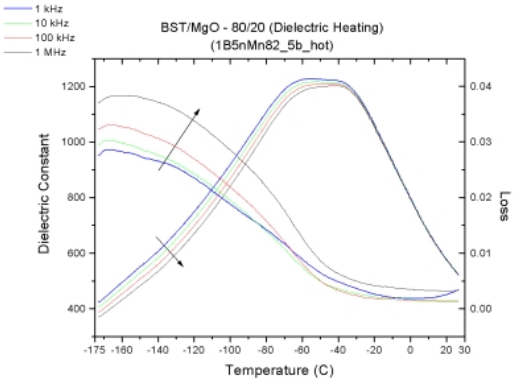


Figure 7. 80/20 Dielectric

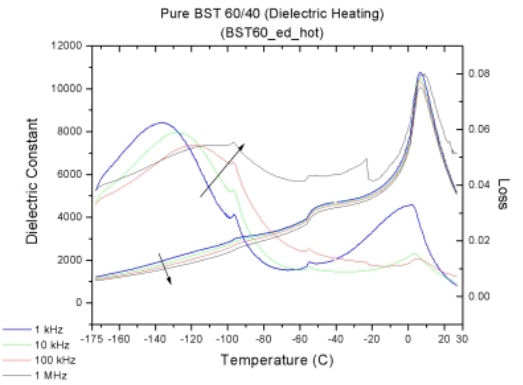


Figure 8. Pure BST Dielectric

While there is a slight deviation of the pyroelectric peak temperature, this is most likely caused by variances in the sample, as there is no trend related to the composition of each sample. As one would expect, the pyroelectric coefficient of the composite increases as it approaches pure BST.

The pyroelectric curves below show fairly consistent behavior from each of the samples.

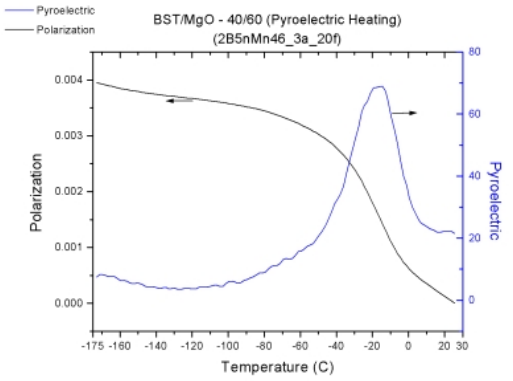


Figure 8. 40/60 Pyroelectric

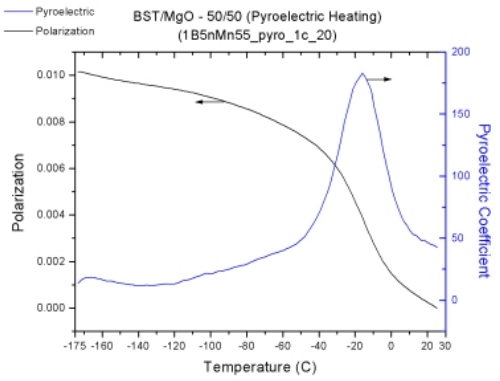


Figure 9. 50/50 Pyroelectric

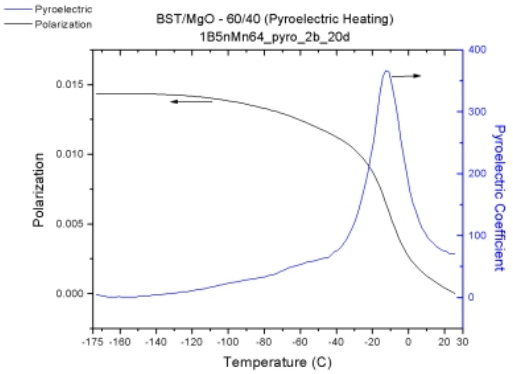


Figure 10. 60/40 Pyroelectric

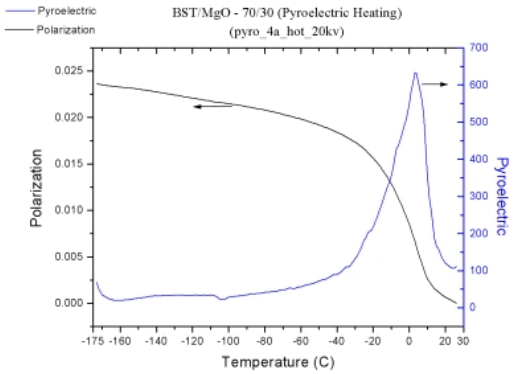


Figure 11. 70/30 Pyroelectric

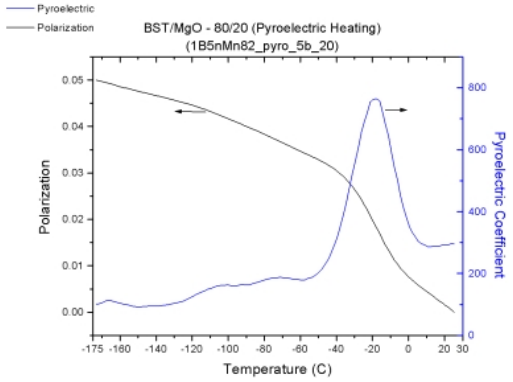


Figure 12. 80/20 Pyroelectric

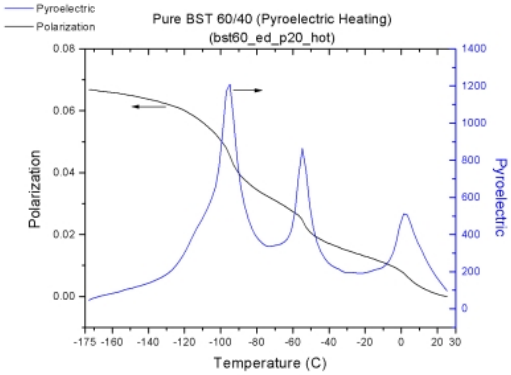


Figure 13. Pure BST Pyroelectric

First Run Effect

There was an interesting event that occurred during the first pyroelectric measurement in some of the samples. Dielectric measurements were performed first subjecting the samples to the temperature ranges off 100-300K. When the sample was poled for the first time at 20kV/cm from 300K to 100K and then tested for pyroelectric properties, a small anomaly occurred. A considerably larger pyroelectric peak that was usually shifted to the right (higher temperature)

was measured in most of the samples. There are many possible explanations for this including stress on the material which would cause the shift and possible experimental errors. All of the tests were conducted using the same equipment and for some samples all tests were conducted without ever disturbing the sample or the test chamber. This effectively eliminates certain sources of error. A pure BST sample was tested to verify its characteristics and the equipment was able to record the expected values. The samples were retested and more consistent behavior was observed. The source of this anomaly will be interesting to study in ongoing research.

CONCLUSION

The addition of MgO to BST 50 proved to be advantageous in improving the pyroelectric detection properties. The dielectric constant was lowered considerably and that is reflected in the figure of merit. The pyroelectric coefficient was also decreased comparatively and the ratio of p/k increased. The dielectric loss (tanδ) was reduced by approximately a full order of magnitude making the figure of merit for this material greater than pure BST. The largest figures of merit were observed in the 70/30 and 80/20 samples. These results look promising in producing better IR sensors.

The figure of merit for the IR sensors was improved by an order of magnitude over values obtained with pure BST. It is desirable to explore composites of other BST compositions and MgO to select the best BST composition for IR sensing applications. The interesting effect the MgO had on the phase transitions is also worth looking at closer.

It is noticed that the C-T transition is very sensitive to the composite formation. It appears to be weaker in the composite where as the T-O transition becomes stronger. As a result, pyroelectric coefficients are higher where as the corresponding dielectric constant and loss at that temperature are lower, giving high IR sensing characteristics. Another aspect of these composites that deserves attention is their structure and method of synthesis.

Microwave sintering apparently helped to maintain the chemical integrity of the components in the composite as is evident from the electrical measurements and phase transitions of BST. A complete microstructure and property relationship study was not performed in the current project but there are definite positive effects seen by using nano-size particles based on work carried out by other members of the group[6].

For future work in this area more detailed studies are needed to examine the composite effects on the various phase transitions in the case of different $(Ba_{0.5}Sr_{0.5})TiO_3$ composites and their ultimate effect on IR pyroelectric sensors in different temperature region applications.

ACKNOWLEDGEMENT

I would like to first thank Professor Bhalla for his patience and attentive mentoring which allowed me to conceptually understand the pyroelectric phenomena in a short period of time. His expert opinion and wealth of knowledge helped me understand my results. Shashnk Agrawal was indispensable as a resource throughout my research. His experience and guidance in the lab helped make this research possible. Also, I would like to thank Dr. Edward Alberta for his help with the equipment and for helping verify our results.

This material is based upon work supported by the National Science Foundation under Grant No. EEC-0244030.

REFERENCES

[1] A.S. Bhalla, R. Guo, "Pyroelectricity," *Wiley Encyclopedia of Electrical and Electronics Engineering*, 17, 465-469, 1999.

[2] L.C. Sengupta, S. Sengupta, "Breakthrough advances in low loss, tunable dielectric materials," Mat Res Innovat, 2:278-282, 1999.

[3] O.P. Thakur, C. Prakash, D.K. Agrawal, "Dielectric Behavior of $Ba_{0.95}Sr_{0.05}TiO_3$ ceramics sintered by microwave," Materials Science and Engineering, B96, 221-225, 2002.

[4] Wontae Chang, Louise Sengupta, "MgO-mixed $Ba_{0.6}Sr_{0.4}TiO_3$ bulk ceramics and thin films for tunable microwave applications," Journal of Applied Physics, 92, 3941-3946, 2002.

[5] Shashnk Agrawal, R. Guo, D. K. Agrawal, A. S. Bhalla, R. R. Neurgaonkar, C. B. Murray, "Dielectric Tunability of BST:MgO composites prepared by using nano particles," Ferroelectric Letters. (To Be Published.)

[6] S. Agrawal et al. (unpublished results)

MODELING DIFFRACTION PATTERNS FROM PERIODIC NANOSTRUCTURES USING RIGOROUS COUPLED-WAVE ANALYSIS

Andrew N. Fontanella,* Kebin Shi[+], and Zhiwen Liu[#]

Department of Electrical Engineering
The Pennsylvania State University
University Park, PA 16802

*Undergraduate Student of
Department of Electrical Engineering
The Pennsylvania State University
University Park, PA 16802

ABSTRACT

This paper presents a general description of the rigorous coupled-wave analysis and its implementation in MATLAB based simulation software. The focus of this work is the accurate modeling of diffraction patterns produced by binary grating structures. Further, the intention is to apply this procedure to supercontinuous white light sources. In doing so, a unique diffraction pattern will be observed as a function of wavelength. From this diffraction pattern profile, a unique grating structure can be inferred. This procedure thus has the potential for application towards nano-scale metrology and optical memory. As an initial step towards this goal, the diffraction efficiency patterns of a binary structure are simulated as a function of wavelength. This is done for both the TE and TM modes of polarization. Experimental and published results are then analyzed to support the validation of this technique.

[#] Faculty Mentor
[+] Graduate Mentor

INTRODUCTION

As the reduction in feature size of integrated circuit components continues into the range of tens of nanometers and smaller, it has become necessary to devise novel methods towards the extraction of information regarding the accuracy of the physical fabrication process. At these sizes, many difficulties arise in gathering information about the physical characteristics of the fabricated patters. Direct optical observation techniques become impossible, and scanning electron microscope metrology can be damaging to the fragile structure of the circuitry. Ideally, real-time techniques would be employed following fabrication in order to identify defective wafer batches and alert to the need for adjustment to the process [1]. This necessity has led to the development of diffraction based optical methods which are both efficient and non-destructive. Theses are based on techniques in which a wafer test-region imprinted with a periodic grating structure is exposed to spatially-coherent light at a range of angles along the plane formed by the grating vector and grating normal, or in another approach, at a single frequency and varying wavelengths. A reflected zero-order diffraction angle profile is thus generated as a function of some variable. This yields a solution to the forward problem of obtaining a unique diffraction intensity function which correlates to the physical parameters of the grating structure. The inverse problem entails the reconstruction of the two-dimensional grating profile using the information thus obtained. To this end, physical parameter prediction methods compare experimental data to theoretical models using computational techniques, and a best-fit model is produced [2].

A unique refinement to this process proposes the use of a supercontinuum white-light source. In implementing this proposal, an additional chromatic dimension of variability is introduced to the process without the need for a time dependant wavelength sweep. This may be used in conjunction with angular based techniques in order to enhance collected data, or it may find application solely, thus circumventing the need for the dynamic metrology setup or time dependant measurements that are necessary for angular based methods and sweeping wavelength sources, respectively.

The highly spatially coherent continuum of wavelengths required for such a process has only recently been realizable thanks to the development of “white-light laser” technology. This involves the coupling of a short duration laser pulse into a specialized optical fiber. The nonlinear effects of the fiber effectively expand the frequency domain spectrum of the electromagnetic pulse such that the emergent beam is a superposition of wavelengths ranging from the infrared to the low ultraviolet.

Crucial to the successful implementation of this wavelength based method is the ability to predict diffraction patterns based on a set of grating physical parameters. Rigorous coupled-wave analysis was chosen for this purpose, as it has been shown to produce numerically accurate results for periodic structures while minimizing the computational complexity of the procedure. This paper

comprises a step towards the satisfaction of the need for a multiple wavelength simulation tool based on the formulation of a computational procedure founded on the principles of rigorous coupled-wave analysis [3][4].

PROCEDURES[#]

The binary grating profile was subdivided into three regions consisting of two homogenous spaces (Region I and Region II) and an intervening layer composed of the ridges and groves of the grating (Fig. 1). The periodic boundary layer is expressible as a Fourier series of the form

$$\varepsilon(x) = \sum_{h=-\infty}^{\infty} \varepsilon_h \exp\left(j \frac{2\pi h}{\Lambda}(x) \right), \tag{1}$$

where

$$\varepsilon_{h=0} = (n_{II})^2 f + (n_I)^2 (1-f), \tag{2}$$

$$\varepsilon_h = [(n_{II})^2 - (n_I)^2] \left(\frac{\sin(\pi h f)}{\pi h} \right), \tag{3}$$

Λ represents the period of the grating, f represents the fraction of the grating period which the ridges occupy, and n_I and n_{II} are the refractive indices of the media in Region I and II, respectively. Although the summation of $\varepsilon(x)$ is over an infinite range, sufficient numerical accuracy was obtained with the retention of 101 Fourier harmonics symmetrical about h=0.

The wave vector components for the various diffracted orders are defined as

$$k_{xm} = k_0 [n_I \sin\theta - m(\lambda_0 / \Lambda)], \tag{4}$$

$$K_{L,zm} = \begin{cases} k_0 [(n_L)^2 - (k_{xm}/k_0)^2]^{1/2}, & k_0 n_L > |k_{xm}| \\ -jk_0 [(k_{xm}/k_0)^2 - (n_L)^2]^{1/2}, & k_0 n_L < |k_{xm}| \end{cases} \quad L\text{=I, II} \tag{5}$$

where m represents successive integer values from $-\sigma/4$ to $\sigma/4$, and σ is equal to one less than the number of Fourier harmonics retained in Eq.(1). Λ is the grating period, λ_0 is the wavelength of the impingent wave in Region I, θ is the angle between the incident beam and the grating normal, and $k_0 = 2\pi/\lambda_0$. The number n will be defined as $\sigma/2+1$.

[#] The computer implementation of the RCWA was based, to a large degree, upon the mathematical foundation laid by Moharam *et al* [4]. The mathematics which appears herein is an adaptation of their work.

For the TE mode of propagation (i.e. the orientation shown in Fig. 1 where the electric field vector runs along the y-axis, which is perpendicular to the grating vector), the tangential component of the electric field within the grating region was expressed as the sum of the n spatial harmonic components. The n spatial harmonic components of the tangential magnetic field were then derived from the electric field components. Through the use of Maxwell's equations, the harmonic components were then expressed solely in relation to each other. This yielded a set of two coupled-wave equations which were reduced to a single set of equations whose size was determined by the number of spatial harmonics retained within. Put into matrix form, this set of equations is expressible as

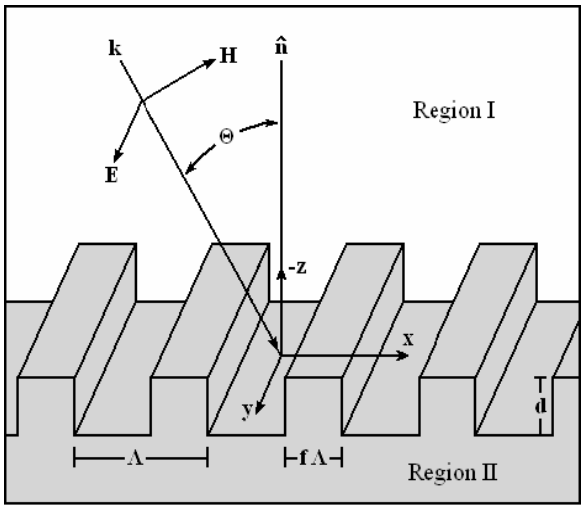


Fig. 1 Geometry for the TE mode.

$$[\partial^2 \mathbf{S}_y / \partial(k_0 z)^2] = [\mathbf{A}][\mathbf{S}_y], \tag{6}$$

where $\mathbf{S_y}$ represents the set of normalized-amplitude spatial harmonics of the electric field within the grating region. The matrix **A** is defined by the equation

$$\mathbf{A} = \mathbf{K}_x^{\,2} - \mathbf{E}, \tag{7}$$

where $\mathbf{K_x}$ is an n×n diagonal matrix whose (m,m) element is equal to k_{xm}/k_o. The n×n matrix **E** consists of elements for which the (i,j) element is equal to $\varepsilon_{(i-j)}$, where $\varepsilon_{(i-j)}$ is determined from Eqs. (1) and (2). The eigenvalues and eigenvectors of **A** were then computed. The eigenvalues were set into a diagonal n×n matrix **Q**, whose elements are equal to the positive square root of the eigenvalues of **A**. The eigenvectors were set into the columns of an n×n matrix **W**. From these matrices, the matrix **V** was computed, where **V**=**WQ**.

The next step involved matching the tangential electric and magnetic field components at the boundary regions of the grating (i.e. z=0 and z=d). A set of unknown constants, c_i^+ and c_i^-, (for which n integer values of i are retained) arise from and are dependant upon the boundary conditions of the grating. At z=0,

$$\delta_{m0} + R_m = \sum_{i=1}^{n} w_{m,i}[c_i^{\,+} + c_i^{\,-} \exp(-k_0 q_i d)], \tag{8}$$

$$j[n_1 \cos(\theta)\delta_{m0} - (k_{I,zm}/k_0)R_m] = \sum_{i=1}^{n} v_{m,i}[c_i^{+} - c_i^{-} \exp(-k_0 q_i d)], \qquad (9)$$

where d is the groove depth of the grating; $w_{m,i}$, $v_{m,i}$, and q_i are the (m,i) elements of the **W** and **V** matrices and the (m,m) element of the **Q** matrix, respectively; R_m is the amplitude of the m^{th} component of the backwards diffracted field; and δ_{m0}=1 for m=0 and δ_{m0}=0 for m≠0. Substituting for the unknown R_m, a single set of equations is generated. A similar procedure is followed in the manipulation of the z=d boundary condition equations

$$T_m = \sum_{i=1}^{n} w_{m,i}[c_i^{+} \exp(-k_0 q_i d) + c_i^{-}], \qquad (10)$$

$$j[(k_{II,zm}/k_0)T_m = \sum_{i=1}^{n} v_{m,i}[c_i^{+}(-k_0 q_i d) - c_i^{-}]. \qquad (11)$$

T_m is the amplitude of the m^{th} component of the forward diffracted field. The two sets of equations thus obtained are set into matrix form and the unknown constants c_i^{+} and c_i^{-} are solved for.

Solutions for R_m and T_m are obtained by solving the matrix equations

$$\begin{bmatrix} \delta_{mo} \\ jn_I \cos(\theta)\delta_{m0} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -j\mathbf{Y_1} \end{bmatrix} [\mathbf{R}] = \begin{bmatrix} \mathbf{W} & \mathbf{WX} \\ \mathbf{V} & -\mathbf{VX} \end{bmatrix} \begin{bmatrix} \mathbf{c}^{+} \\ \mathbf{c}^{-} \end{bmatrix}, \qquad (12)$$

$$\begin{bmatrix} \mathbf{I} \\ j\mathbf{Y_2} \end{bmatrix} [\mathbf{T}] = \begin{bmatrix} \mathbf{WX} & \mathbf{W} \\ \mathbf{VX} & -\mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{c}^{+} \\ \mathbf{c}^{-} \end{bmatrix}, \qquad (13)$$

where $\mathbf{Y_1}$ and $\mathbf{Y_2}$ are diagonal, n×n matrices whose (m,m) elements are $(k_{I,zm}/k_0)$ and $(k_{II,zm}/k_0)$, respectively. The diagonal, n×n matrix **X** has $\exp(-k_0 q_i d)$ for it's (i,i) element. Finally, the diffraction efficiencies for each particular spatial harmonic m are defined by the equations

$$DE_{rm} = R_m R_m^{*} \operatorname{Re}\left(\frac{k_{I,zm}}{k_0 n_1 \cos\theta} \right), \qquad (14)$$

$$DE_{tm} = T_m T_m^{*} \operatorname{Re}\left(\frac{k_{II,zm}}{k_0 n_1 \cos\theta} \right), \qquad (15)$$

Taking the real component of the $k_{L,zm}$ terms effectively filters out any evanescent waves and retains a limited number of propagating diffraction orders.

For the TM mode of propagation (i.e. the orientation shown in Fig. 2 where the magnetic field vector runs along the y-axis, which is perpendicular to the grating vector), a similar solution method was followed, this time with the tangential component of the electric field vector derived from that of the magnetic field using Maxwell's equations. This eventually lead to the derivation of the matrix equation

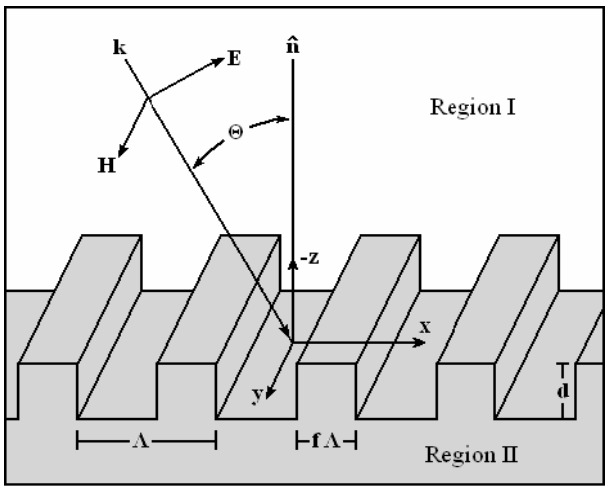


Fig. 2. Geometry for the TM mode.

$$[\partial^2 \mathbf{U}_y / \partial(k_0 z)^2] = [\mathbf{EB}][\mathbf{U}_y], \tag{16}$$

where $\mathbf{U_y}$ represents the set of normalized-amplitude spatial harmonics of the magnetic field within the grating region. The matrix **B** is defined by the equation

$$\mathbf{B} = \mathbf{K}_x^{\,2}\mathbf{E}^{-1}\mathbf{K}_x - \mathbf{I}, \tag{17}$$

where $\mathbf{K_x}$ and **E** are as defined previously, and **I** is an n×n identity matrix. The eigenvalues and eigenvectors of **EB** were then computed. The eigenvalues were set into a diagonal n×n matrix **Q**, whose elements are equal to the positive square root of the eigenvalues of **EB**. The eigenvectors were set into the columns of an n×n matrix **W**. From these matrices, the matrix **V** was computed, where

$$\mathbf{V}=\mathbf{E}^{-1}\mathbf{W}\mathbf{Q}\,. \tag{18}$$

As in the TE case, the tangential electric and magnetic field components at the boundary regions of the grating (i.e. z=0 and z=d) are matched and a set of unknown constants are solved for using the boundary condition equations. At z=0,

$$\delta_{m0} + R_m = \sum_{i=1}^{n} w_{m,i}[c_i^{\,+} + c_i^{\,-}\exp(-k_0 q_i d)], \tag{19}$$

$$j\left[\left(\frac{\cos\theta}{n_I}\right)\delta_{m0} - \left(\frac{k_{I,zm}}{k_0 n_I^{\,2}}\right)R_m\right] = \sum_{i=1}^{n} v_{m,i}[c_i^{\,+} - c_i^{\,-}\exp(-k_0 q_i d)] \tag{20}$$

Substituting for the unknown R_m, a single set of equations is generated. A similar procedure is followed at z=d:

$$T_m = \sum_{i=1}^{n} w_{m,i}[c_i^{+} \exp(-k_0 q_i d) + c_i^{-}], \tag{21}$$

$$j\left(\frac{k_{II,zm}}{k_0 n_I^{\,2}}\right) T_m = \sum_{i=1}^{n} v_{m,i}[c_i^{+}(-k_0 q_i d) - c_i^{-}]. \tag{22}$$

As in the TE mode, the two sets of equations thus obtained are set into matrix form and the unknown constants c_i^{+} and c_i^{-} are solved for.

Solutions for R_m and T_m are obtained from the matrix equations

$$\begin{bmatrix} \delta_{mo} \\ jn_I \cos(\theta)\delta_{m0}/n_I \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -j\mathbf{Z_1} \end{bmatrix} [\mathbf{R}] = \begin{bmatrix} \mathbf{W} & \mathbf{WX} \\ \mathbf{V} & -\mathbf{VX} \end{bmatrix} \begin{bmatrix} \mathbf{c}^{+} \\ \mathbf{c}^{-} \end{bmatrix}, \tag{23}$$

$$\begin{bmatrix} \mathbf{I} \\ j\mathbf{Z_2} \end{bmatrix} [\mathbf{T}] = \begin{bmatrix} \mathbf{WX} & \mathbf{W} \\ \mathbf{VX} & -\mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{c}^{+} \\ \mathbf{c}^{-} \end{bmatrix}, \tag{24}$$

where $\mathbf{Z_1}$ and $\mathbf{Z_2}$ are diagonal, n×n matrices whose (m,m) elements are $[k_{I,zm}/(k_0 n_I^{\,2})]$ and $[k_{II,zm}/(k_0 n_{II}^{\,2})]$, respectively. The diagonal, n×n matrix $\mathbf{X}$ has $\exp(-k_0 q_i d)$ for it's (i,i) element. The diffraction efficiencies for each spatial harmonic m are defined by the equations

$$DE_{rm} = R_m R_m^{*} \operatorname{Re}\left(\frac{k_{I,zm}}{k_0 n_1 \cos\theta}\right), \tag{25}$$

$$DE_{tm} = T_m T_m^{*} \operatorname{Re}\left(\frac{k_{II,zm} n_I}{k_0 n_1 \cos\theta}\right), \tag{26}$$

The experimental procedure for the verification of this program consisted of shining the "white light" beam through a transparent diffraction grating. At an angle of incidence of zero degrees, only the zero-order transmitted diffraction efficiency was measured. This is the order possessing the greatest power. Additionally, since non-zero diffraction order angles vary as a function of wavelength, measurements would have to be made along a range of angles to acquire non-zero order diffraction efficiencies. Measuring only the zero-order diffraction greatly simplified the experimental procedure, since for this order, all of the wavelengths are superimposed upon one another.

RESULTS AND DISCUSSION

An important aspect of the verification of the code procedure is the conservation of energy. The sum of all normalized propagating reflected and transmitted diffraction orders must converge to unity with the retention of increasing numbers of spatial harmonics. For 51 harmonics (the number retained in computational analysis) the accuracy of the MATLAB code for conservation of energy was better than one part in 10 billion.

Analysis of the accuracy of the software was tested through the simulation of a binary grating profile with the physical parameters defined in the simulation of Moharam *et al*, whose mathematical analysis of the coupled-wave technique largely laid the foundation for the present work. In their publication, the dimension of variability was the groove depth *d*. The comparable results obtained from our technique appear in Fig. 3 and Fig. 4. The results presented in this paper show very good agreement with their work.

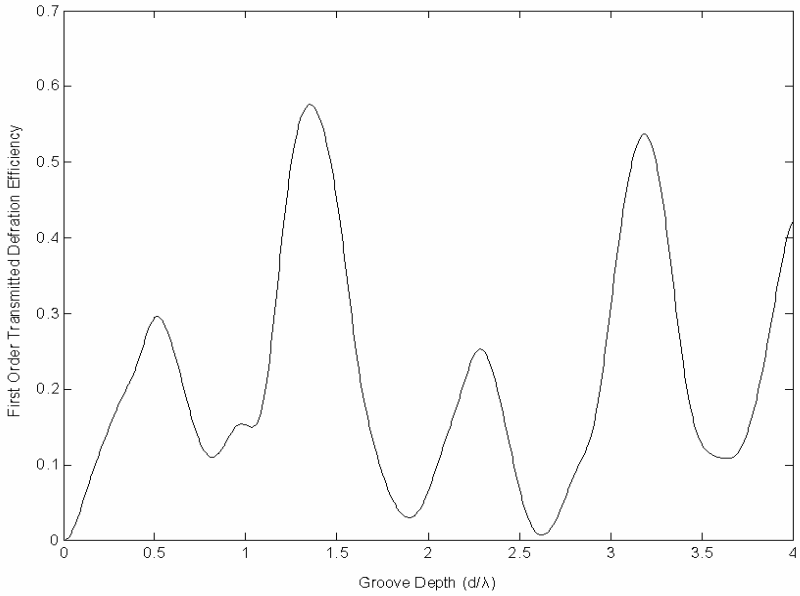


Figure 3. First order TE transmitted diffraction efficiency as a function of grove depth for a theoretical grating with n_I=1, n_{II}=2.02, $\Lambda=\lambda_0$, and f=0.5. The angle of incidence was 10° with λ_0=633 nm.

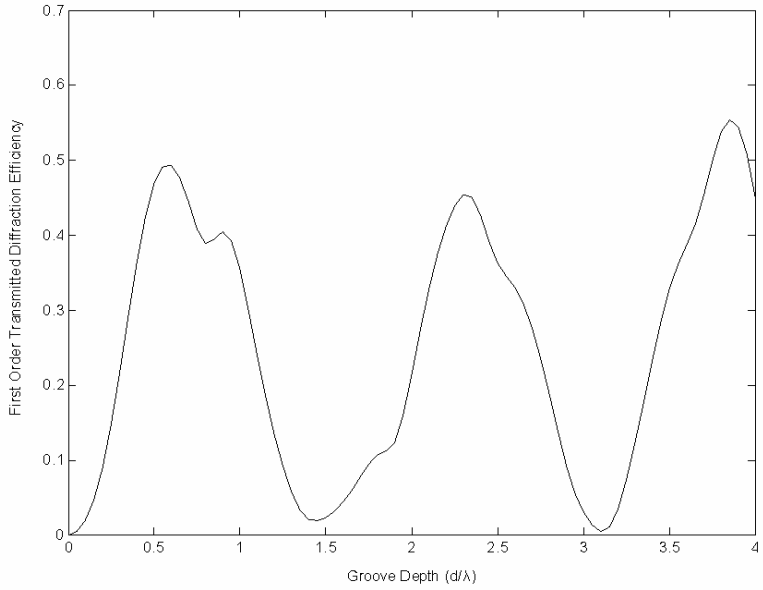


Figure 4. First order TM transmitted diffraction efficiency as a function of grove depth for a theoretical grating with n_I=1, n_{II}=2.02, $\Lambda=\lambda_0$, and f=0.5. The angle of incidence was 10° with λ_0=633 nm.

Although the aim of this work was the implementation of RCWA along multiple wavelengths, no published data for multiple-wavelength binary grating diffraction patterns are known to exist. In order to verify our results for this, a number of experiments were performed with the "white-light laser." Theoretical and experimental results appear in Fig. 5 and Fig. 6. Although the correlation between the two is not strikingly strong, this is most likely do to a number of experimental factors and is not an indication of a failure of this technique. The grating used did not correspond precisely to the parameters set in the computational analysis. The predicted spectral profile corresponds to a binary grating (i.e. a grating with a periodic square shaped profile), although the experimental grating did not possess this ideal characteristic. The degree of discrepancy between the predicted and experimental results was most likely do to the presence of multiple physical parameters which were unaccounted for (e.g. sidewall slope, corner curvature) which may have played a much larger role in the shaping of the diffraction patters the initially assumed. It is possible that the second peak observed at approximately 600 nm was the result of some imprecise etching procedure by which the grove depth was not constant and may have effectively produced a secondary grove depth at which constructive interference could occur for this wavelength.

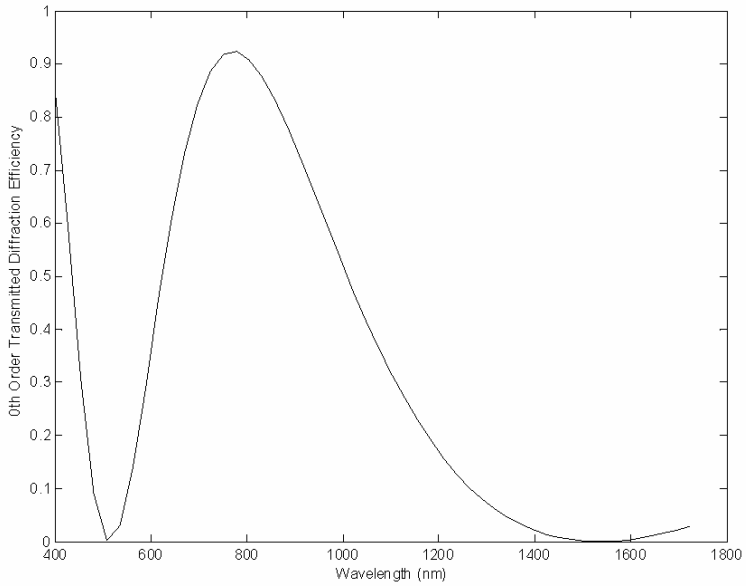


Figure 5: The theoretical results for 0^{th} order transmitted diffraction efficiency with random polarization for a grating with properties n_{II}=1.5, Λ=77 μm, d=3.127 μm, and f=0.5. The angle of incidence was 0°.

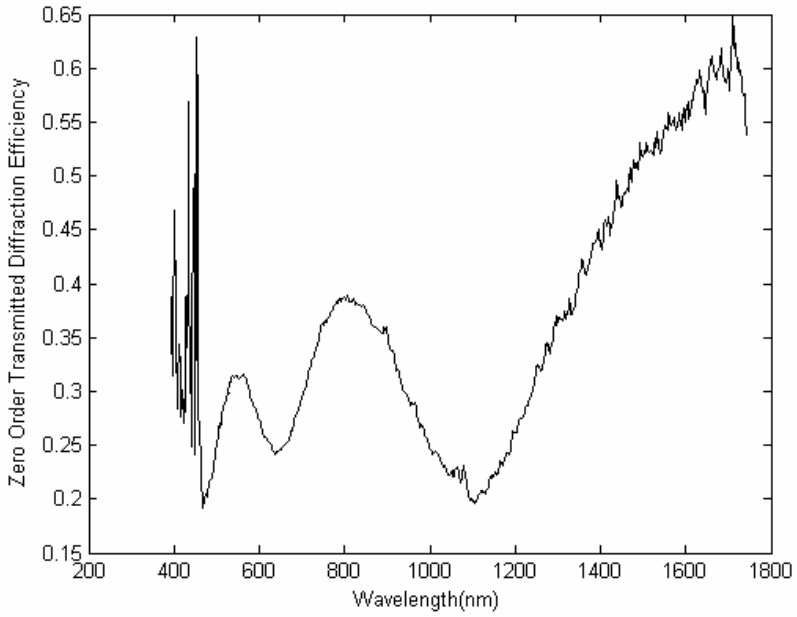


Figure 6: The experimental results for 0^{th} order transmitted diffraction efficiency with random polarization for a grating with properties approximated by n_{II}=1.5, Λ=26.15 μm, d=1.56 μm, and f=0.5. The angle of incidence was 0°.

Despite the imprecise correlation between simulation and observation, it should be noted that the simulation was accurate in predicting a spike in diffraction efficiency at around 800 nm. This agrees with an intuitive examination of the grating structure. At a grove depth of 1.56 μm, two spatially coherent waves striking a ridge and a grove will undergo a relative optical path length difference of $(n_{II}-1)d$. This corresponds to one full wavelength at 780 nm and two full wavelengths at 390 nm, meaning constructive interference will occur at these two wavelengths. This intuitive observation lends support to our method and gives hope that despite the obstacles to verification encountered, finely adapted experiments will yield more precise results.

SUMMARY

Although irrefutable experimental results were not acquired to verify the proper implementation of the program in the case of a continuous spectrum of wavelengths, published results for traditional monochromatic diffraction measurements show strong support for this method. This suggests promise for the technique when it is applied to white light based technology. The hope is that future experiments will verify what was aspired to in this experiment. Further research may yet validate the results obtained presently, showing that discrepancies between theoretical and experimental results were the result of the binary approximations used in modeling the grating structure.

The successful implementation of computational analysis for white-light diffraction based metrology would be a great step towards the eventual implementation of this technology. This method shows great promise in that it may be capable of producing quick and accurate results while reducing the need for dynamic equipment. Additionally, the accuracy and sensitivity of this technique to small variations in feature sizes is not well known and may show to produce a significant improvement over current diffraction based metrology methods; however, more work is necessary in the simulation of this technique. Further research should involve the multi-layer approach to the rigorous coupled-wave analysis of the grating structure. With the successful implementation of this invaluable modeling method, grating structures with complex physical features may be modeled and simulated. This would eventually lead to the better approximation of fabricated grating structures with non-ideal physical properties and the highly accurate modeling of the diffraction patterns of such gratings. If this were accomplished, it would surely lead to the validation of this technology and its expansion into the forefront of diffraction based metrology.

ACKNOWLEDGMENTS

The author wishes to thank Peng Li for his input and assistance in the formulation and compilation of this paper. Also, much gratitude is owed to Zhiwen Liu for his invaluable assistance and financial support.

This work is partially supported by the National Science Foundation under Grant No. EEC-0244030.

REFERENCES

1. H. Hogan. "Tiny Measurements Present Big Challenges." *Photonics Spectra*. Apr. 2004
2. E. M. Drege, J. A. Reed, D. M. Byrne. "Lineraized Inversion of Scatterometric Data to Obtain Surface Profile Information," Society of Photo-Optical Instrumentation Engineers, Vol. 41, No. 1. Pgs. 225-236. Jan. 2002.
3. M. G. Moharam, T. K. Gaylord. "Diffraction Analysis of Dielectric Surface Relief Gratings," *Journal of the Optical Society of America*, Vol. 72, No. 10. Pgs. 1385-1392. Oct. 1982
4. M. G. Moharam, E. B. Grann, D. A. Pommet, T. K. Gaylord. "Formulation for Stable and Eefficient Implementation of the Rigorous Coupled-Wave Analysis of Binary Gratings," *Journal of the Optical Society of America*, Vol. 12, No. 5. Pgs. 1068-1076. May 1995.

NSF EE REU PENN STATE Annual Research Journal Vol. II (2004)

STIMULATED ORIENTATIONAL SCATTERING IN NEMATIC LIQUID CRYSTAL WAVEGUIDE STRUCTURES

Xi Lin*, Andres Diaz[ǂ], Jianwu Ding[+], and I. C. Khoo[#]
Department of Electrical Engineering
Pennsylvania State University, University Park, PA 16802

*Undergraduate student of
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801

ABSTRACT

We have studied the effects of stimulated orientational scattering in a nematic liquid crystal waveguide structure. A linearly polarized 1.55 μm continuous-wave laser is coupled into a liquid crystal waveguide structure through a 10 mm focal lens. In the waveguide, the liquid crystal director is parallel to the polarization of the input laser. Under appropriate experimental conditions, the intensive interaction between the input laser beam and the liquid crystal triggers energy transfer from the input beam to a cross-polarized scattered beam mediated by director axis fluctuations. The complete analysis of the output pump and noise beam vs. the input laser was implemented. Computer simulations of the output beam intensities were made and compared with the experimental results.

[#] Faculty Mentor
[ǂ] Postdoctoral Mentor
[+] Graduate Mentor

INTRODUCTION

Background on Nematic Liquid Crystals

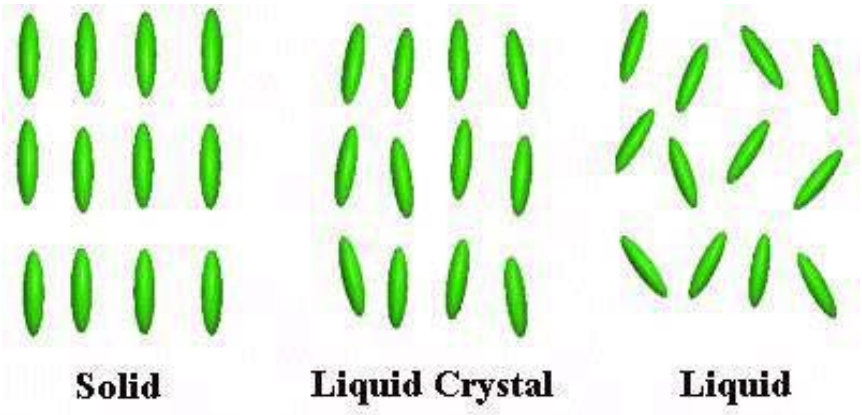


Figure 1. Molecular arrangements of solids, liquid crystals, and liquids.

Liquid crystals exhibit an intermediate phase (Figure 1) between liquids and crystals, hence the name [1]. Their molecular arrangement gives them properties similar to liquids, such as viscosity and surface tension [2]. In addition, liquid crystals have large optical non-linearities and are optically anisotropic materials [3].

Thermotropic liquid crystals are the most widely used and most extensively studied class of liquid crystals. Their various phases are a function of the temperature. The molecules are typically represented as elongated rods arranged in distinctive ordered structures [1, 2]. Thermotropic liquid crystals can be further divided into three classes: nematic, cholestoric, and smetic [1].

Nematic liquid crystals are the most widely studied and widely used class of liquid crystals. They possess the two properties that exemplify liquid crystals: fluidity and crystalline structure [1].

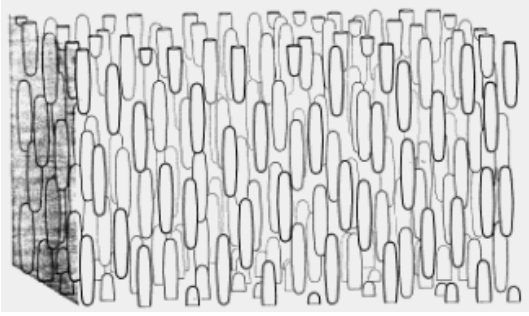


Figure 2. Molecular arrangement of nematic liquid crystals.

The positions of the molecules in nematic liquid crystals (Figure 2) can be seen as random with no ordering of the centers of gravity [2]. The molecules collectively have a directional correlation. They are all oriented in the general direction ($\hat{n}$) of the "director axis" and the direction of each molecule fluctuates about $\hat{n}$ [1, 3].

Nematic liquid crystals are also centrosymmetric, i.e. their macroscopic physical properties remain the same when $\hat{n}$ is inverted (i.e. $+\hat{n} \rightarrow -\hat{n}$). Thus, if each molecule had a dipole moment, the sum of the moments would vanish. The centrosymmetric property of nematic liquid crystals follows from the fact that their second-order nonlinear optical susceptibility $\chi^{(2)} = 0$ [1, 4].

The nonlinear properties of nematic liquid crystals stems from the large magnitude of its third-order nonlinear optical susceptibility $\chi^{(3)}$. This allows the numerous nonlinear effects to be observed with low power, continuous wave (cw) lasers instead of the higher power, pulsed lasers [1, 4]. The effect described in this paper, stimulated orientational scattering, is observed with the use of low power cw infrared lasers.

Nematic liquid crystals also exhibit birefringence, meaning they have different refractive indices for parallel and perpendicular polarizations. This anisotropy in the refractive index is denoted by $\Delta n = n_{\parallel} - n_{\perp}$ [1, 4].

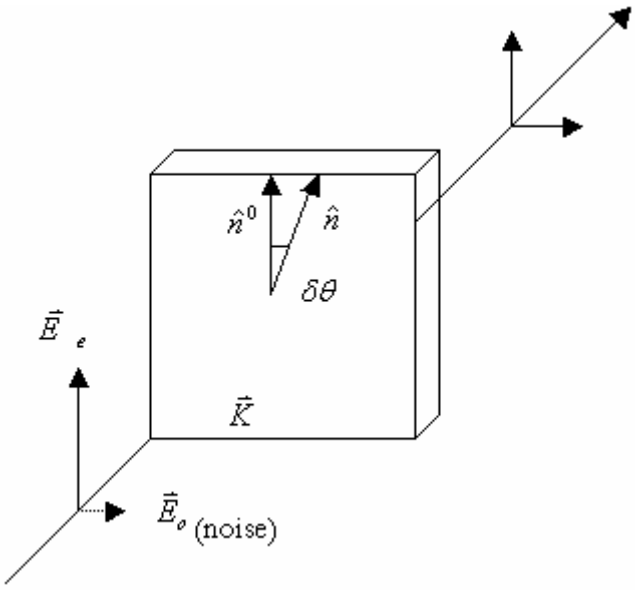


Figure 3. Linearly polarized laser incident on nematic liquid crystal.

Stimulated Orientational Scattering: Theory and Experiments

Stimulated orientational scattering arises due to the nonlinear properties of nematic liquid crystals. An optical plane wave is sent through a nematic liquid crystal (Figure 3) such that its polarization is either parallel (extraordinary, e-ray) or perpendicular (ordinary, o-ray) to the director axis ($\hat{n}$). The plane wave also has a small noise component perpendicular to the original wave. The large

anisotropy and weak intermolecular forces of the nematic liquid crystal cause the director axis to undergo small fluctuations about its director axis.

The first experimental observation of SOS was reported by Zel'dovich, *et al.* A high-powered pulse laser was used to examine the SOS effects in planar nematic liquid crystal. In particular, the transient effects of SOS were examined. Khoo, *et al*, showed that SOS effects, both transient and steady state, can be observed with low power pulsed lasers. P. Etchegoin studied SOS effects, with varying input powers to see the effects on output powers. Recent experiments by Khoo, *et al*., studied the dependence of output power of o-wave on input power of e-wave in nematic liquid crystal film and have shown it to be in agreement with theory [5-7, 9].

All previous studies on SOS used film structures to hold the nematic liquid crystal, whose thickness is on the order of a few hundred microns. No existing studies on SOS have been done using nematic liquid crystal housed in a waveguide structure, which is approximately ten times thicker than film. With increased interaction length, we expect to see higher gain and a lower power threshold for the SOS effects to occur. If SOS can be triggered at very low power, then the applications for SOS are much broader.

EXPERIMENTAL DESCRIPTION

The liquid crystal used is E7 from EM Chemicals (Figure 4) in a planar waveguide at room temperature. E7 is a mixture of four liquid crystals and has properties very different from each of its constituents. It has a nematic-isotropic phase of 63ºC.

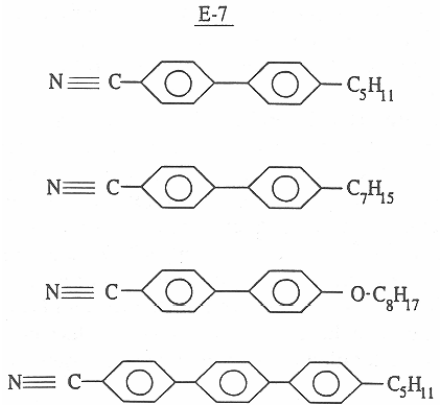


Figure 4. Molecular structure of the constituents of E7.

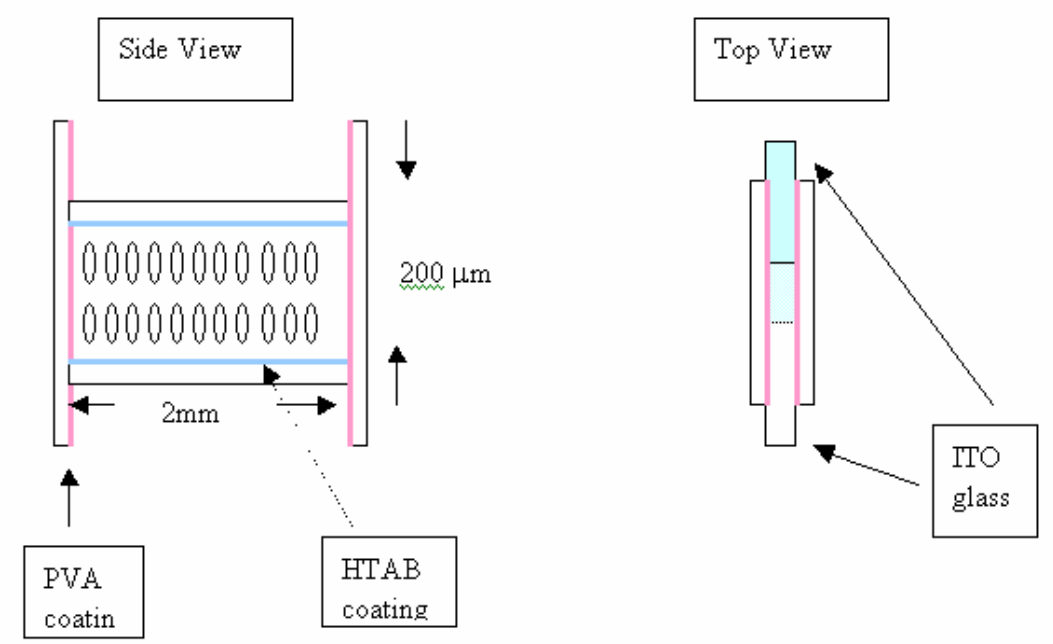


Figure 5. Waveguide design.

The liquid crystal is held in by two pieces of glass slides in the vertical direction and two narrow pieces of ITO glass in the horizontal direction. The vertical surfaces are coated with PVA and the horizontal surfaces with HTAB to ensure correct alignment of the liquid crystal molecules. The waveguide is 200 µm wide and 2 mm deep (Figure 5).

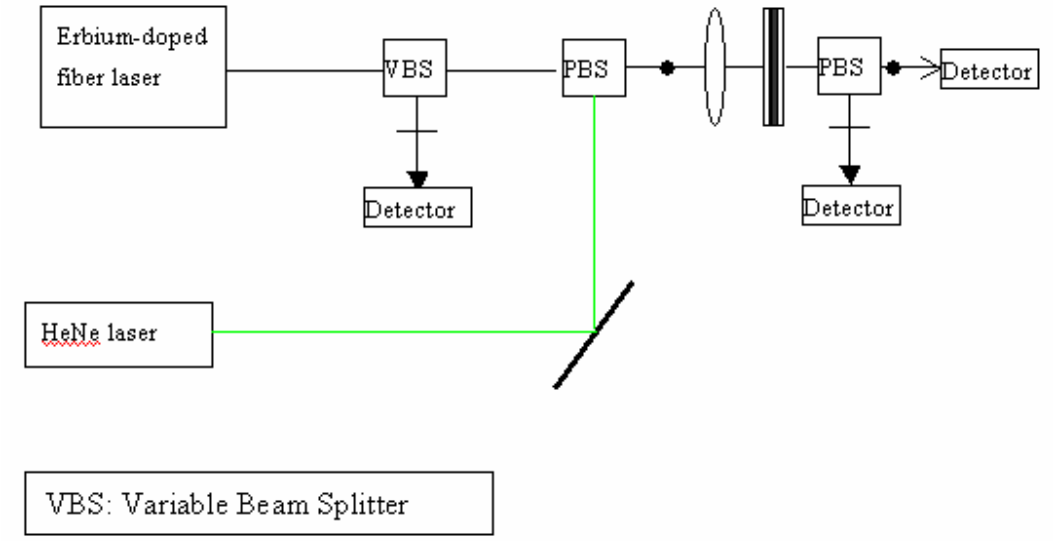
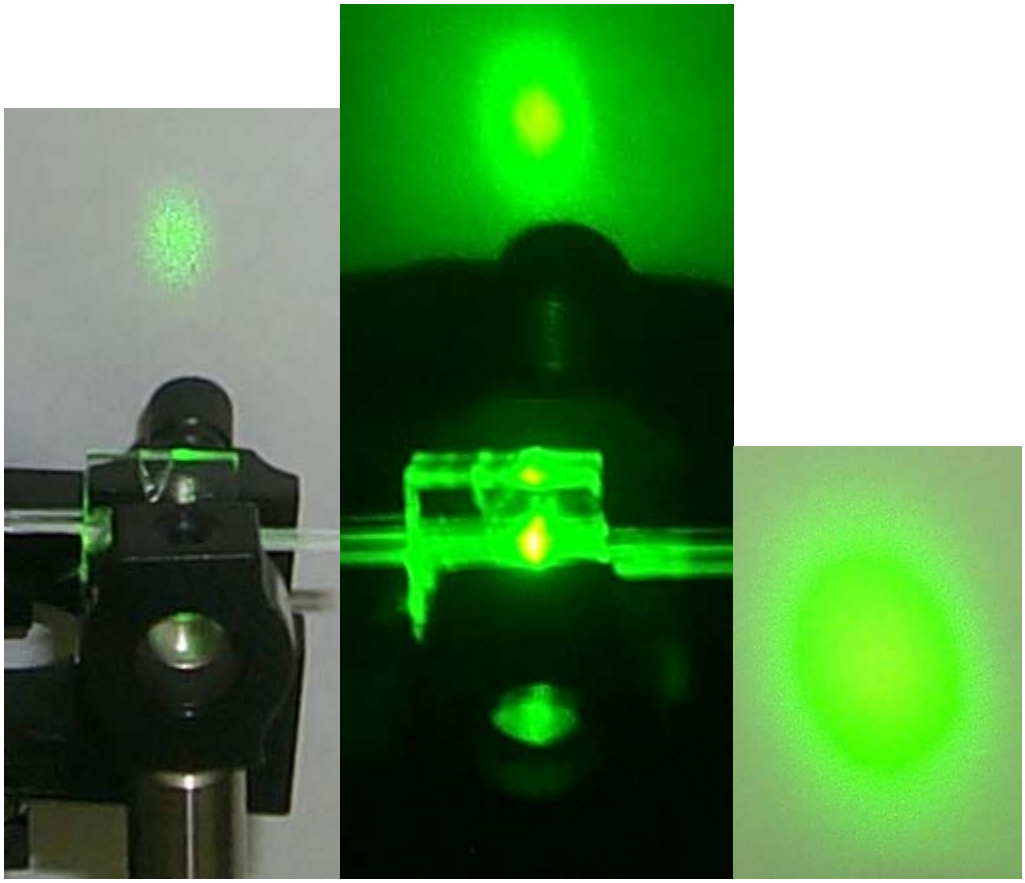


Figure 6. Experimental setup, PBS: Polarizing Beam Splitter.

A linearly polarized low-power cw infrared erbium fiber laser with a wavelength of 1.55 μm is focused by a 10 mm focal length lens onto the nematic liquid crystal waveguide as an e-wave. The focused spot diameter of the laser on the sample is approximately 80 μm.

A variable beam splitter serves to attenuate the power of the laser output so various powers can be achieved. The first polarizing beam splitter polarizes the laser beam to an e-wave that is incident on the waveguide. The second polarizing beam splitter separates the output beam from the waveguide into e-wave and o-wave components.

An infrared detector serves as a reference to the input beam intensity and the other two detectors serve to measure the intensity of the output e-wave and o-wave intensities. The latter two detectors are connected to an oscilloscope to analyze steady-state fluctuations in the intensities.



Figures 7(a) 7(b) 7(c)

Figure 7. Lens focusing laser through waveguide and resulting spot.

A helium-neon laser is used as a reference for proper alignment of the beam path and to check the coupling of the output beams.

RESULTS

Experimental Data

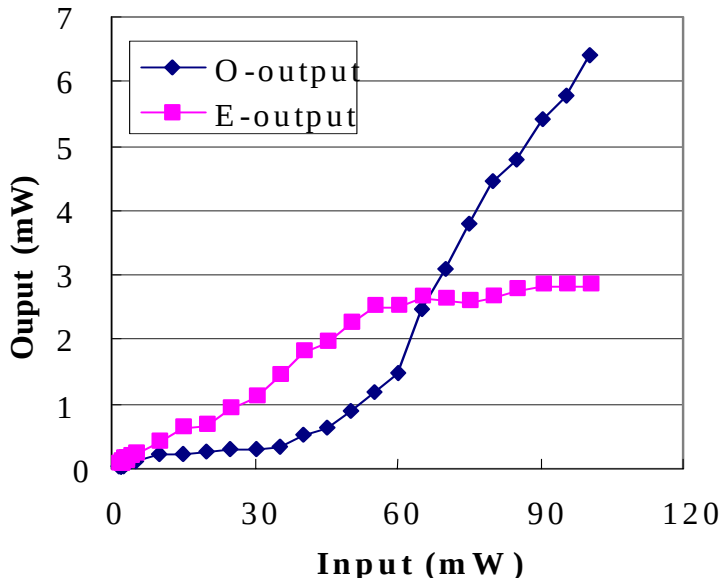


Figure 8 (a) Measured transmitted e and o waves as function of input power.

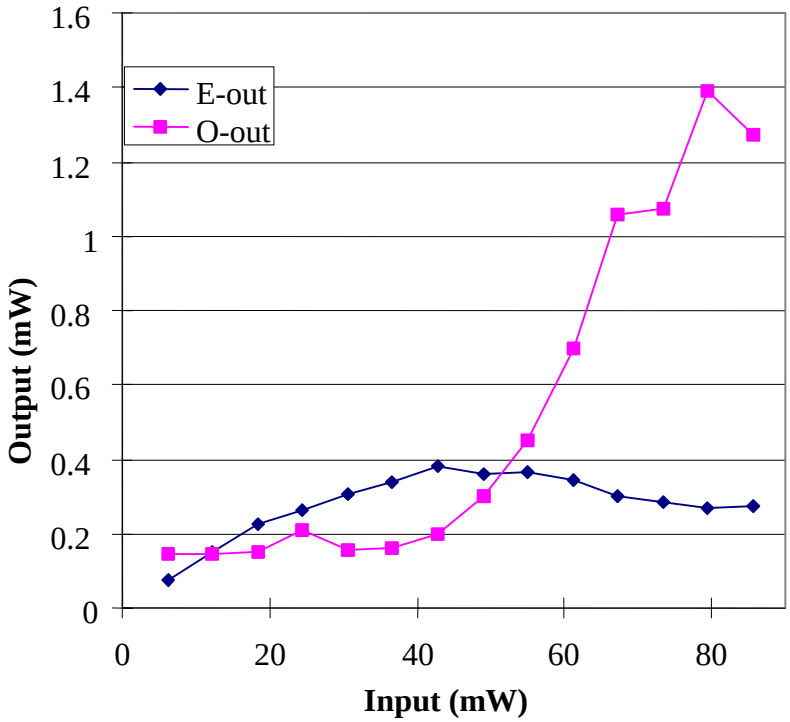


Figure 8 (b) Measured transmitted e and o waves as function of input power.

Computer Simulation

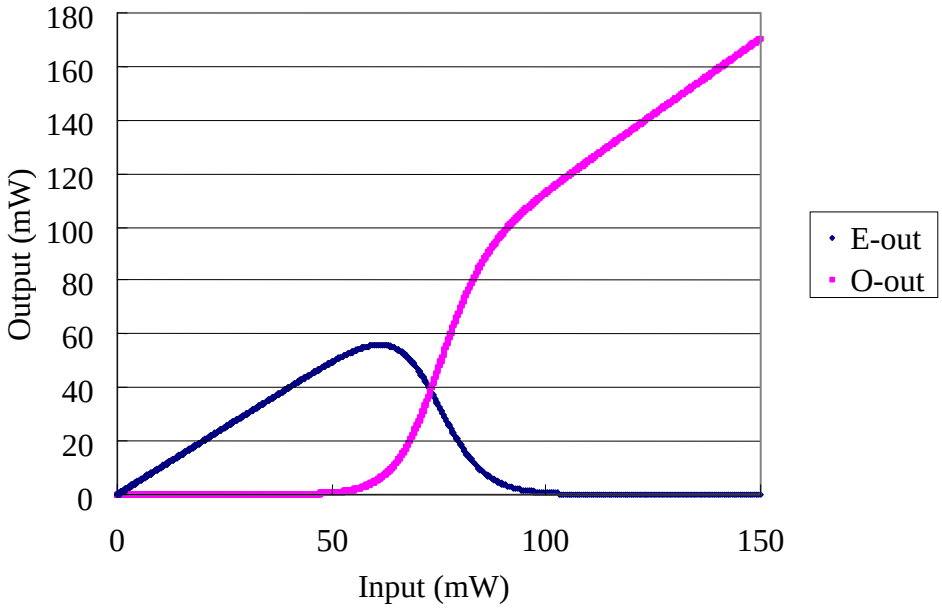
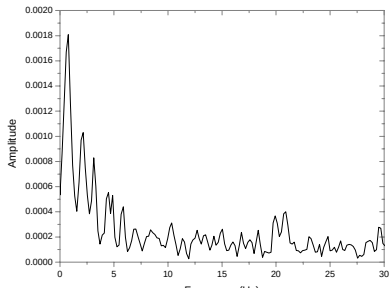


Figure 9. Calculated output e and o waves as function of input power.

The results from two different samples of E7 liquid crystal in waveguide are shown above (Figures 8 (a) and (b)). They agree with the predicted result from theory whereby, initially, the e-wave is linear and the o-wave is constant. After threshold, the o-wave experiences exponential growth and the e-wave is constant. The threshold obtained from the experimental observations agrees with that from the theoretical computation (Figure 9), at about 50 mW.

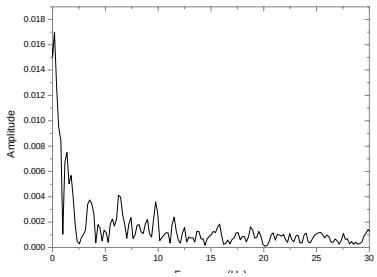
Fast Fourier Transform Analyses

Input Intensity 4.89 mW



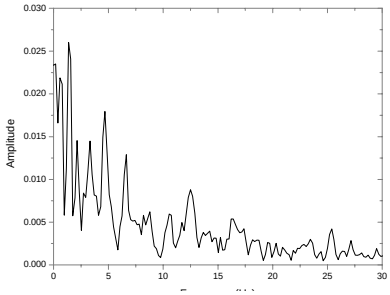
Figures 10 (a)

Input Intensity 15.9 mW



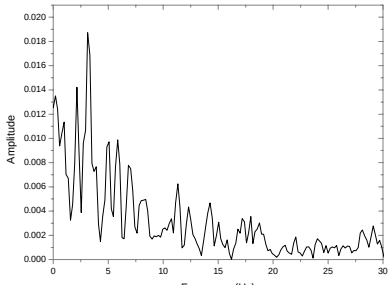
(b)

Input intensity 44.0 mW



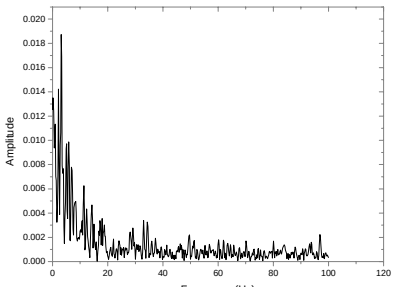
(c)

Input intensity 48.9 mW



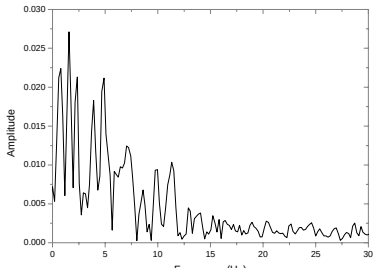
(d)

Input intensity 48.9 mW (different frequency scale)



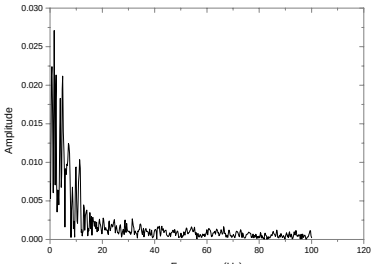
(e)

Input intensity 69.7 mW



(f)

Input intensity 69.7 mW (different frequency scale)



(g)

Figure 10 Fast Fourier Transform for different values of input intensity.

The Fast Fourier Transform (FFT) of 250 ms time-domain o-wave intensities for various input intensities are shown above. Up until approximately 50 mW, the frequency domain has only a DC component (Figures 10 (a) through (c)). For intensities higher than 50 mW, we see a characteristic frequency of approximately 3 Hz along with more rapid oscillations (Figures 10(a) through (g)).

We get infrared transmission ratios of approximately 7% for 2 mm waveguide compared to 85% for 400 μm film [7].

DISCUSSION

The FFT analyses show characteristic frequencies at approximately 3 Hz for input intensities above threshold. This corresponds to the frequency difference, $\Omega = \omega_x - \omega_y$, of the electric field, $\mathbf{E} = E_x e^{i(\mathbf{k}_x \cdot \mathbf{r} - \omega_x t)} \hat{\mathbf{i}} + E_y e^{i(\mathbf{k}_y \cdot \mathbf{r} - \omega_y t)} \hat{\mathbf{j}}$, inside the liquid crystal cell. It is predicted that the frequency difference depends on the particular liquid crystal sample.

Previous experiments involving SOS used liquid crystal housed in film structures. The threshold for those experiments was approximately 150 mW to 200 mW [7, 9]. In our experiments, the threshold for waveguide structures is approximately 50 mW. This agrees with the theory advanced by Khoo, *et al.*, which shows the small signal gain to be exponential with the interaction length of the structure [5].

The output intensities of the experiment and the simulation differ by two orders of magnitude. We believe this is due to the theoretical model not taking into account light absorption in the waveguide structure.

CONCLUSION

In summary, we have shown that, using a waveguide structure instead of film structure, the threshold for the appearance of stimulated orientational scattering effects has been lowered by a factor of two to three, in agreement with theory. This result allows for more applications in which high laser intensities are not possible since the waveguide structure allows the SOS effect to occur at lower intensities.

Possible future experiments in this field include modifying the dimensions of the waveguide to achieve both higher transmission rates and lower threshold level for SOS effects. Further FFT analyses can be done on time-domain intensities, especially at higher intensities where the oscillations are especially rapid.

ACKNOWLEGEMENT

This material is based upon work supported by the National Science Foundation under Grant No. EEC-0244030.

REFERENCES

[1] I. C. Khoo, *Liquid Crystals: Physical Properties and Nonlinear Optical Phenomena*, Wiley, New York, 1995.

[2] Andres Diaz, *Theory and Applications of Molecular and Collective Nonlinear Optical Behavior of Nematic Liquid Crystals*, PhD Thesis, The Pennsylvania State University, 2004.

[3] Jianwu Ding, *Stimulated Orientational Scattering in the Nematic Liquid Crystals*, MS Thesis, The Pennsylvania State University, 2002.

[4] P. Etchegoin and R T. Phillips, "Stimulated orientational scattering and third-order nonlinear optical processes in nematic liquid crystals," *Physical Review E*, **55** (5), 5603-5612, (1997).

[5] I. C. Khoo and Y. Liang, "Stimulated orientational and thermal scatterings and self-starting optical phase conjugation with nematic liquid crystals," *Physical Review E*, **62** (5), 6722-6733, (2000).

[6] I. C. Khoo and A. Diaz, "Nonlinear dynamics in laser polarization conversion by stimulated scattering in nematic liquid crystal films," *Physical Review E*, **68** (042701), 1-4, (2003).

[7] I. C. Khoo and J. Ding, "All-optical cw laser polarization conversion at 1.55 μm by two-beam coupling in nematic liquid crystal films," *Applied Physics Letters*, **81** (14), 2496-2498, (2002).

[8] B. Ya. Zel'dovich, S. K. Merzlikin, N. F. Pilipetskii, A. V. Sukhov, "Observation of stimulated forward orientational light scattering in a planar nematic liquid crystal," *JETP Letters*, **41** (10), 514-517, (1985).

[9] I. C. Khoo, Y. Liang, H. Li, "Observation of stimulated orientational scattering and cross-polarized self-starting phase conjugation in a nematic liquid-crystal film," *Optics Letters*, **20** (2), 130-132, (1995).

INTENSITY PROFILES OF TRANSMITTED LIGHT IN FERROELECTRIC SINGLE CRYSTAL FIBERS

Pape A. Sene*, Hongbo Liu[+], and Ruyan Guo[#]

Department of Electrical Engineering and Materials Research Institute
The Pennsylvania State University, University Park 16802

*Undergraduate Student of
The Department of Electrical and Computer Engineering
Polytechnic University
Brooklyn, NY 11201

ABSTRACT

Ferroelectric single crystal fibers have naturally waveguide structure in addition to having large electrooptic and nonlinear optic effects, which make them ideally suited for advanced optical and electrooptical applications needed in various fields such as information processing and optical communications. Ferroelectric crystal fibers have indices of refraction sensitive to the polarity and intensity of laser light thus having important applications in nonlinear optics including frequency controls and miniature lasers when doped properly. This research focuses on experimental studies of the light intensity profiles transmitted through single crystal $LiNbO_3$ fibers. Theoretical simulations of such transmitted intensity profiles of laser light are also carried out for the $LiNbO_3$ single crystal fibers using beam propagation method. The simulation and experimental studies are extended to the commercially available single mode and multimode silica optical fibers to compare the results obtained for $LiNbO_3$ crystals. Erbium doped $LiNbO_3$ single crystal fibers were also tested for their gain characteristics and their light intensity profiles to identify their potential in 1550 nm lasing and optical amplifications.

INTRODUCTION

Ferroelectric materials in their ferroelectric phase have spontaneous polarizations,[1] which is the result of cooperative alignment of permanent dipoles

---

[#] Faculty Mentor
[+] Graduate Mentor

in these materials. Even without the application of an electric field, a properly poled ferroelectric material or a ferroelectric single crystal (with single domain) will have net surface charges (which is the reason why the materials have remanent polarization) and have anisotropic electrical, mechanical, and optical properties. Ferroelectric materials have very important applications in micromechanical systems (MEMS), ultrasonic motors, pyroelectric detectors, capacitors, power conversion devices, electromechanical-actuators, and various sensors [1].

Lithium niobate crystal is one of the well-known ferroelectric crystals in optical and electrooptic applications. It has a wide range of transparency from 450 nm to 5.2 μm [7]. It has also a very good optical damage resistance when operated above 800 nm. Although its electrooptic coefficient is not one of the highest (31 pm/V) its temperature stability is excellent. $LiNbO_3$ thus has been widely used as electro-optic modulator and Q-switch for Nd:YAG or Ti:Sapphire lasers as well as modulator for fiber optics communications[1]. With specially designed domain patterns, $LiNbO_3$ crystals have also been used for second harmonic generating (SHG) devices that generate light in the visible wavelength range from a diode pump light of infrared wavelengths [1].

$LiNbO_3$ has a ferroelectric phase transition temperature (Curie temperature) of 1170^{o}C that is lower than its melting point of 1250^{o}C [7]. Large crystals up to 2 inches in diameter can be grown commercially. However, a-axis oriented crystals of self-cladded waveguide geometry is not commercially available and the technique producing planar waveguide structure from a bulk crystal by, e.g., ion beam implantation, is both costly and inefficient in light coupling. Using an advanced growth technique known as LHPG (laser heated pedestal growth) method [e.g., 2,3] available at the Materials Research Institute, Penn State; single crystal lithium niobate fibers can be grown with desired orientation and dimension, for potential use in various optical applications. In addition, the LHPG technique allows the growth of doped crystals with excellent uniformity in both the radial and the length directions, which ensure long interaction distance with large light intensity, both are critical factors for nonlinear optical applications.

In this research it is investigated how the laser light propagates inside an anisotropic ferroelectric single crystal fiber by simulation and experimental studies. The modal structures for the light propagating in single mode glass fibers (SMF) and multimode glass fibers (MMF) were also investigated, to make a comparison to the light propagation in a lithium niobate single crystal. In addition, preliminary result of emission and gain characteristics of erbium-doped $LiNbO_3$ crystal fibers was obtained. The successful completion of this study will help establish a research platform for fiber crystal device design. Coupled with the unique expertise in materials science and single crystal fiber growth at Penn State, it would accelerate investigation on the modal structures of crystals with various shapes, orientations, and domain configurations, and that of temperature and electric field dependence, which will lead to synthesis and design of

miniature monolithic SHG lasers and for wavelength sensitive optical couplers and switches.

THEORETICAL BACKGROUND

(A) Modal Structure

Light beams are electromagnetic waves that are made up of electric and magnetic fields; both fields oscillate together at the frequency of light[5]. The propagation of laser light, through a polarizable medium is governed by the following set of equations, known as the Maxwell's equations in differential form[5]:

$$\begin{cases} \nabla \times \vec{E} = -\dfrac{\partial \vec{B}}{\partial t} & \text{(Faraday's Law)} \\ \nabla \times \vec{B} = \mu_0 \mu_r \varepsilon_0 \varepsilon_r \dfrac{\partial \vec{E}}{\partial t} & \text{(Maxwell modified Ampere's Law)} \\ \nabla \cdot \vec{\mathrm{D}} = \nabla \cdot \vec{E} = \rho & \text{(Gauss's Law)} \\ \nabla \cdot \vec{\mathrm{B}} = 0 \end{cases} \tag{1}$$

When we deal with linear, isotropic, and ideal dielectric materials (no conduction current **J**=0 and no free charges ρ=0), these equations take the form [5]:

$$\begin{cases} \nabla \times \vec{E} = -\dfrac{\partial \vec{B}}{\partial t} \\ \nabla \times \vec{B} = \mu_0 \mu_r \varepsilon_0 \varepsilon_r \dfrac{\partial \vec{E}}{\partial t} \\ \nabla \cdot \vec{\mathrm{D}} = \nabla \cdot \vec{E} = 0 \\ \nabla \cdot \vec{\mathrm{B}} = 0 \end{cases} \tag{2}$$

where $\vec{D} = \varepsilon \vec{E} = \varepsilon_0 \varepsilon_r \vec{E} = \varepsilon_0 \vec{E} + \vec{P}$ and $\vec{B} = \mu \vec{H} = \mu_0 \mu_r \vec{H}$. Dielectric permittivity $\varepsilon = \varepsilon_\circ(1+\chi)$, where χ is the susceptibility of the material. The average electric dipole moment per unit volume or the dielectric polarization of the material [5], $\boldsymbol{P} = \varepsilon_\circ \chi \boldsymbol{E}$.

The general form of a wave equation is

$$\nabla^2 f = \frac{1}{v^2} \frac{\partial^2 f}{\partial t^2} \tag{3}$$

where v is the phase velocity of the wave. For free space: $c = \dfrac{1}{\sqrt{\varepsilon_0 \mu_0}} = 3\text{x}10^8$ [m/s]. The speed of light in the dielectric material is defined by $\nu = \dfrac{1}{\sqrt{\mu\varepsilon}} = \dfrac{c}{n}$,

where the refractive index of the material is defined [5] by $n = \sqrt{\frac{\varepsilon}{\varepsilon_o}} = \sqrt{(1+\chi)}$ or $n^2=\mu_r\varepsilon_r$.

For electromagnetic fields $\mathbf{E}(\mathbf{r},t)=Ee^{\boldsymbol{j}(\omega t-\mathbf{k}.\mathbf{r})}$ and $\mathbf{H}(\mathbf{r},t)=He^{\boldsymbol{j}(\omega t-\mathbf{k}.\mathbf{r})}$, where **k** is the wave vector which contains three components $\mathbf{k}=\boldsymbol{i}k_x+\boldsymbol{j}k_y+\boldsymbol{k}k_z$ and the k_x, k_y and k_z are wave numbers along x, y, and z-direction respectively, using (Eq.3), one obtains:

$$\nabla^2E= -\mu\varepsilon\omega^2E \qquad (4)$$
$$\nabla^2H= -\mu\varepsilon\omega^2H \qquad (5)$$

In free space $\omega^2=(2\pi/\lambda_0)^2(c)^2$, let $k_0=(2\pi/\lambda_0)$ as the wave number in free space, the wave equations are of the form known as *Helmholtz Wave Equations*:

$$\nabla^2E + n^2k_0^2E =0 \qquad (6)$$
$$\nabla^2H + n^2k_0^2H =0 \qquad (7)$$

Simulating transmitted light intensity profile is a rather complex process, which involves finding solutions for Helmholtz equations that satisfy both E and H field boundary conditions in a specific fiber. For glass fibers of cylindrical core/cladding boundary condition the solutions may be obtained by the method of variable separation (solve the radial, angular and longitudinal components individually) to find the guided propagating light field strength. For the light field component along the propagation direction $E_z(r,\phi,z) = R(r)\cdot\Phi(\phi)\cdot Z(z)$, the solution is of the form:

$$E_z(r,\phi,z,t) = R(r)\cdot\Phi(\phi)\cdot Z(z)\cdot e^{i\varpi t} = \begin{cases} AJ_m(\kappa\cdot r)e^{im\phi}e^{i\beta z}e^{i\varpi t} & \text{for } r \le a \\ CK_m(\gamma\cdot r)e^{im\phi}e^{i\beta z}e^{i\varpi t} & \text{for } r > a \end{cases} \qquad (8)$$

where A and C are constants, $(J_m)$and (K_m) are the Bessel functions, κ and γ are known as phase and decay constant, respectively, and *m* and β are the propagation constants. The solutions of Helmholtz Wave Equations are not continuous, because $\Phi(\phi)=\Phi(\phi+2\pi)$ for m=0, ±1, ±2, … so that $\Phi(\phi)$ and that of $E_z(r,\phi,z)$ are periodic with discrete solutions. Each of the mathematically discrete solution of the wave equation corresponds to a propagating guided mode in the optical fiber.

By repeating the process for E_z one can also find E_x:

$$E_x(r,\phi,z,t) = E_0 \begin{cases} \dfrac{1}{J_0(\kappa a)} J_0(\kappa \cdot r) \cdot e^{im\phi} \cdot e^{i\beta z} \cdot \mathrm{e}^{\mathrm{i}\varpi \mathrm{t}} & \text{for } r \le a \\ \dfrac{1}{K_0(\gamma a)} K_0(\gamma \cdot r) \cdot e^{im\phi} \cdot e^{i\beta z} \cdot \mathrm{e}^{\mathrm{i}\varpi \mathrm{t}} & \text{for } r > a \end{cases} \quad (9)$$

Knowing the electric field distribution in cross section of a fiber after a particular propagation distance along z-direction, the total power and power distribution can be obtained, as power density is proportional to $|E_x|^2$, and the power in an area can be obtained through integration.

Physically the mode concept is not complicated. Light will not be propagating inside a fiber if it does not satisfy the total internal reflection condition that is governed by Snell's law [4]. Total internal reflected light would not be continually guided inside the waveguide except those lights that are in-phase with each previous reflection. Each such guided propagating wave at a specific total internal reflection angle is regarded as a propagation mode [4].

(B) Beam Propagation Simulating Method

An advanced simulation program (BBV Slab Software) is used in this work, which is appropriate for calculating the two dimensional beam propagation of laser light into a material. It is capable of calculations for both TE and TM polarizations. To execute the simulation, information about the material would be entered. The structure of the material gets defined in the Structure box, that includes the definition of the structure of the material, its shape and the measures of its width, height, length, cladding size if it has one, and core size. The properties of the material such as its index of refraction are entered in the Material and Variable boxes. The wavelength of the laser light is also entered. This software is used a lot in the field of integrated optics. It takes about three hours, depending the resolution desired, for the software to calculate in one cross section what the output of the light coming out of the crystal should look like. There are different methods of calculating the output of the modes, but the software promises better results if the automatic calculation is chosen leaving it the choice to use the best method for the given material characteristics.

EXPERIMENTAL DESCRIPTIONS

(A). Transmitted Light Intensity in SMF and MMF:

A commercially available single mode glass fiber (9 μm diameter core and 125 μm of diameter cladding) and a commercially available multimode fiber (50 μm core and 125 μm cladding) were used for the direct observation of the transmitted light intensity profiles. The fibers have index of refraction of the core n_1=1.48, and index of refraction of cladding n_2=1.46, and their structures are shown in Fig. 1.

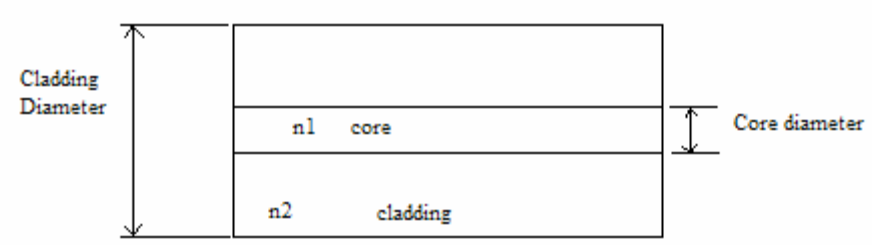


Fig. 1. Illustration of the side cross-section of an optical glass fiber with core diameter 2*a*.

The experiment was carried out in such a way (Fig. 2) that one end of the glass fiber lined up next to the laser source, and the other end of the fiber facing a focus lens, which was directed to the microscope that brings the CCD camera into focus, to capture the whole laser light coming out of the crystal. The He-Ne 633 nm laser light was used in this case.

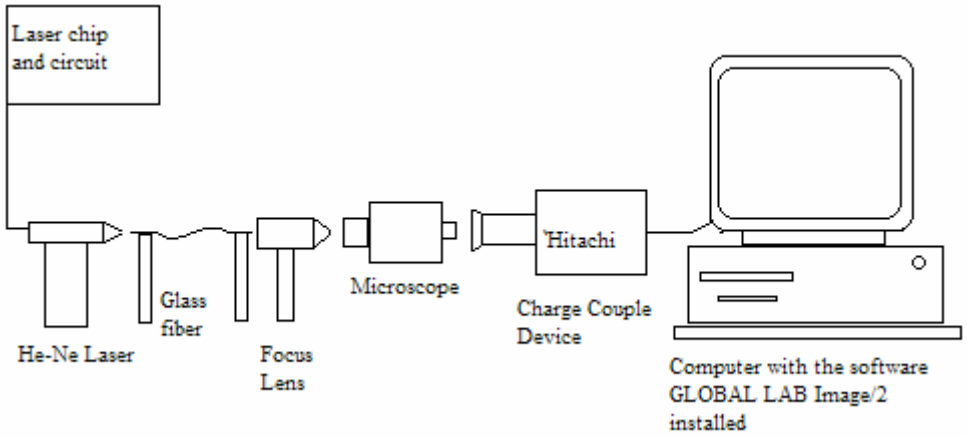


Fig. 2. Experimental configuration of intensity profile observation and capture.

(B). Transmitted Light Intensity in Single Crystal $LiNbO_3$ Fiber

Almost the same experimental set up (Fig. 2) was used to observe the output modes structure of the 633 nm laser light through the lithium niobate crystal. The $LiNbO_3$ crystal used was grown by the LHPG technique with high optical quality. The crystal has a length direction along its crystallographic a-axis and one of the transverse directions parallel to the polar c-axis. A typical (~130um) view of such crystal fibers [6] is shown in Fig. 3 where the shorter direction in the cross section coincides with the c-axis.

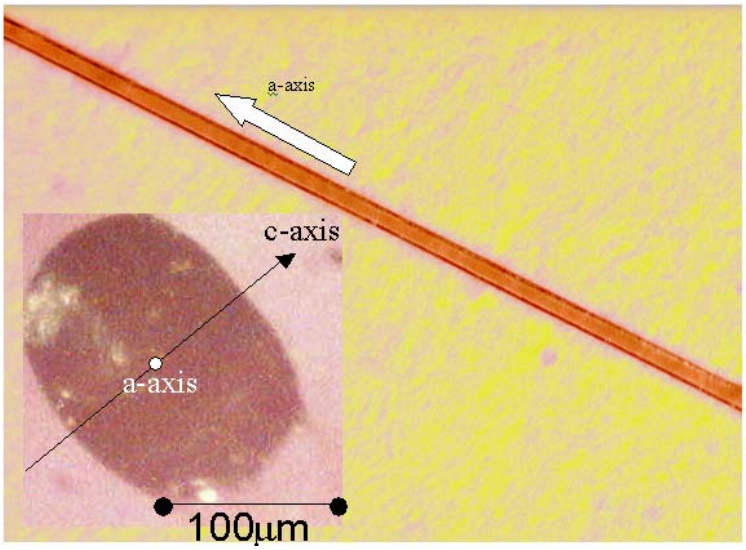


Fig. 3. Single crystal $LiNbO_3$ a-axis fiber grown by the laser heated pedestal growth method. The inset shows the cross section of such a fiber in an enlarged scale.

(C). Observation and Analysis of Emitted Light Intensity in Doped Single Crystal $LiNbO_3$ Fiber

An LHPG grown Erbium doped $LiNbO_3$ crystal fiber was used in this experiment to test the effectiveness of the doping and to study the gain characteristic of the Er-doped fiber crystal. The experiments were carried out in the microwave and photonic laboratory Electrical Engineering Department, Lehigh University, through collaboration.

For testing the optical pumped light emission, an 808 nm diode fiber pump laser was used. A modular controller provided both temperature stability for pump, the sample, and the continued adjustment of pumping light intensity. With careful alignment ensuring effective coupling into the crystal fiber, both pump light and emitted light are observed using a highly sensitive CCD camera (Cohu Model 7512) and the beam profile was analyzed by Spiricon Laser beam diagnosis software. The actual spectrum content of the emitted light was analyzed using optical spectrum analyzer. Neutral density filters and/or bandpass (1550nm) filters are used in front of the CCD camera to avoid saturating the sensor. The experimental configuration is shown in Fig. 4 and a schematic diagram of the experiment is presented in Fig. 5.

For the spectrum analysis, a pigtailed 1550 nm optical fiber probe connected to the OSA is aligned directly at the exit end of the $LiNbO_3$ fiber. The similar alignment was also used to measure the spectrum content of the 808 nm diode laser.

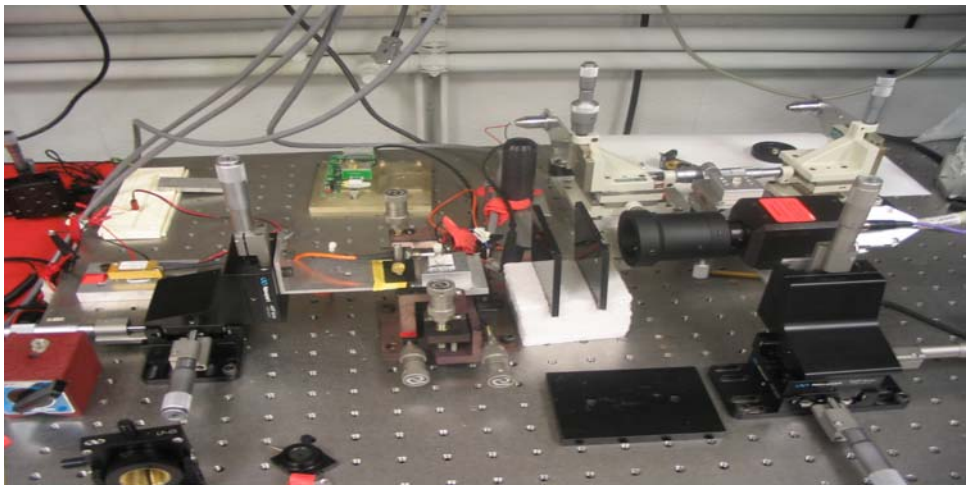


Fig. 4. Actual experimental configuration for laser pumped emission test for Er:$LiNbO_3$ crystal fiber.

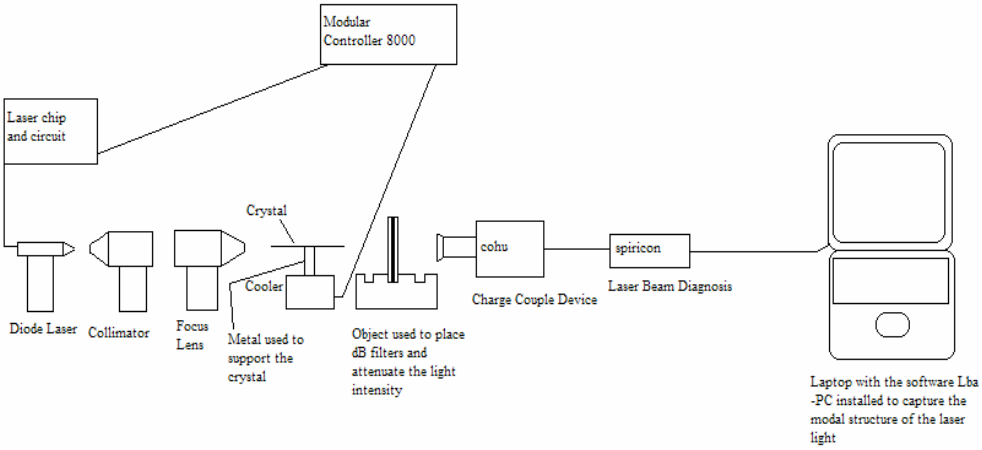


Fig. 5. Schematic diagram of the experiment for laser pumped emission test for Er:$LiNbO_3$ crystal fiber.

RESULTS

(A). Observation of Transmitted Light Intensity in SMF and MMF:

Simulated result calculated using BBV software, (Fig. 6a), and the captured experimental result (Fig. 6b) for a single mode optical silica fiber are presented in Fig. 6. Corresponding results for multimode optical silica fiber are obtained and shown in Fig. 7.

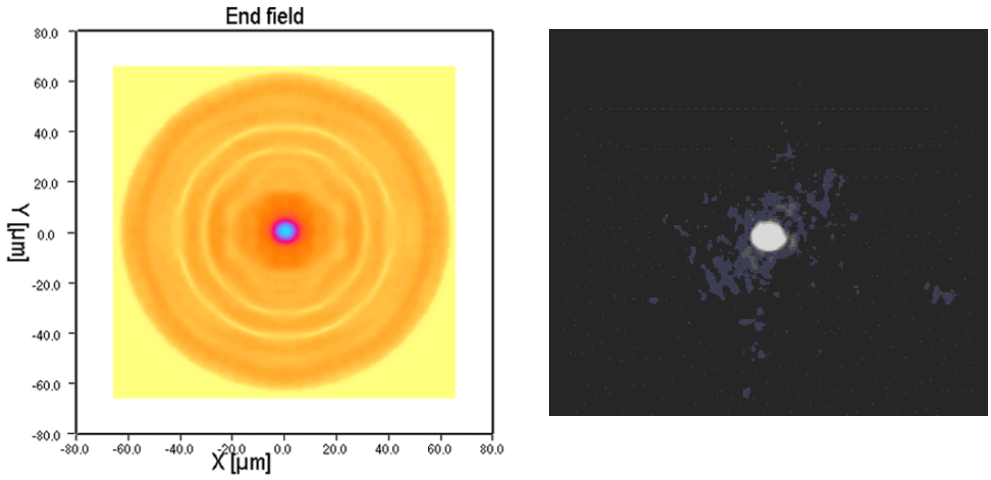


Fig. 6. (a, left) Simulated light intensity profile and (b, right) the measured light intensity profile at the cross section of the exit end for a single mode glass fiber.

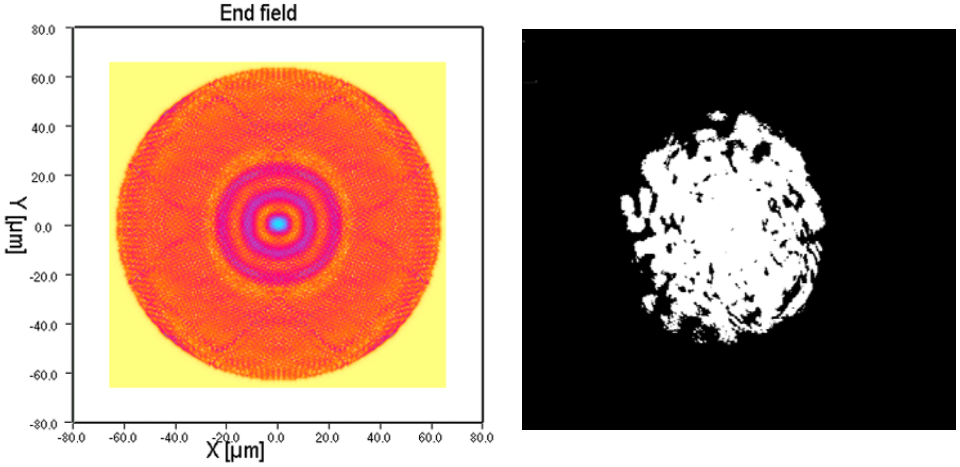
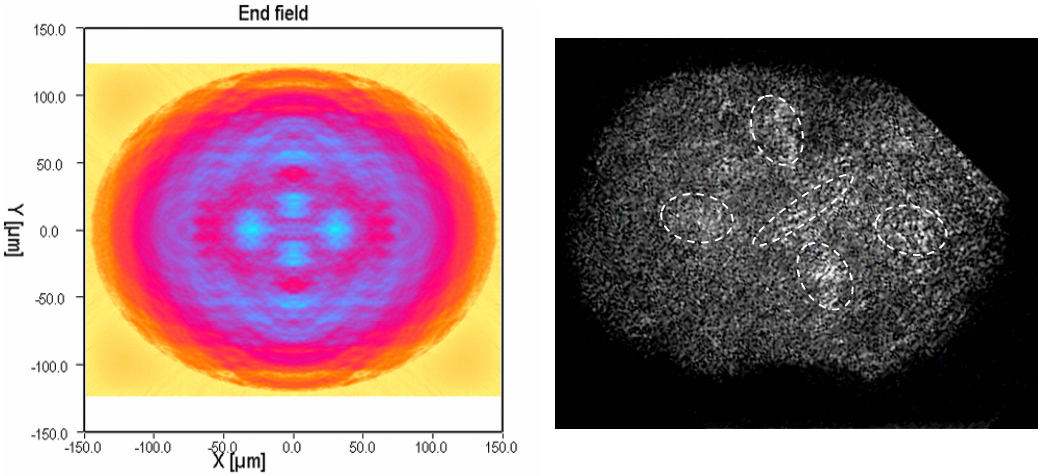


Fig. 7. (a, left) Simulated light intensity profile and (b, right) the measured light intensity profile at the cross section of the exit end for a multimode glass fiber.

(B). Observation of Transmitted Light Intensity in Single Crystal $LiNbO_3$ Fiber

The simulation for single crystal $LiNbO_3$ fiber is much more complicated as the crystal does not have circular cross section and its indices are anisotropic. Simulated result obtained using the BBV software (Fig. 8a) and the captured experimental result (Fig. 8b) for a single crystal $LiNbO_3$ fiber are presented in Fig. 8. The dashed elliptical enclosures were added in Fig. 8(b) to aid the illustration of the intensity distribution observed in the crystal fiber.

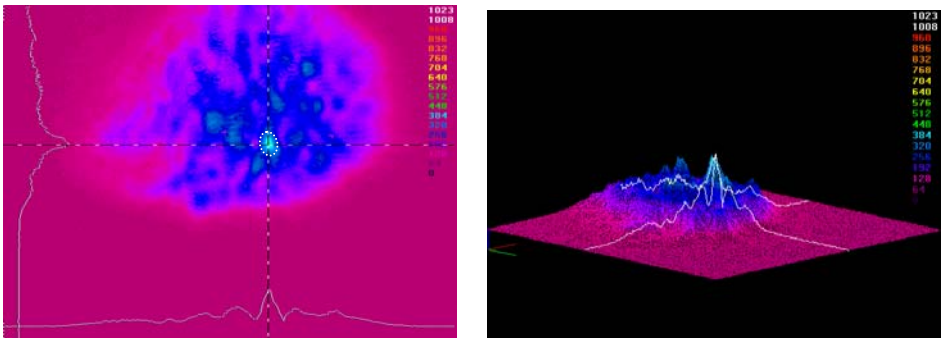


(a) (b)

Fig. 8. (a) Simulated light intensity profile and (b) the measured light intensity profile at the cross section of the exit end for a single crystal $LiNbO_3$ fiber. The cross section is the (100) plane.

(C). Observation and Analysis of Emitted Light Intensity in Doped Single Crystal $LiNbO_3$ Fiber

Figure 9 demonstrated the light profile observed using high sensitivity CCD camera at the exit end of the Er-doped $LiNbO_3$ crystal in both 2D and 3D views. The light intensity profiles shown were obtained at a pumping condition V_b=1.843 V and I_b=594.8 mA for the pump light (808 nm). Although appropriate filters were used to attenuate the laser and avoid damaging the CCD camera, the pump light in Fig. 9 was not filtered; a dotted ellipsoid was added in Fig. 9(a) to indicate the location of the crystal fiber. At sufficient power level of the pump laser, it is evident that the light through the $LiNbO_3$ crystal is significantly strong with intensities well above the background pump light.



(a) (b)

Fig. 9. Light intensity observed at the exit end of the Er-doped $LiNbO_3$ crystal (a) in a 2D view and (b) in a 3D view. The dotted elliposoid encompass the crystal. The background light field is pump light (30 dB neutral density filter applied).

The spectrum content of the emitted or transmitted light from the Er-doped $LiNbO_3$ crystal is analyzed using optical spectrum analyzer (OSA). Fig. 10 demonstrates the frequency contents of the pump light used that is centered around 808 nm. Light emitted from Er-doped $LiNbO_3$ fiber contains emitted light with wavelength in the rather broad spectrum around 1550±20 nm, as is evident as shown in Fig. 11. When continued light capture is made, the light emitted around 1550 nm is further verified, as shown in Fig. 12, where the upper trace represents the maximum-recorded values, and the lower trace represents the instantaneous values.

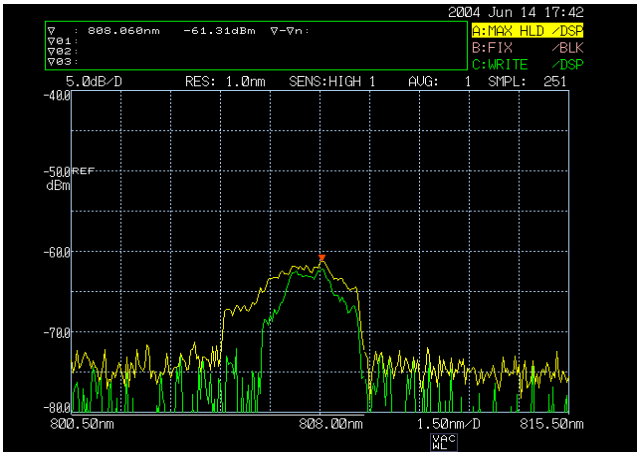


Fig. 10. The spectrum content obtained for the diode pump light.

DISCUSSIONS

The modal structure of the 633 nm laser light coming out of the single mode fiber was fairly predictable and did confirm the theory of light propagation in single mode fibers. Examining and comparing the figures of simulated and experimentally measured intensity profiles, it is apparent that different appearance between simulated results and the corresponding experimental results do exist. The different appearance comes from several reasons, imperfection, parameters offset, and the most importantly the light interference of the multiple modes. When there is only one propagating mode, although interference does present caused by the dust particles in the air with the coherent light and gives spotty look of the intensity profile, the basic profile does not change. For MMF and single crystal cases, light from different modes travels along slightly different paths with different phases when emerges out of the fiber therefore their interference pattern could be very complicated. The intensity profile one captures thus is actually the specklegram contains multiple orders of interference.

Taking the interference into consideration, the theory about light propagation inside a multimode fiber is somewhat verified. The resemblance of the intensity distribution found for $LiNbO_3$ fiber with the theoretical simulation results is rather remarkable. Within the crystal fiber, total light power is carried by individual modes so that, at the fiber's output, these small portions combine, producing an output beam with its power.[4]

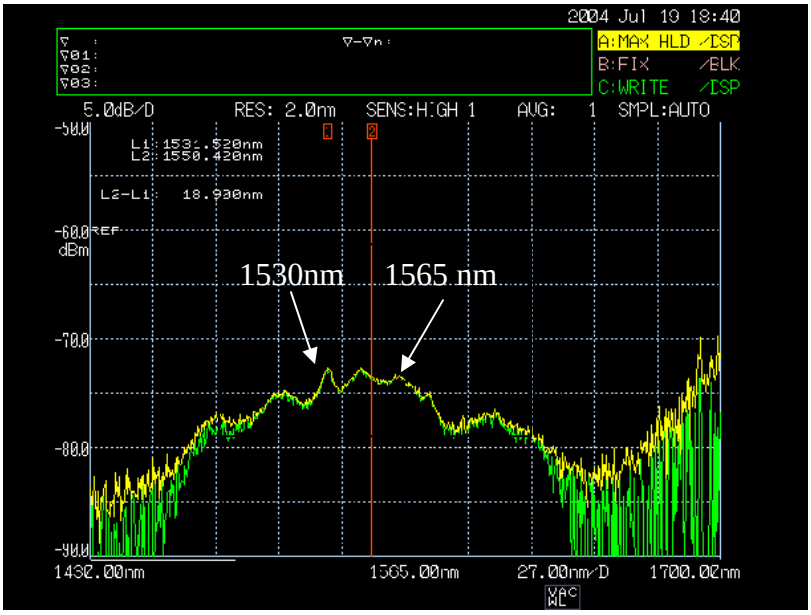


Fig. 11. Spectrum content of the light emitted from Er-doped $LiNbO_3$ crystal fiber.

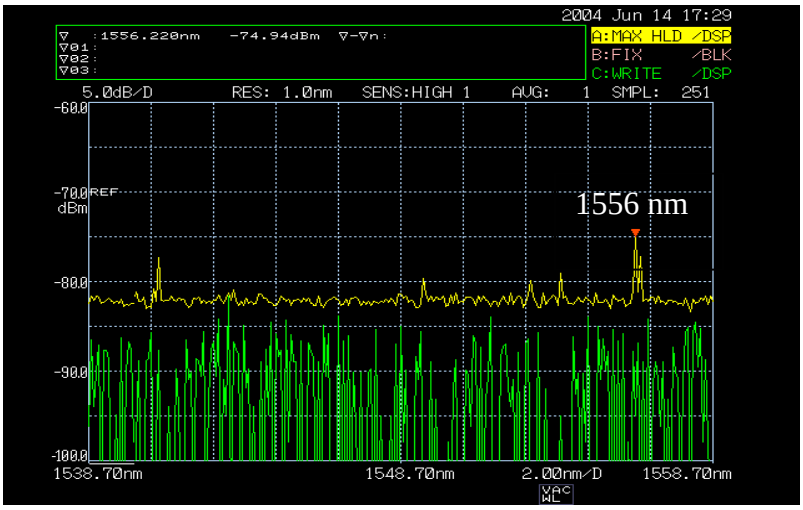


Fig. 12. Accumulated light intensity (upper trace) emitted from the Er-doped $LiNbO_3$ single crystal fiber.

Although lack of further experimental detail to warrant quantitative interpretation for the ferroelectric crystal case, its mode structure with off-center maxima presents very interesting pattern and requires consideration for potential applications. Further study on their polarization and electric field dependence is

to be carried out. More quantitative simulation will also need to be carried out, which takes into consideration of the refractive indices as function of wavelength, temperature, biasing stress or electric field, and light field intensity.

It is well known that the lithium niobate crystal has an ordinary and extraordinary indices of refraction that depend on the wavelength λ (in microns) of the laser light going through the crystal. The indices can be calculated using the Sellmeier equations.[7] At room temperature, 24.5°C, the equations [7] for the congruent Lithium Niobate are represented as follow:

$$n^2_o(\lambda) = 4.9048 + \frac{0.11768}{(\lambda^2 - 0.04750)} - (0.027169 \cdot \lambda^2) \quad (10)$$

$$n^2_e(\lambda) = 4.5820 + \frac{0.099169}{(\lambda^2 - 0.04443)} - (0.021950 \cdot \lambda^2) \quad (11)$$

Therefore, for the congruent lithium niobate crystal, the calculated refractive indices for each laser used are:

632 nm laser;	$n_o = 2.28656$;	$n_e = 2.20286$
808 nm laser:	$n_o = 2.25421$;	$n_e = 2.17501$
1550 nm laser:	$n_o = 2.21122$;	$n_e = 2.13806$

The values of the coefficients of the Sellmeier equations depend on the temperature at which the lithium niobate crystal is operating and should be considered in further studies.

Doping the crystal with rare earth ions such as Erbium would also change the coefficients value of the Sellmeier equations; in addition, the resistance against photorefractive effect or optical damage of the crystal would increase [7]. In this study it was assumed that doped lithium niobate has the same index value as the undoped crystals, as the coefficients of the Sellmeier equations for the Erbium doped crystal are not currently available.

Several interest phenomena were observed during the experiment with doped lithium niobate crystal fibers. First of all, incident light at 808 nm (near IR) caused the Er:$LiNbO_3$ crystal to glow with bright green-blue light (450-570 nm), near the incident end. This effect is mostly caused by band-edge absorption and emission. Although it is also possible that an intense laser light can cause sufficient SHG light in $LiNbO_3$ crystal the wavelength of SHG at 404 nm will yield purplish light. The other observation was that the light was self-focused inside the crystal, due to the gradual change of the crystal refractive indices with light travels faster in center of the crystal where light intensity is highest from the focused incident light.

One of the most important results of this research is the verification of 1550 nm light gain phenomena in the Er-doped $LiNbO_3$ fibers. A ferroelectric waveguide structure with both gain and modulation functions are regarded highly for various photonic applications. However from the preliminary result the gain

seems weak. Quantitative evaluation about gain efficiency would be desirable. Crystals of higher doping level should be also synthesized and tested.

CONCLUSION

Both experimental measurements and theoretical simulations on the transmitted light intensity profiles through a ferroelectric single crystal $LiNbO_3$ fiber are carried out. Theoretical simulations provided good account for the light intensity profile for anisotropic single crystal fibers. The off-center output intensity profile in a $LiNbO_3$ crystal fiber is reported for the first time. Without accounting for the interference among multiple modes, limited success can be expected for the simulation method to interpret the complex specklegram obtained. The simulation and experimental studies were extended to the commercially available single mode and multimode silica optical fibers with satisfactory results. Erbium doped $LiNbO_3$ single crystal fibers were tested for their gain characteristics and it is verified 808 nm pump light can generate broad gain spectrum in the doped crystal fiber near the 1550 nm wavelength, in addition to band edge emission in blue-green region.

In summary anisotropic and nonlinear ferroelectric lithium niobate single crystal fibers may be similarly simulated using advanced beam propagation software, for their transmitted intensity profile. Rare earth (e.g., Er) doped lithium niobate fibers have important potential in communication frequency lasing and frequency conversion. With additional studies on the electric and optical field intensity dependent index, and on proper doing combination and levels, novel applications using such crystal fibers should be explored further.

ACNOWLEDGEMENT

Special thanks go to Dr. Ruyan Guo who is an excellent role model. Dr. Kenneth Jenkins and Mrs. Linda Becker for their awesome contribution in making this research program a good experience and unforgettable moment of my life. I must thank those who did help me in enjoying and doing the experiments: Mr. Liu Hongbo, a Penn Sate graduate student, Mr. Mamour Ba, a graduate student at Lehigh University and his professors Dr. James C. M. Hwang and Dr. W. D. Jemison. Nevertheless, I must also thank my mother and my father, who have always been there for me.

This material is based upon work supported by the National Science Foundation under Grant No. EEC-0244030. The partial financial support of Pennsylvania Center for Optical Technology (COT) is also acknowledged.

REFERENCES

[1] J. C. Burfoot and G. W. Taylor, "*Polar dielectrics and their applications*" pp. 257, 261, 360, 373, MacMillan Press Ltd New York (1979).

[2] R. Guo, A. S. Bhalla, and L. E. Cross, "$Ba(Mg_{1/3}Ta_{2/3})O_3$ Single Crystal Fiber Grown by Laser Heated Pedestal Growth Technique," *J. Appl. Phys.,* **75(9),** pp. 4704-4708 (1994).

[3] R. Guo, Y. Jiang, A. S. Bhalla, and L. E. Cross, "Single Crystal Fibers of Rare Earth Ortho-Niobate: Growth, Ferroelastic Property, and Domain Studies," *Proceedings of the Tenth IEEE International Symposium on Applications of Ferroelectrics,* East Brunswick, NJ, pp. 899-902 (1996)

[4] D. K. Mynbaev and L. L. Scheiner, "*Fiber-Optic Communications Technology*" pp. 46-47, 57-59 Prentice Hall, New Jersey (2001)

[5] S. J. Orfanidis, "Electromagnetic Waves and Antennas" pp. 1-5 Rutgers University (2003) http://www.ece.rutgers.edu/~orfanidi/ewa/ch01.pdf

[6] Man Gu et. al. (unpublished result).

[7] K. K. Wong, "*Properties of Lithium Niobate*"; pp. 25, 68, 115-116, 141 INSPEC, IEEE London (2002)

ANALYSIS OF TUNABLE HORSESHOE RESONATORS USING DIELECTRIC LOADING TECHNIQUES

Rehan Shariff* and Michael Lanagan[#]

Department of Materials Science and Engineering
The Pennsylvania State University
University Park, PA 16802

*Undergraduate student of
The Department of Electrical and Computer Engineering
University of Illinois at Chicago
Chicago, IL 60607

ABSTRACT

Advances in microwave communication systems have stimulated the growth of new materials and continue to challenge microwave filter designers with increasingly rigorous requirements for high performance filters. Microstrip bandpass filters made with high dielectric constant materials can help to reduce device size and cost, while retaining high performance and a broad frequency range. Specifically, this research focused on a folded horseshoe microstrip resonator, which has a narrow bandpass characteristic and low insertion loss.

The objective of this study was to analyze the tunability of the resonator by loading dielectric pads at strategic locations, thereby shifting its resonance frequency. The Finite Difference Time Domain (FDTD) method was explored to better understand resonance phenomena and analyze electromagnetic field distributions in the resonators. Prototype resonators were fabricated using the Low Temperature Co-fired Ceramics (LTCC) process and microwave analysis was performed on the samples using a network analyzer. Finally, the optimum dielectric loading conditions were determined.

[#] Faculty Mentor

INTRODUCTION

Filters play crucial roles in microwave communication and measurement systems. Microwave filters are used to select desired signals within the electromagnetic spectrum. Depending on requirements such as power, frequency, and integration with other microwave devices, filters may be realized through various types of transmission lines including waveguide, coaxial, and microstrip lines. The advance of innovative materials and fabrication technologies has stirred the progress of new microstrip filters for microwave applications [1]. In general, microstrip lines have a broad frequency range and are most easily integrated with microwave devices. The microstrip line illustrated in Figure 1 consists of a dielectric substrate in between a ground plane and a conductor. Although the microstrip line has larger loss than waveguide and coaxial lines, its advantages over these transmission lines include more reliability, more stability, smaller weight, and a wider bandwidth [2].

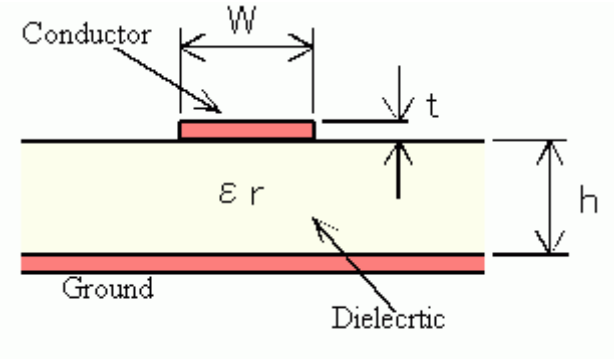


Figure 1: Structure of a microstrip line [2]

Filters are composed of many resonators that are placed in series in a transmission line structure. The response magnitude for a single resonator is shown in Figure 2.

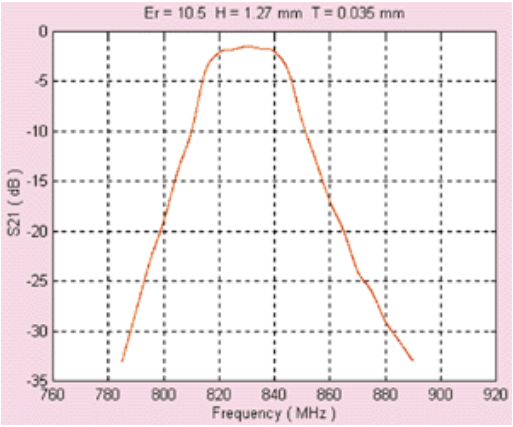


Figure 2: Response of a single resonator [3]

The response magnitude of a filter is the sum of all the individual resonator response magnitudes [2]. The response of a typical bandpass filter is illustrated in Figure 3.

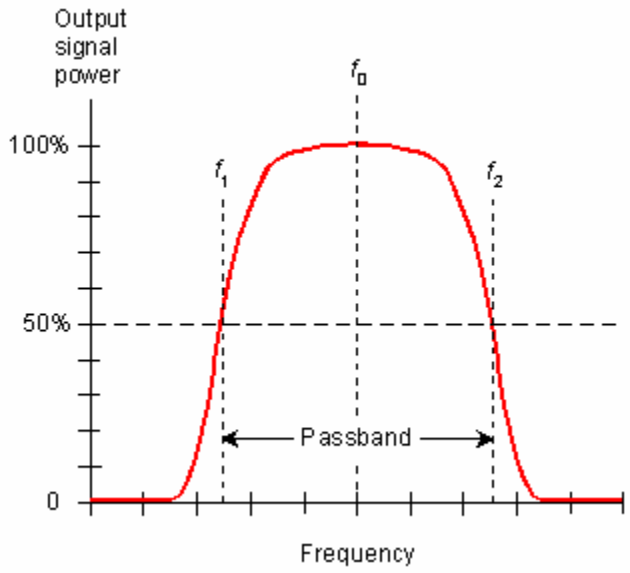


Figure 3: Response of a typical bandpass filter [4]

A microwave ring circuit, one type of microstrip resonator, can be used in many circuit applications including filters, oscillators, and mixers. It has a simple design and its performance is easy to predict [5]. Although ring resonators have great frequency selectivity, their insertion loss is too high for use in practical applications. Instead, a horseshoe resonator may be used. The horseshoe resonator is composed of two horseshoes consisting of a feedline and two capacitive pads joined by a half-ring as shown in Figure 4 [6].

The resonator has good bandpass characteristics and low insertion loss as shown in Figure 5. It can be constructed to utilize either magnetic or electric coupling. While magnetic coupling occurs across two half rings when the two horseshoes are placed in series, electric coupling occurs between two adjacent capacitors when the horseshoes are placed in parallel. The resonance frequency can be adjusted by varying the radius of the rings and the size of the capacitors.

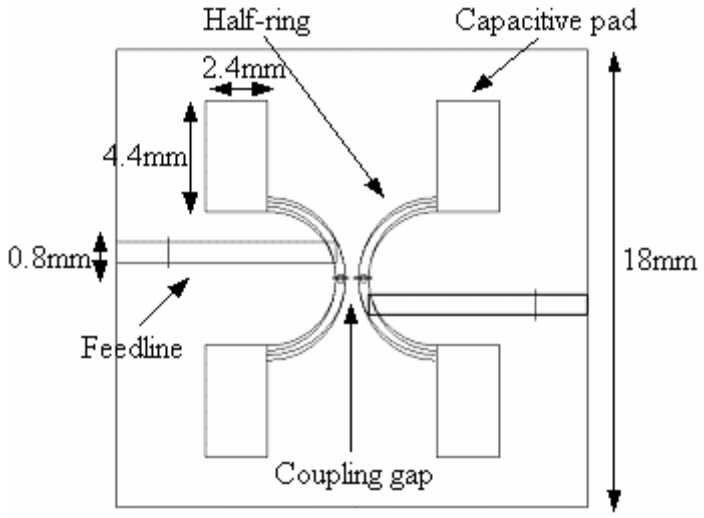


Figure 4: Horseshoe resonator design [2]

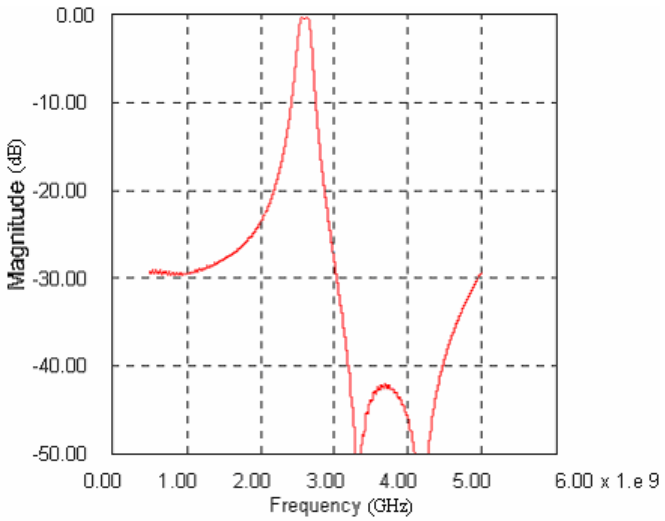


Figure 5: S_{21} response characteristic [2]

The structure of the horseshoe resonator was improved by folding it across the ring gap. The new structure, which was introduced by Takaki Murata, is composed of a dielectric substrate inserted between two horseshoes as shown in Figure 6 [2]. The ground plane is situated at the center of the substrate and is partially cut between the edges of the rings to obtain magnetic coupling.

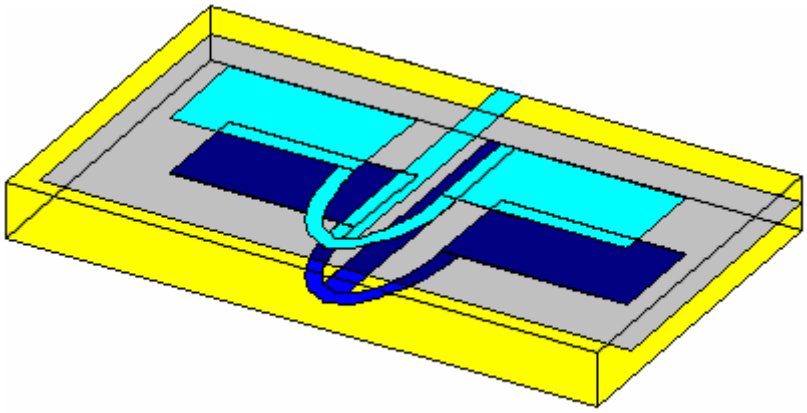


Figure 6: Folded horseshoe resonator design [2]

The S_{21} response characteristic of the folded structure is very similar to the original, but the size has been reduced by half. The folded structure can be constructed to utilize electric coupling by cutting apertures in the ground plane between the capacitors. Attenuation poles can be observed in the S_{21} response by removing the magnetic coupling and using only electric coupling. Figure 7 shows a graph of the S_{21} response characteristic of a resonator using only electric coupling. The folded horseshoe resonator has very good filter characteristics, deep attenuation poles, and may serve as a practical microwave filter [2].

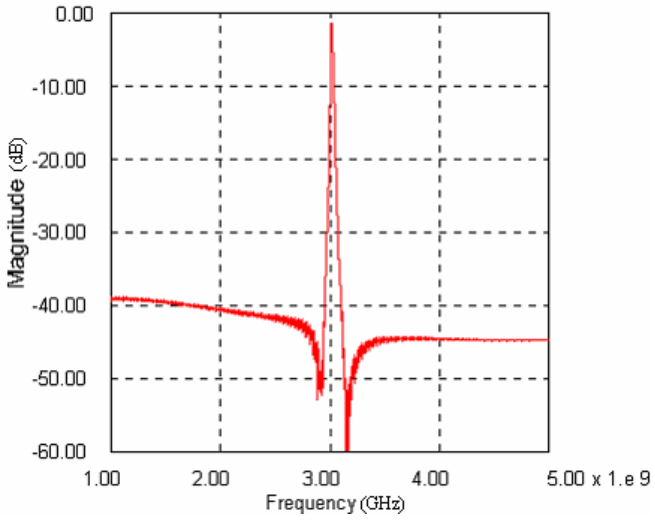


Figure 7: S_{21} response characteristic of a resonator with only electric coupling [2]

Microstrip resonators have many advantages for producing tunable filters over other types of transmission lines. Therefore, investigating the tunability of the folded horseshoe structure may provide interesting future applications. The areas of highest electric field in the horseshoe structure are in the capacitors at the

resonance frequency. When a dielectric pad is loaded onto one of the capacitors, the fringing field passes through the pad as shown in Figure 8. This results in a shift of resonance frequency [2].

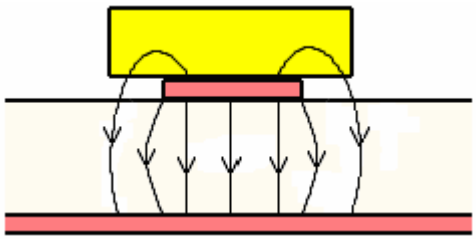


Figure 8: Electric field in a microstrip with a dielectric pad [2]

EXPERIMENTAL DESCRIPTION

The primary purpose of this research was to analyze the tunability of the folded horseshoe resonator by loading high permittivity pads onto the capacitors of the resonator. The frequency tuning approach was explored in two stages. First, Finite Difference Time Domain (FDTD) modeling simulations were conducted by using dielectric pads of several different sizes and dielectric constants. Many different simulations were conducted to determine the optimal loading condition. Next, the simulation results were verified by measuring properties of a prototype using a network analyzer.

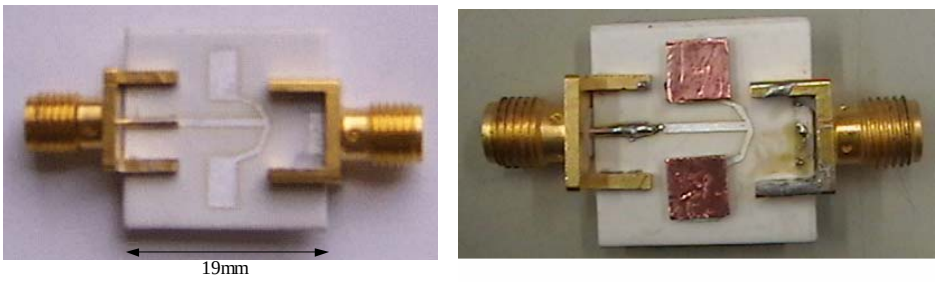
First, to better understand resonance phenomena in the resonators, the FDTD modeling approach was used. It is the best method to analyze electromagnetic field distributions in the resonators. The algorithm solves both electric and magnetic fields in time and space using the following Maxwell's equations [2].

$$\nabla \times \boldsymbol{E} + \frac{\partial B}{\partial t} = J_m \tag{1}$$

$$\nabla \times \boldsymbol{H} - \frac{\partial D}{\partial t} = J_e \tag{2}$$

Next, prototypes of the resonators were fabricated using the Low Temperature Co-fired Ceramics (LTCC) process. LTCC is a multilayer substrate technology that combines low dielectric constant materials with high conductivity metals like silver [2]. By using a LTCC prototype and a few dielectric pads with different dielectric constants, we can easily verify the simulations experimentally. Figure 9a shows the top surface of the folded horseshoe resonator that is attached to a coax to microstrip adapter on the left side of the resonator. The right adapter in Figure 9a is attached to the bottom horseshoe (not shown). The ground plane is attached to the coaxial shield by vias and a silver patch on the surface, which is

seen adjoining the left coax/microstrip adapter. Figure 9b shows the resonator with dielectric pads loaded onto the capacitors.



a) b)

Figure 9: Folded horseshoe resonator prototype a) without any dielectric pads and b) with pads loaded

Furthermore, a network analyzer was used to perform microwave analysis. The Hewlett Packard 8510 Microwave Network Analyzer, shown in Figure 10, was used in this experiment. The frequency range that the HP 8510 is capable of supporting is 45 MHz to 26 GHz. The network analyzer was used to measure scattering parameters in order to obtain resonance frequency and the loss of the sample resonator. Pads of different dielectric constants including ε_r=9, ε_r=22, and ε_r=88 were placed on the capacitors of the sample resonator and the resonance frequency and loss were measured.

Figure 10: HP 8510 Network Analyzer

RESULTS AND DISCUSSION

Simulation Results

The following results were obtained by running several FDTD simulations of 3D models of the folded horseshoe resonator. Simulations were conducted to

study the differences between Takaki Murata's model and the prototype. Simulations were also conducted to determine the loading condition for optimal tunability.

First, the differences between the prototype and Murata's model were examined. The substrate thickness of the prototype was only 0.96 mm, while that of Murata's model was 1.32 mm. In addition, the prototype required vias to connect the ground plane to the input and output ports since the ground plane is located at the center of the substrate. Murata's model did not take this type of a connection into consideration. The vias were modeled at only one of the two ports to simplify the model. Figures 11, 12, and 13 show the results of the simulations of these models.

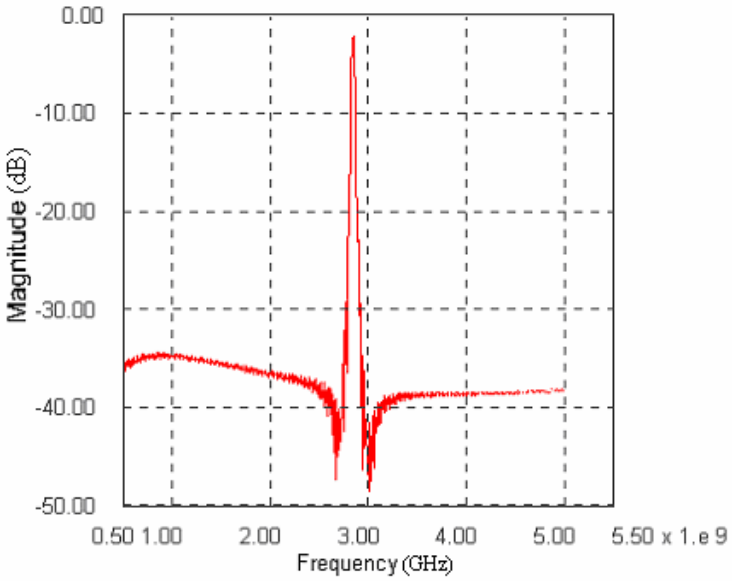
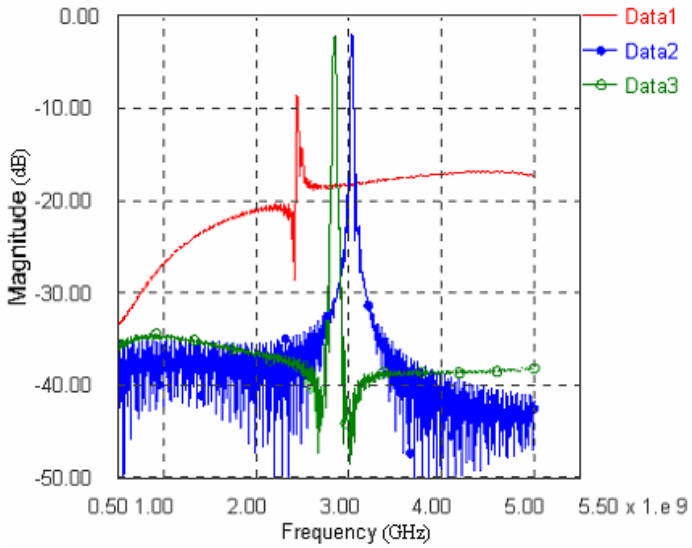


Figure 11: FDTD simulation of the transmission response of a folded horseshoe structure with dielectric substrate thickness of 0.96 mm

The results in Figure 12 indicate that as the thickness of the dielectric substrate decreases, the central frequency also decreases. This occurs because the distance between capacitors becomes shorter, thereby increasing the capacitance and decreasing the resonance frequency. However, when the thickness of the substrate is very low, the characteristics of the S_{21} curve become undesirable.



(Data1 - 0.64 mm, Data2 - 0.96mm, Data3 -1.32mm)

Figure 12: FDTD simulation of the transmission response for varying thickness of dielectric substrate

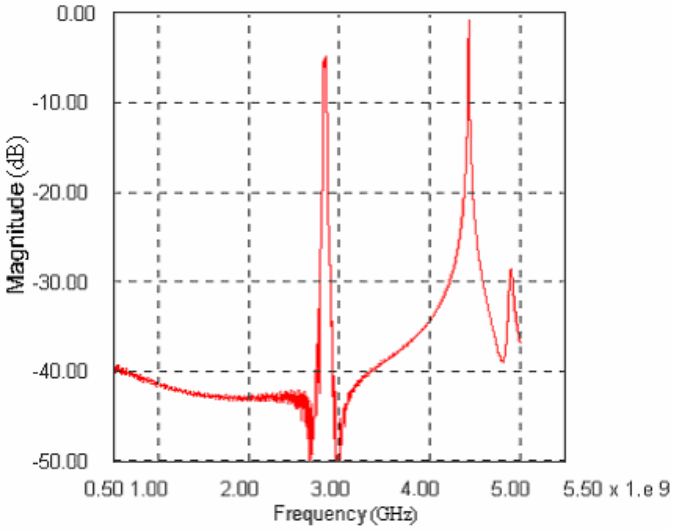
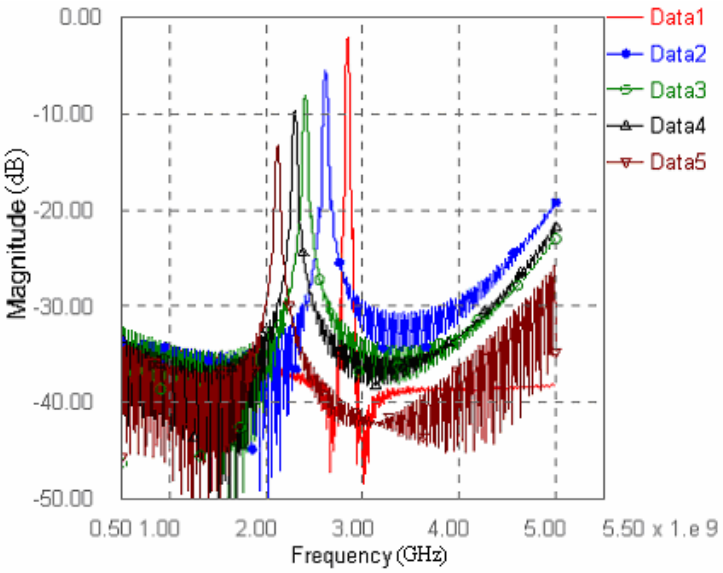


Figure 13: FDTD simulation of the transmission response of a folded horseshoe structure with a modeled connection to ground plane through vias for one connector

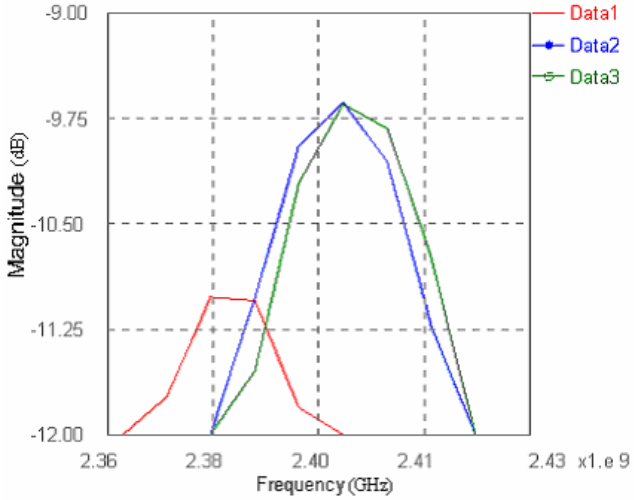
Figure 13 takes into account the coax to microstrip connection to the horseshoe resonator which shows that a single connector has a loss of approximately 5 dB. Therefore, there is a total loss of 10 dB from both connections. The high insertion loss can be corrected with a better connection design to the internal ground plane.

The following graphs are results of FDTD simulations that were conducted with various high permittivity superstrates on the horseshoe resonators in order to study the effect of loading dielectric pads on the resonator. The placement, length, width, and thickness of the pad were varied to determine the optimal loading condition.



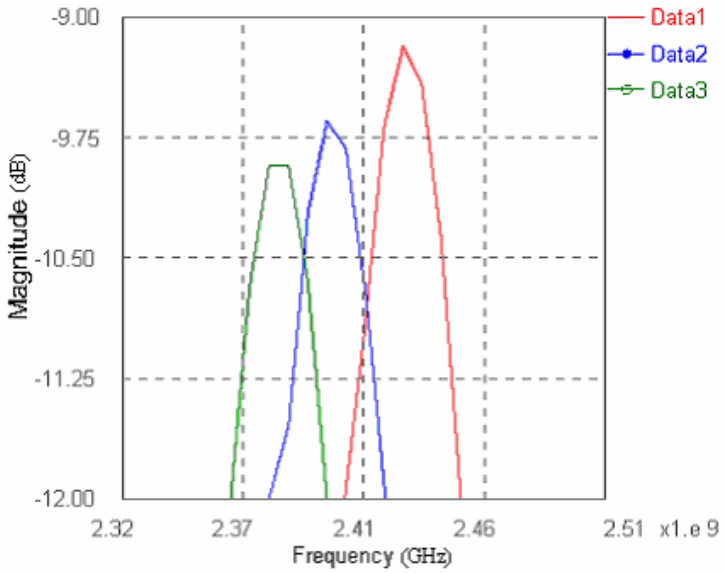
(Data1 – no pads, Data2 – ε_r = 9, Data3 – ε_r = 22, Data4 – ε_r = 35, Data5 – ε_r = 88)

Figure 14: Relationship between K and shifts in central frequency



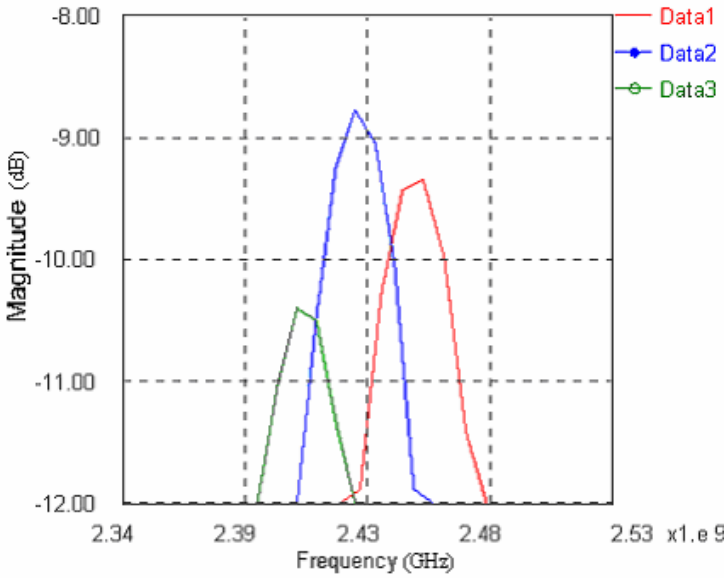
(From right to left: Data1 - .4mm distance from inner edge of capacitor to edge of pad, Data2- .2mm, Data3 - both edges coincide)

Figure 15: FDTD simulation of the transmission response for varying placement of dielectric pads relative to capacitors



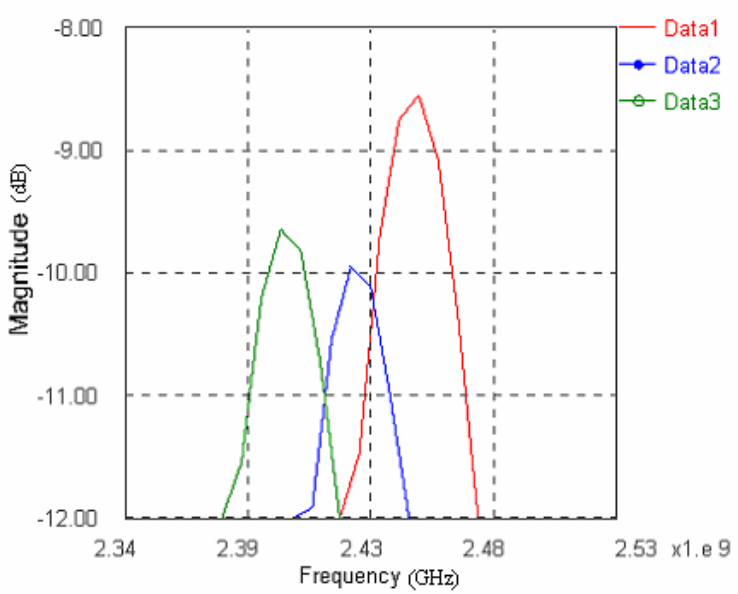
(From left to right: Data1- 1.12 mm, Data2 – 1.32 mm, Data3 – 1.52 mm)

Figure 16: FDTD simulation of the transmission response for varying thickness of dielectric pads



(From left to right: Data1 - 4.8 mm x 4.8 mm, Data2 - 5.0 mm x 5.0 mm, Data3 - 5.2 mm x 5.2 mm)

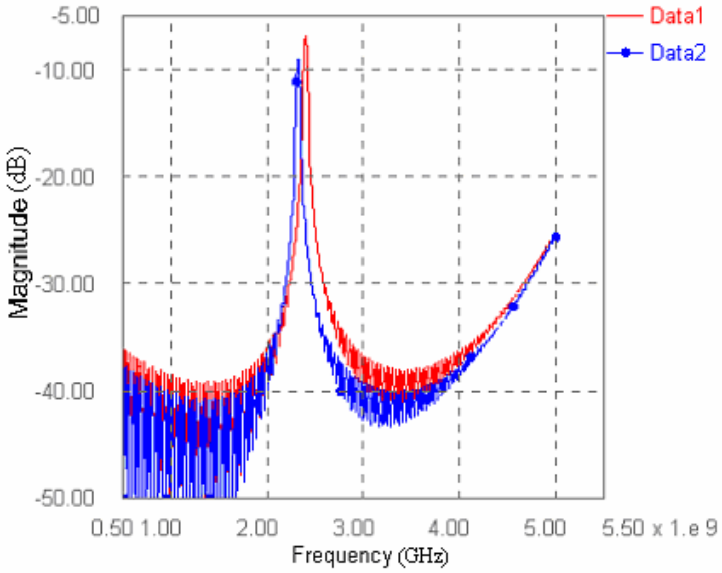
Figure 17: FDTD simulation of the transmission response for varying length and width of square dielectric pads



(Data1 – 4.6 mm x 5.0 mm, Data2 – 4.8 x 5.2 mm, Data 3 – 5.0 x 5.4 mm)

Figure 18: FDTD simulation of the transmission response for varying length and width of rectangular dielectric pads

The graph in Figure 14 indicates that as the dielectric constant of the pads increase, the resonance frequency decreases and the loss increases. Also, these results show that the attenuation poles are lost when the resonator is tuned by loading dielectric pads. Figure 15 shows that the loss increases as the dielectric pads are placed farther apart from the ring edges. Figure 16 shows that increasing the thickness of the pads resulted in an increase in resonance frequency and a reduction in loss. Figures 17 and 18 indicate that an increase in area did not necessarily result in lower loss. The optimum area of the pad was 4.6 mm x 5.0 mm. However, it is clear that as the area of the pads increases, the resonance frequency decreases. This is because the pads can affect more of the electric fringing fields, which leads to an increase in capacitance. Finally, conductive plates were placed on top of the dielectric pads to see the effect on the resonance frequency and on the loss.



(Data1 – without conductive plate, Data2 – with conductive plate)

Figure 19: FDTD simulation of the transmission response for placing conductive plates on the dielectric pads

The results in Figure 19 show there was a slight decrease in the resonance frequency and increase in loss as metal is added on top of the pad.

Experimental Results

The following graphs are of data that was measured using the network analyzer. First, the S21 characteristic of the prototype was taken. Figures 20 and 21 show these results. The results show an insertion loss of approximately 13 dB. This verifies the simulation on the connections of the prototype since we obtained a minimum loss of 10 dB. Finally, the dielectric pads were loaded onto the capacitors of the resonator with high dielectric constants including $\varepsilon_r = 9$, $\varepsilon_r = 22$, and $\varepsilon_r = 88$. The S_{21} characteristics were recorded to see the shift in resonance frequency and the effect on the loss. Figure 22 shows the results.

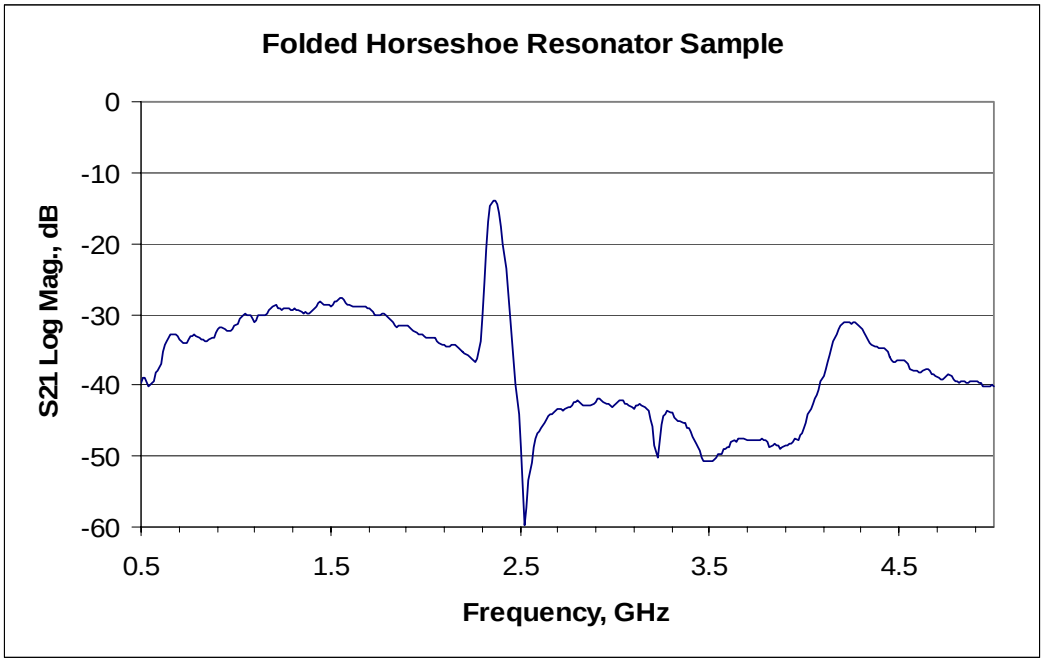


Figure 20: Folded horseshoe resonator sample S_{21} characteristic

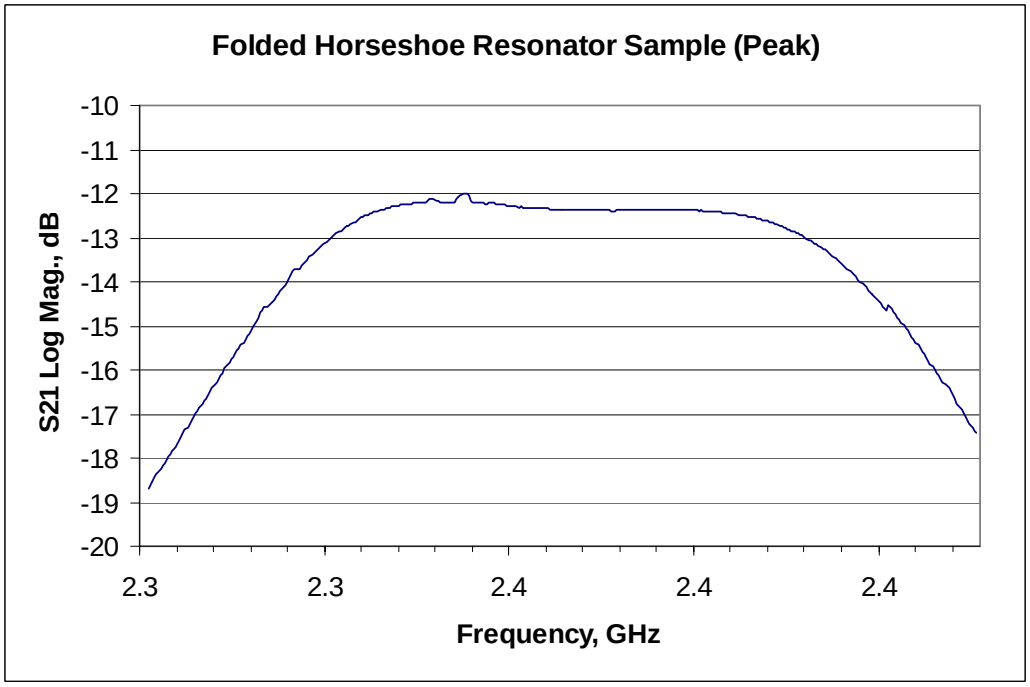


Figure 21: Peak of S_{21} characteristic

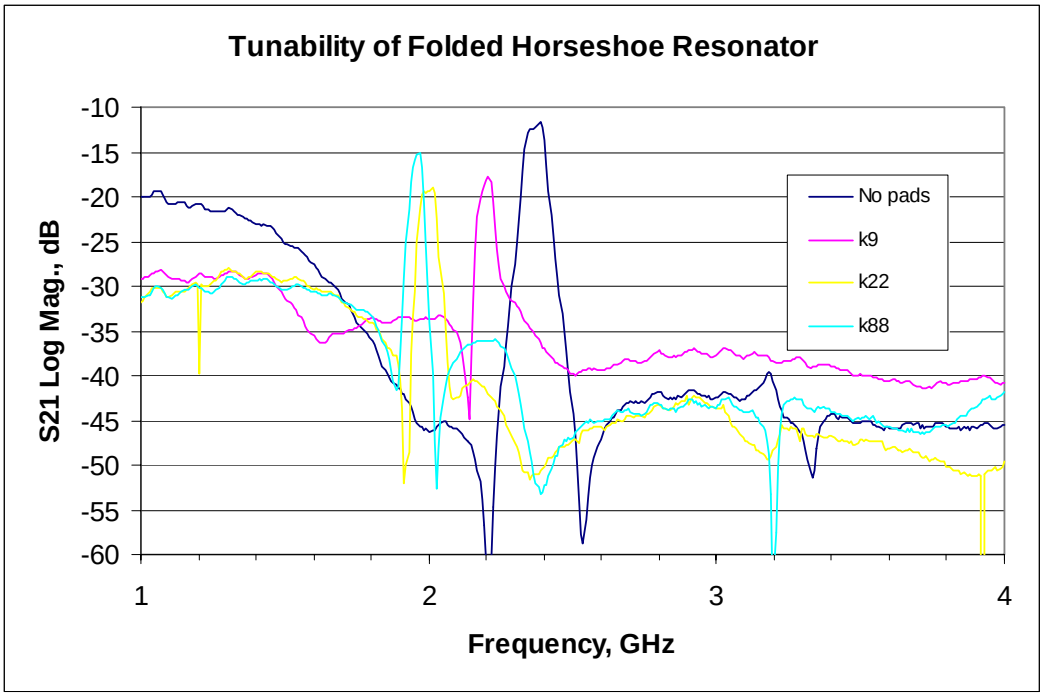


Figure 22: Tunability of the folded horseshoe resonator

These results confirm our simulations. As the dielectric constant increases, the resonance frequency decreases. However, there is no visible trend between permittivity and insertion loss as with the simulations. This is because the simulations do not take the loss of the dielectric material into account.

SUMMARY

The objective of this study was to analyze the tunability of the folded horseshoe resonator by loading dielectric pads onto the resonator and determining the optimal loading condition. The FDTD method was explored and the results indicated that the vias to ground plane connection of the prototype was causing a loss of 10 dB. Although the folded horseshoe structure is much smaller than the original one, it is not as practical for applications until the connection is modified. Also, the attenuation poles were not visible when the dielectric pads were loaded. Other FDTD simulation results showed that as the dielectric constant increased, there was a decrease in frequency and an increase in insertion loss. Also, as the area of the dielectric pads increased, there was a decrease in resonance frequency and an increase in insertion loss. Prototypes were fabricated and microwave analysis was performed on the sample resonators using a network analyzer. The measurements on the samples confirmed the simulation results, and demonstrated that dielectric loading onto a resonator is a promising approach to tunable microwave filters.

ACKNOWLEDGEMENTS

I would like to thank the following individuals whose contributions were essential to this research: Dr. Michael Lanagan, for his knowledge and support as he guided me through the research process; Mr. Steve Perini, for instructing me on the operation of the HP 8510 Network Analyzer; Mr. Khalid Rajab, for his assistance with the Finite Difference Time Domain modeling software; Ms. Amanda Baker, for providing schematics and samples of horseshoe resonators. Special thanks go to the National Science Foundation, the Electrical Engineering Department at the Pennsylvania State University, and Dr. Ruyan Guo for providing this research opportunity. This material is based upon work supported by the National Science Foundation under Grant No. EEC-0244030.

REFERENCES

[1] J. Hong and M.J. Lancaster, "Microstrip Filters for RF/Microwave Applications"; pp. 1-2 in *Introduction*, Edited by Kai Chang, John Wiley & Sons, Inc., New York, 2001.

[2] T. Murata, "Fundamental Aspects of Microstrip Resonators and Their Tunability," *Masters Thesis in Material Science and Engineering*, 2004.

[3] R. Martins, "Techniques Yield Tiny Hairpin-Line Resonator Filters." *Microwaves & RF*, November 1999, <http://www.mwrf.com/Globals/PlanetEE/dsp_article.cfm?ArticleID=9356&Extension=html> (July 21, 2004).

[4] Whatis?com, Nov 11, 2002, <http://whatis.techtarget.com/definition/0,,sid9_gci331403,00.html>, (July 19, 2004).

[5] K. Chang, "Microwave Ring Circuits and Antennas", pp.1-2 in *Introduction*, Edited by Kai Chang, John Wiley & Sons, inc., New York, 1996.

[6] A. Hennings, G.B.Semouchkin, E.A.Semouchkina and M. Lanagan, "Design Optimization of Microstrip Square-Ring Bandpass Filter with Quasi-Elliptic Function," *33rd European Microwave Conf. Proc.*, Munich, pp. 175-178, October 2003.