UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1995

GRADUATE STUDENT RESEARCH PROGRAM FINAL REPORTS

VOLUME 8

PHILLIPS LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way

Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major David Hart

Program Manager, RDL
Scott Licoscos

Program Administrator, RDL
Gwendolyn Smith

Program Administrator
Johnetta Thompson

Reproduced From Best Available Copy

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1995

19981218 064

REPORT DOCUMENTATION PAGE

Form Approved

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the tir
and maintaining the data needed, and completing and reviewing the collection of information. Send comments re
information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for inf
1204, Arlington, VA 22202-4302, and to the Office of management and Budget, Paperwork Reduction Project (0704-01

AFRL-SR-BL-TR-98-

0817

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT T
	December, 1995	Final

4. TITLE AND SUBTITLE
USAF Summer Research Program - 1995 Graduate Student Research Program Final Reports, Volume 8, Phillips Laboratory

5. FUNDING NUMBERS

6. AUTHORS
Gary Moore

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research and Development Labs, Culver City, CA

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NI
4040 Fairfax Dr, Suite 500
Arlington, VA 22203-1613

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Contract Number: F49620-93-C-0063

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The United States Air Force Graduate Student Research Program (USAF- GSRP) is designed to introduce university, college, and technical institute graduate students to Air Force research. This is accomplished by the graduate students being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory Technical Directorates and Air Force Air Logistics Centers. Each participant provided a report of their research, and these reports are consolidated into this annual report.

14. SUBJECT TERMS
AIR FORCE RESEARCH, AIR FORCE, ENGINEERING, LABORATORIES, REPORTS, SUMMER, UNIVERSITIES

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

This document is one of a set of 16 volumes describing the 1995 AFOSR Summer Research Program. The following volumes comprise the set:

VOLUME	TITLE
1	Program Management Report
	Summer Faculty Research Program (SFRP) Reports
2A & 2B	Armstrong Laboratory
3A & 3B	Phillips Laboratory
4	Rome Laboratory
5A, 5B, & 5C	Wright Laboratory
6A & 6B	Arnold Engineering Development Center, Wilford Hall Medical Center and Air Logistics Centers
	Graduate Student Research Program (GSRP) Reports
7A & 7B	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10A & 10B	Wright Laboratory
11	Arnold Engineering Development Center, Wilford Hall Medical Center and Air Logistics Centers
	High School Apprenticeship Program (HSAP) Reports
12A & 12B	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A&15B	Wright Laboratory
16	Arnold Engineering Development Center

DEVELOPMENT OF IONOSPHERIC TOMOGRAPHY RECONSTRUCTION ALGORITHMS USING A THREE DIMENSIONAL IMAGING VOLUME

Vala Afshar
B.S., University of Massachusetts Lowell
Department of Electrical Engineering

University of Massachusetts Lowell
One University Avenue
Lowell, Massachusetts 01854

Final Report for:
Graduate Student Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Laboratory
Hanscom Air Force Base

September 1995

DEVELOPMENT OF IONOSPHERIC TOMOGRAPHY RECONSTRUCTION ALGORITHMS USING A THREE DIMENSIONAL IMAGING VOLUME

Vala Afshar
B.S., University of Massachusetts Lowell
Department of Electrical Engineering

Abstract

The basic assumption in most ionospheric tomography algorithms is that the ground stations, and satellite orbit are approximately coplanar. However, typical ray paths from the receiver, to an orbiting satellite are not coplanar. Real world receiver locations are determined by convenience and availability of sites, and depending on the orbit involved, the satellite could be as much as 20 degrees removed in longitude from the chain of receivers. Therefore the receivers will not all be in a line, nor will they be directly under satellite passes. Thus the total electron content *(TEC)* data ray paths may have passed through regions of the ionosphere far different from an assumed *imaging plane*. If these "real world" geometric considerations are not addressed, it is possible that certain details of the ionosphere, such as small-scale features, will not be accurately reconstructed in the presence of highly non-coplanar geometry.

In this project, a robust and efficient algorithm to allow three-dimensional reconstruction was developed. The algorithm utilizes a three-dimensional *imaging volume*. The linear relationship between measured *TEC* data and electron density has been extended to a full three dimensional case, where the lengths of the ray paths are from the orbiting satellite to the ground-based receiver, through a discrete *imaging volume,* are represented as the elements of a *matrix A*.

Introduction

In general, most ionospheric tomography algorithms are not sensitive to many of the real world configurations, in that they assume that the data measurement paths and a *region of interest* are coplanar. In ionospheric tomography, however, the chain of ground-based receiver stations, the ionospheric region of interest, and the satellite orbit are not necessarily coplanar. This sort of highly non-coplanar geometry forces the algorithms to use some sort of planar projection, a method whereby ray paths are projected on an *imaging plane*. The *imaging plane* represents a region of the earth's ionosphere where the satellite orbit, chain of receivers, and the imposed grid structure are close to being coplanar.

Before developing the three-dimensional algorithm to calculate the lengths of the ray paths from an orbiting satellite to a receiver, through a discrete *imaging volume*, I had produced code to calculate the *matrix A* using a two-dimensional polar grid, having only one longitude cell structure. The algorithm developed for the two-dimensional polar grid structure, utilizing ray projections on to an imaging plane, was robust and quite efficient. However, using two-dimensional ionospheric tomography reconstruction techniques exhibited geometric considerations which had to be addressed in the data reduction process prior to the execution of the algorithm. Data reduction meant targeting the satellite passes which where azimuthally aligned (coplanar), as well as near the chain of ground-based receivers. Also, satellite passes with higher elevation angles were preferred.

The resulting algorithm utilizes a three-dimensional imaging volume, with no geographical restrictions placed on the satellite and or receiver locations, hence completely eliminating the data reduction procedure.

Methodology

Assuming that the electron density is constant within a pixel (cell) of an imposed imaging volume, a set of linear equations can be obtained that relate the total electron content data to the electron density as: [1]

$$\underline{c} = A.\underline{x} \qquad (1)$$

Where $\underline{c}$ is an $(Nx1)$ vector of total electron content data measurements, $\underline{x}$ is an $(Kx1)$ vector of electron density data, and A is an (NxK) matrix with elements Ank that are the lengths of the ray paths for the nth TEC measurement through the kth cell of the electron density array. Further assume that the ionosphere is arbitrary divided into a finite number of pixels (cells), each cell numbered in some agreed manner. This is illustrated in Figure 1. The cells should be small enough so that the electron density is approximately constant. [2]

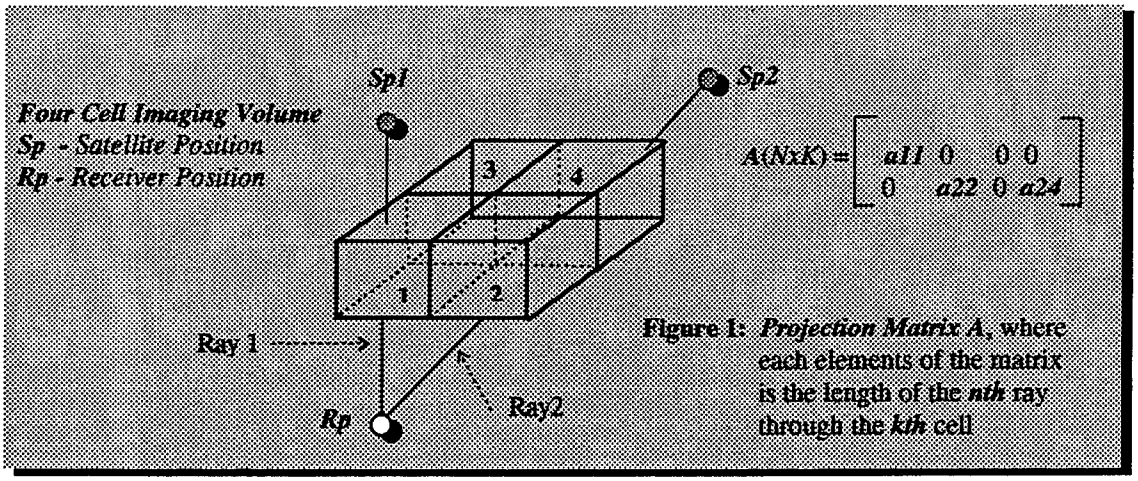


Figure 1: *Projection Matrix A*, where each elements of the matrix is the length of the *nth* ray through the *kth* cell

The following assumptions were made by the principle investigator prior to the development of the resulting algorithm:

. Spherical coordinate system, deemed most appropriate by the principle investigator for the three-dimensional ionospheric tomography reconstruction algorithms, represents the geographical locations of all data points. Producing an *imaging volume* bounded by spheres of constant radii, cones of constant latitude, and planes of constant longitude, utilizes a *spherical grid* with uniform radial, longitude, and latitude cells.

. In the spherical coordinate system, the set of all points whose ρ coordinates are between ρ and $\rho + \Delta\rho$, whose θ coordinates are between θ and $\theta + \Delta\theta$, and whose ϕ coordinates are between ϕ and $\phi + \Delta\phi$, is a solid. The *imaging volume* is shown in Figure 2.

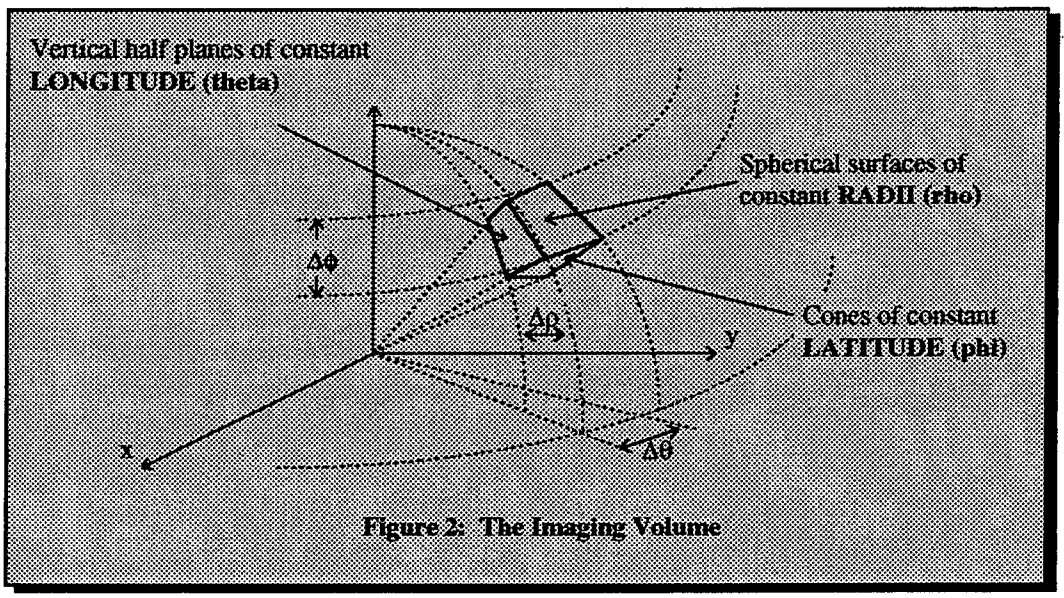


Figure 2: The Imaging Volume

Each cell in the imaging volume is bounded by six surfaces: 2 spheres of radii ρ and $\rho + \Delta\rho$, 2 cones of constant latitude, with half-vertex angles ϕ and $\phi + \Delta\phi$, and 2 planes of constant longitude, with polar angels θ and $\theta + \Delta\theta$. Surfaces of the imaging volume are illustrated in Figure 3.

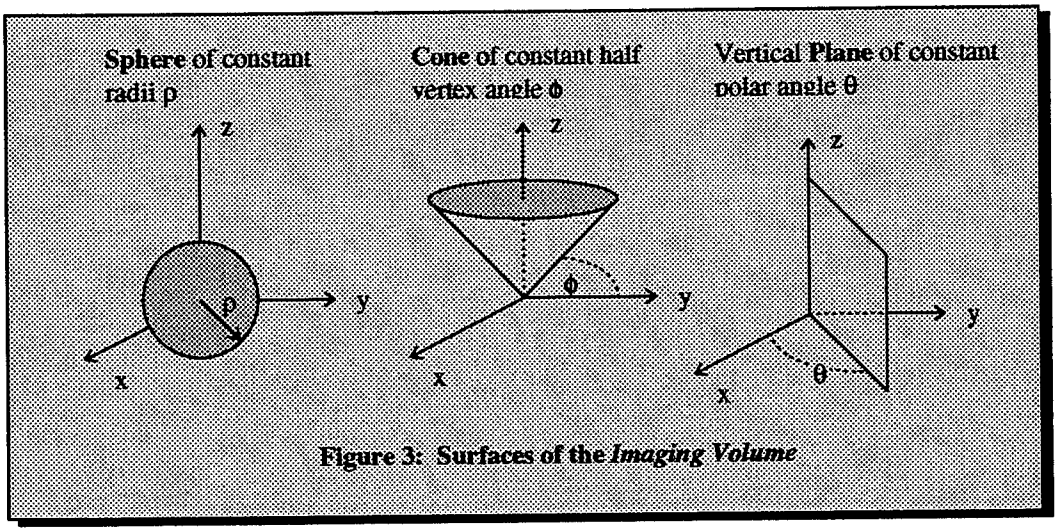


Figure 3: Surfaces of the *Imaging Volume*

. The signal, ray path, is transmitted from the orbiting satellite positioned at *Ps(xs,ys,zs)* and detected at the ground-based receiver positioned at *Pr(xr,yr,zr)*, in three-dimensional space, *characterized* as a straight line. The Equation of the line joining Pr (xr,yr,zr) and Ps (xs,ys,zs) is:

$$(x\text{-}xs) = (y\text{-}ys) / A = (z\text{-}zs) / B \qquad (2)$$

$$A = (yr\text{-}ys) / (xr\text{-}xs) \qquad (3)$$

$$B = (zr\text{-}zs) / (xr\text{-}xs) \qquad (4)$$

. A *valid* ray path must enter the *top spherical* surface, and exit the *bottom spherical* surface of the imaging volume. Hence the imaging volume must extend far enough so that it lies below the orbit of the satellite and above the ground-based receiver, in the region of the ionosphere where the electron density exists.

Algorithm Development

The programs were developed using *MATLAB*, software written in fourth-generation language used in many signals and or image processing applications. The algorithm is structured to execute as a function, having a line of code that contains the syntax definition for the function, consistent with the following format: ***[Matrix A] = A3D (receiver, satellite, grid information)***. A *general* block-diagram of the algorithm is shown in Figure 4.

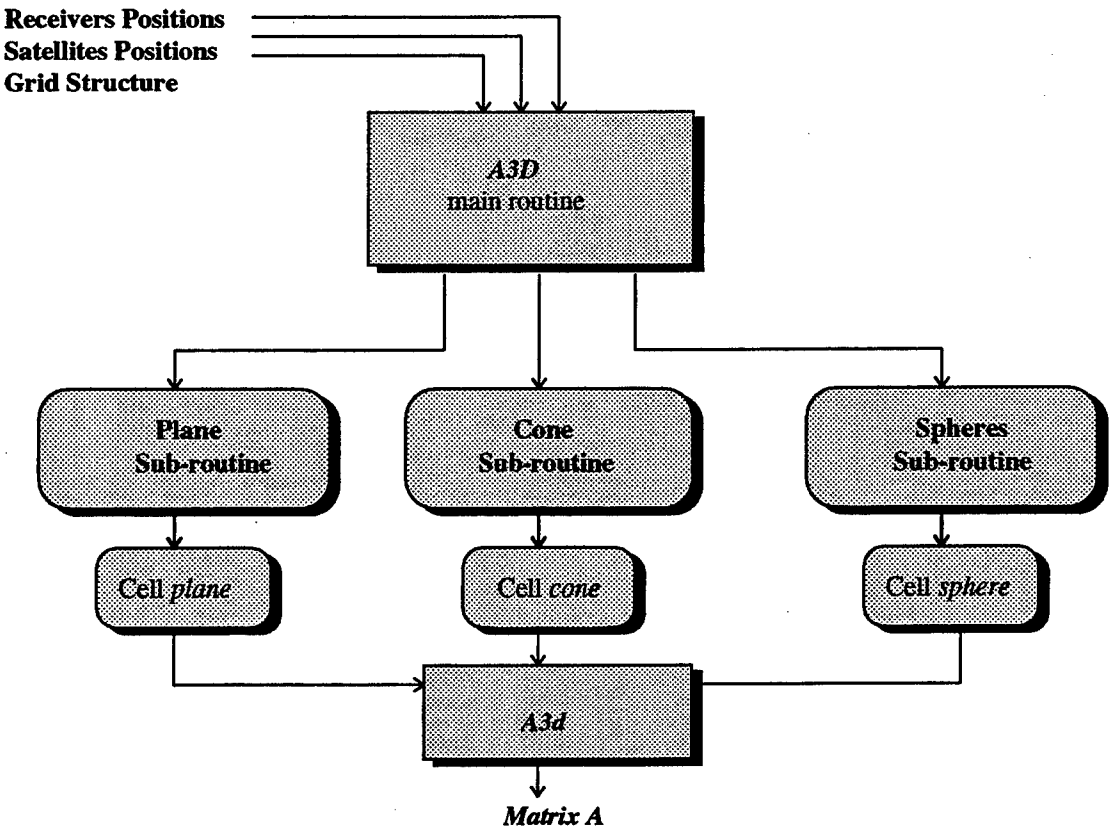


Figure 4: General Block Diagram of the Algorithms

A. Input Data - The required input data are:

1. Locations of the *receivers:* A (*mx3*) matrix where *m* is the number of receivers. The columns of matrix define the latitude, longitude, and altitude of the receiver respectively.

2. Locations of the orbiting *satellite*: A (*nx3*) matrix where *n* is the number of satellite positions. The columns of the matrix define the latitude, longitude, and altitude of the orbiting satellite.

3. *Grid data*: A (*px1*) vector with each element of the vector defining the external boundaries of the *imaging volume*, and the desired partitioning of the imposed *grid* structure. An example of the grid data vector would have the following format:

$$\text{Grid data} = [\ \phi\text{max}\ \ \phi\text{min}\ \ N_\phi\ \ \text{Rmax}\ \ \text{Rmin}\ \ N_R\ \ \theta\text{max}\ \ \theta\text{min}\ \ N_\theta\ ]$$

B. Logic & Structure - The following calculations are performed in the same sequence as they appear in the reconstruction algorithm.

1. Calculate the Cartesian coordinates of the intersection point *P1*, with the ray and the *top* spherical surface of the imaging volume. The following equation:

$$x^2 + y^2 + z^2 = r^2 \qquad (5)$$

and

$$(x\text{-}xs) = (y\text{-}ys) / A = (z\text{-}zs) / B$$

are solved, where *r* is the constant radius of the sphere, shaping the top surface of the imaging volume as shown in Figure 2.

2. Determine if the intersection point in (1), occurs within the longitude and latitude ranges of the imaging volume, governing the boundaries of the *top* surface of the imaging volume.
3. Disregard the satellite and receiver location pair, if the intersection is outside the imaging volume.
4. Calculate the Cartesian coordinates of the intersection point *P2*, with the ray and the *bottom* spherical surface of the imaging volume, using *equations (2) and (5).*
5. Determine if the intersection point in (4) occurs at the longitude and latitude ranges, governing the boundaries of the *bottom* surface of the imaging volume.
6. Disregard the satellite and receiver location pair, if the intersection is outside the imaging volume.
7. If *P1* and *P2* are *valid* (via steps 2 &5), calculate the Cartesian coordinates of all internal surface intersections, within the imaging volume. The following equations, along with the equation of the ray path, *equation (2),* were used to calculate the surface intersections of the ray within the imaging volume:

   a. Ray intersections with *cones* of constant latitude (ϕ):

   $$x^2 + y^2 = r^2 \qquad (6)$$

   where r is the radius of the cone. Knowing r is a function of height, a quadratic-equation can be solved using:

   $$x^2 + y^2 = [\,(x_o / z_o)*z\,]^2 \qquad (7)$$

   $$x_o = r*cos(\phi) \qquad (8)$$

   $$z_o = r*sin(\phi) \qquad (9)$$

   b. Ray intersections with *planes* of constant longitude (θ):

   $$Ax + By + Cz = D \qquad (7)$$

   which reduces to: $x = B*y,$ (8)

   where, $B = cos(\theta) / sin(\theta)$ (9)

   c. Ray intersections with *spheres* of constant radii (ρ) uses *equation (5).*

Note: All the intersection points within the imaging volume are found repeating the using the same procedure.

8. Cell number identification: Given the surface intersections point found in (7), identify the cell number within the imaging volume. A ray path within the imaging volume will only intersect a few cells. The philosophy behind the algorithm is to determine only those cells which have ray intersection. This approach means that the program must perform two main tasks. The first task is to simply calculate the intersection of the ray with the surfaces, within the imaging volume. The second task, is to associate a cell number to those intersection points in an efficient manner. Knowing that within the imaging volume, an intersection point can lie on the surface of as many as two to eight cells, a *checking* mechanism must exist in order to identify the appropriate pair of cell numbers.
9. Construct a *list*, a ($m \times 5$) matrix, where m is number of *valid* rays, recording the coordinates of all intersection points along with corresponding cell numbers. The location of the intersection points in a three-dimensional orientation will be stored in the first three column of the *list*, and the two cell numbers in columns 4-5. The intersection points and corresponding cell numbers are arbitrarily stored.
10. Sort the list.
11. Construct the matrix A, from the *sorted list*, where A is an ($N \times K$) matrix with element *Ank* that are the lengths of the ray paths for the *nth* TEC measurement through the *kth* cell of the imaging volume. Assuming that there are N number of rays, N being a function of the number of satellite/receiver positions, and K number of cells, where $K > N$, then the matrix A is quite sparse, with typically less than one percent of its elements non-zero due to the small number of cells having non-empty intersections with each ray. Using MATLAB's built-in *sparse* routine, the algorithm will produce a *sparse* form of the full matrix A, where all the zero elements have been squeezed out.

The block diagram in Figure 5, illustrates the algorithms. There are three sub-routines which produce the intersection points of the ray path and the surfaces boundaries of the imaging volume. There are also three sub-routines which identify the cell numbers corresponding to the ray path intersecting surfaces of constant radii, latitude, and longitude. In Figure 6, the sub-routines are described in terms of their input and output requirements.

Figure 5: Block Diagram: The Sequence of Operations for a Single Satellite & Receiver Location

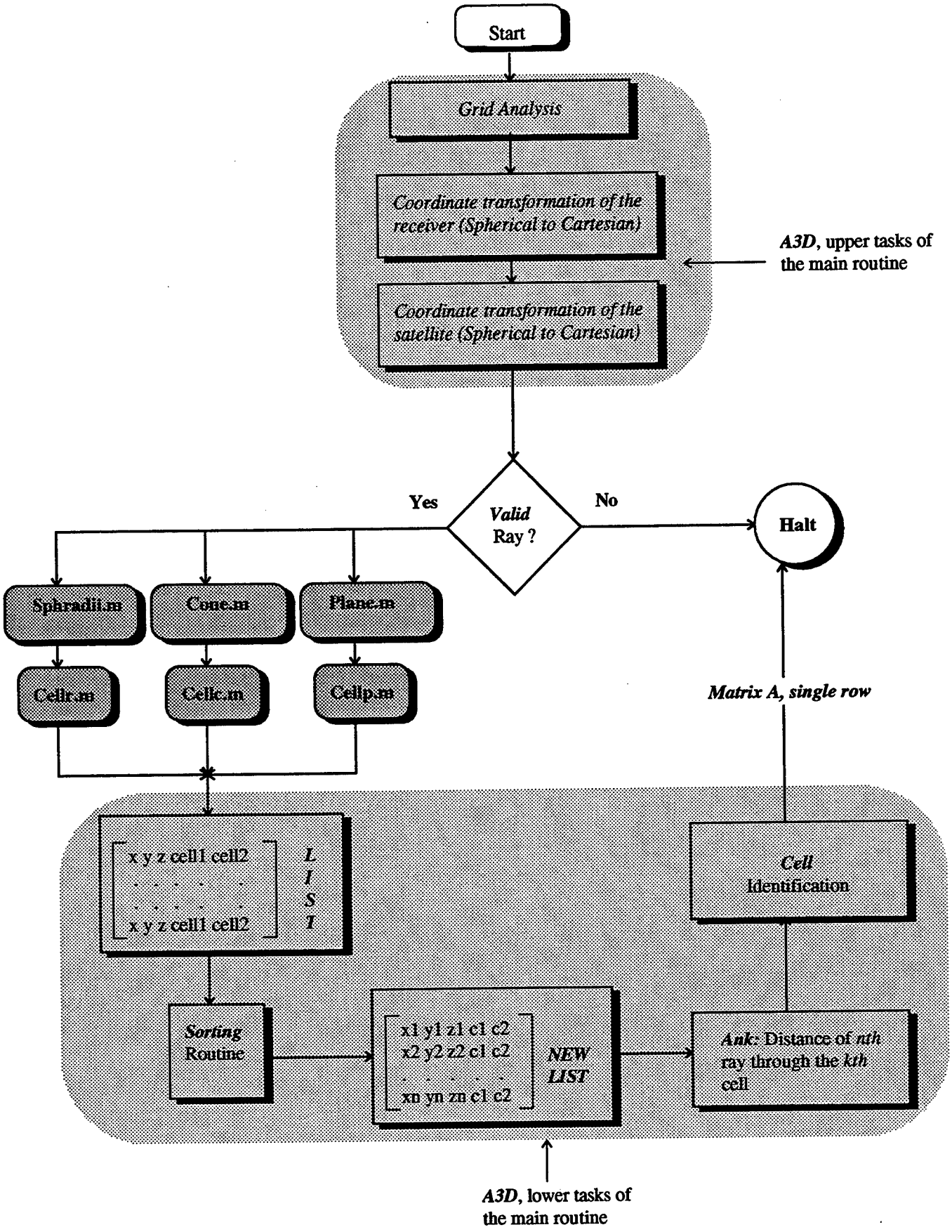


Figure 6. The Sub-Routines: 6a-6c (Intersection Point Calculations) & 6d-6f (Cell Identification)

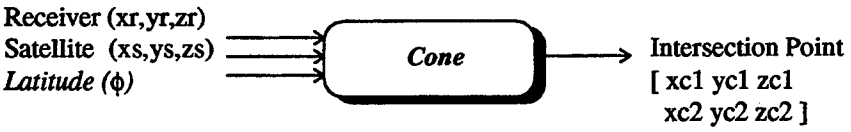
Sphradii function - Routine will calculate the Cartesian coordinates of the intersection of a line with a spheres of constant *radius r*.

(a)

Plane function - Routine will calculate coordinates of the intersection of the ray path with a plane of constant *longitude* θ.

(b)

Cone function - Routine will calculate coordinates of the intersection of the ray path with a cone of constant *latitude* ϕ.



(c)

CellR function - Identify the cell number, a two element vector, within a 3D imaging volume, given a constant *spherical radii* intersection point.

(d)

CellC function - Identify the cell number, a two element vector, with a 3D imaging Volume, given a constant latitudinal cone intersection

(e)

CellP function - Identify the cell number, a two element vector, with a 3D imaging Volume, given a constant longitudinal plane intersection

(f)

Results & Analysis

I tested the algorithms using MATLAB and the following input data: ground-based receiver location, a single satellite location in orbit, and a three-dimensional imaging volume. A single satellite location targeted in orbit, and a ground-based receiver will produce a *(1xK)* vector; one row of the *matrix A*. Using MATLAB, I have generated three test cases, each with different satellite and receiver locations. The following input data for each case are:

case 1: Satellite is directly above the receiver (coplanar)

Receiver location: 42.5N -71.3W 6370km

Satellite location: 43.7N -72.5W 7300km

case 2: Satellite at the north-west and receiver at the south-east of the imaging volume

Receiver location: 24.3N -71.4W 6370km

Satellite location: 58.2N -82.7W 7400km

case 3: Satellite at the north-east and receiver at the south-west of the imaging volume

Receiver location: 22.3N -84.1W 6370km

Satellite location: 56.7N -70.8W 7500km

All three test cases used the same imaging volume with the following data defining the imposed grid structure: Grid data: {60N, 20N, 20, (6370+800)km, (6370+150)km, 26, -70W, -85W, 5}.

For each test case there are two MATLAB generated figures. The first figure illustrates the satellite and receiver locations in a three-dimensional orientation. The surface intersections of the ray path with the imaging volume is also shown on the same illustration. The second MATLAB figure, is a plot of the ray length vs. cell number; one row of the *Matrix A*. The results are shown in Figures 7-9.

Discussion

The linear relationship between measured TEC data and electron density has been extended to a full three-dimensional case, where the lengths of the ray paths are from the orbiting satellite to the ground-based receiver, through a discrete imaging volume, are represented as the elements of the matrix A. The primary objective of the resulting algorithm is to accurately utilize computerized ionospheric tomography, at the same time remain sensitive to many of the real word configurations. The algorithm is found to address all these issues in a robust and efficient manner.

Figure 7. Test Case 1

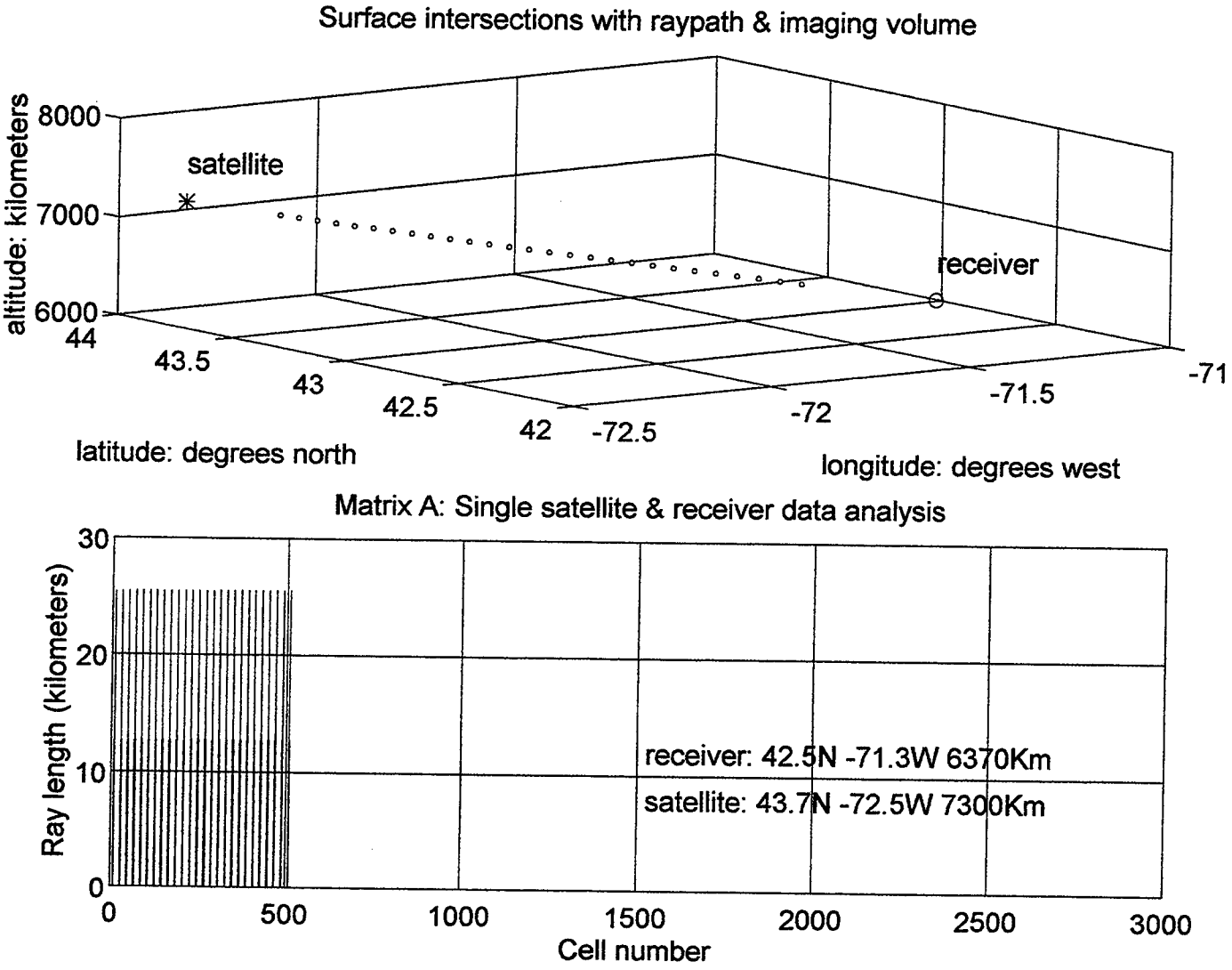


Figure 8. Test Case 2

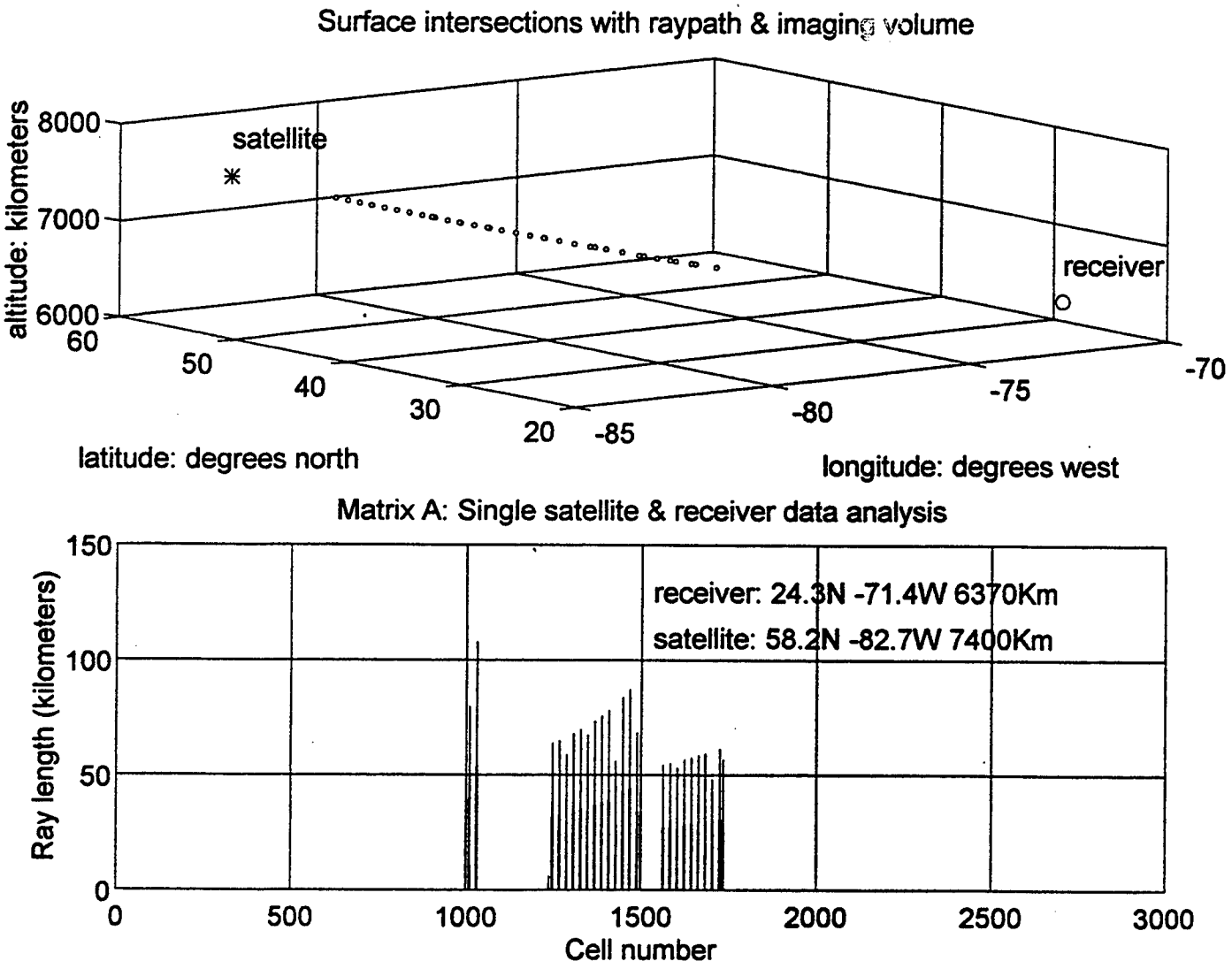
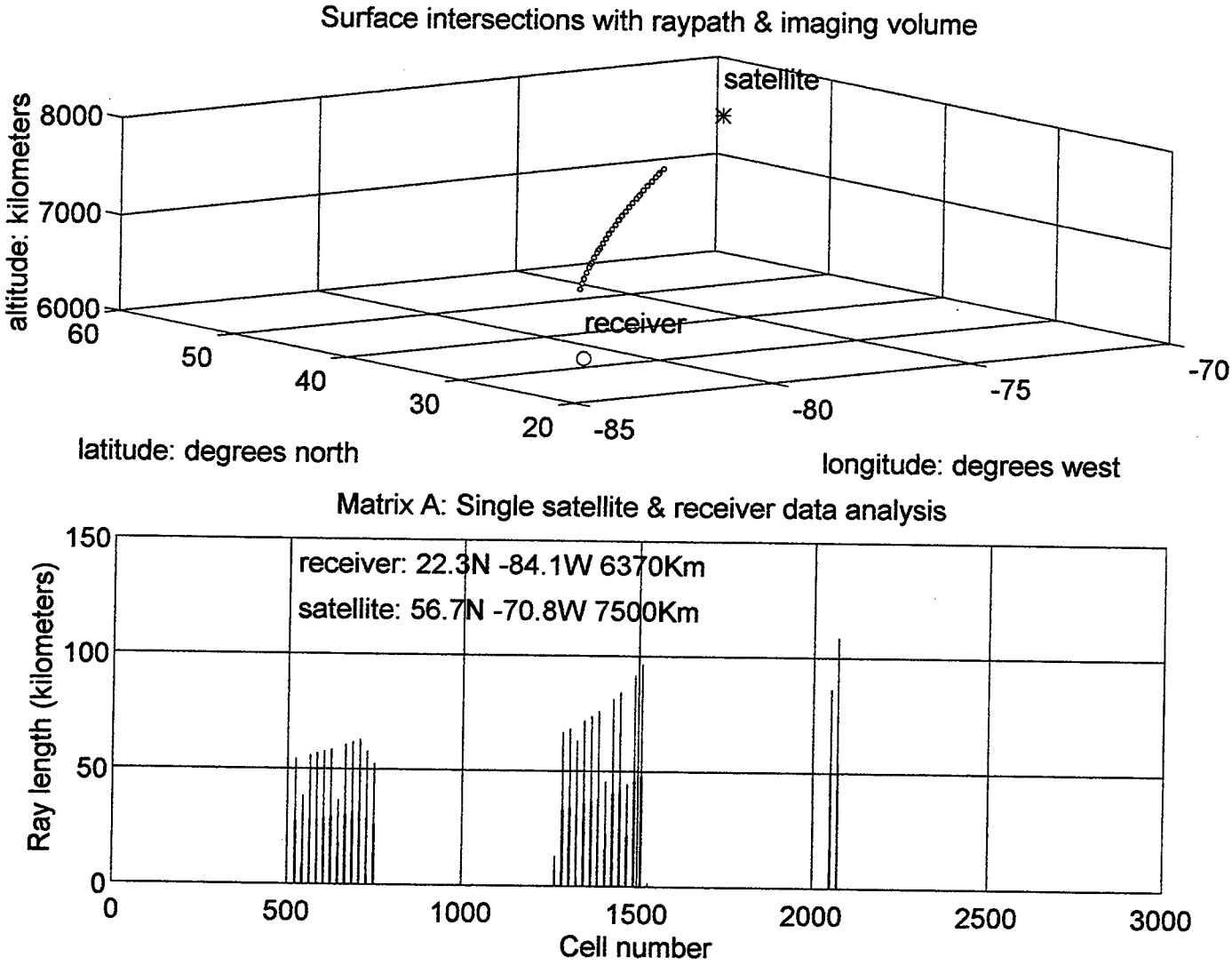


Figure 9. Test Case 3



References

1. Walter S. Kuklinski, "Ionospheric Tomography Using a Model Based Transformation Maximum Entropy Technique", Air Force Office of Scientific Research, pp. 2-4, 1994.

2. Walter S. Kuklinski, "Development of Robust Ionospheric Tomography Algorithms Using Model-Based Maximum Entropy Parameter Estimation", pp. 1-11, 1995.

A STUDY OF LG ATTENUATION IN THE VICINITY OF THE KYRGYZ ARRAY, KYRGYZSTAN

Thomas S. Anderson
Graduate Student
Department of Geology and Geophysics

University of Connecticut
Box U-45, Beach Hall
Storrs, CT 06269

Final Report for:
Graduate Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Laboratory

August 1995

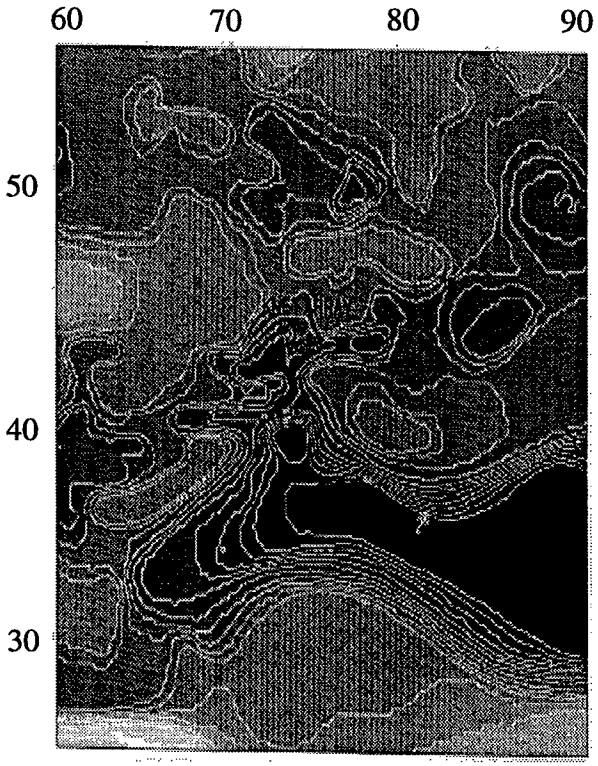


Figure 2a. Kyrgyz array and events on contour of Moho topography.

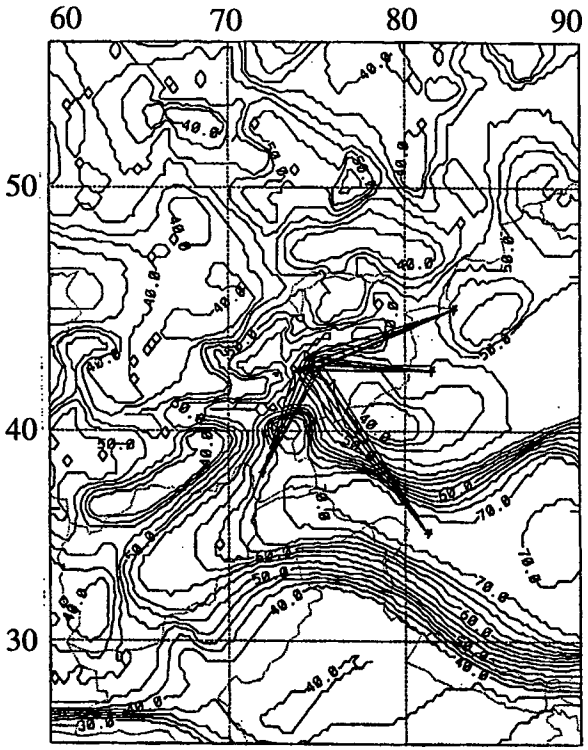


Figure 2b. Contour map of Moho topography with 4 events in this study.

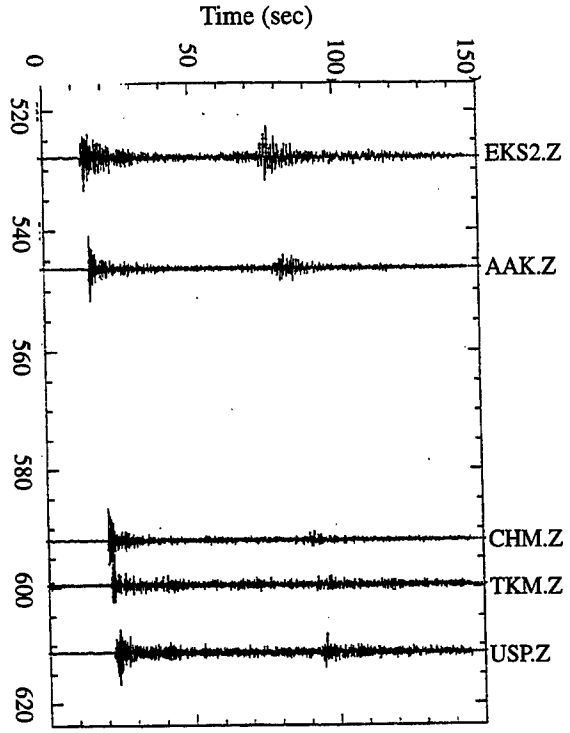


Figure 3a

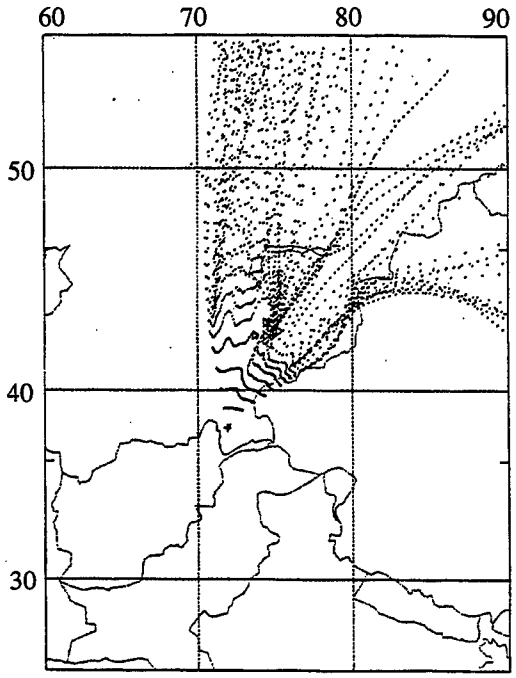


Figure 3b

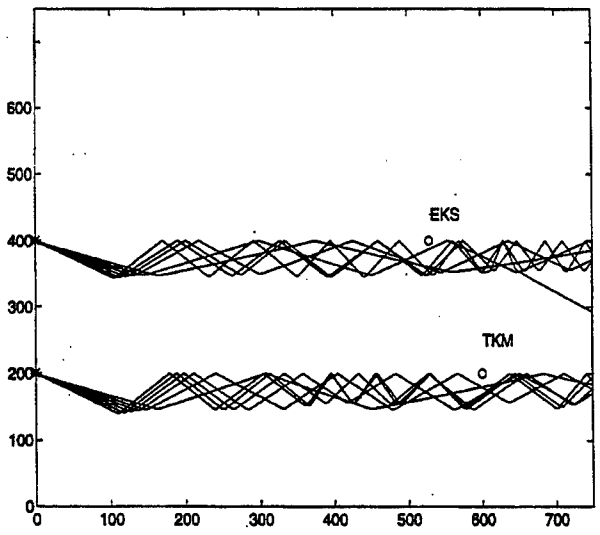


Figure 3c

Figure 3a. Record section plot of event #1, 6 May 92. Data was band passed filtered 1-5 Hz. The signals vary greatly across the array, station TKM has little or no Lg.

Figure 3b. Map view of ray bounce points from Dynamic Ray Tracing for event #1. The wave front is strongly affected by the steep Moho topography. A two-dimensional caustic appears to be forming down the middle of the array.

Figure 3c. Rays for event #1 at take off angles 63-71 degrees for paths to stations TKM and EKS2. Both paths show development of unique focusing and defocusing of rays due to the gradients in the Moho topography.

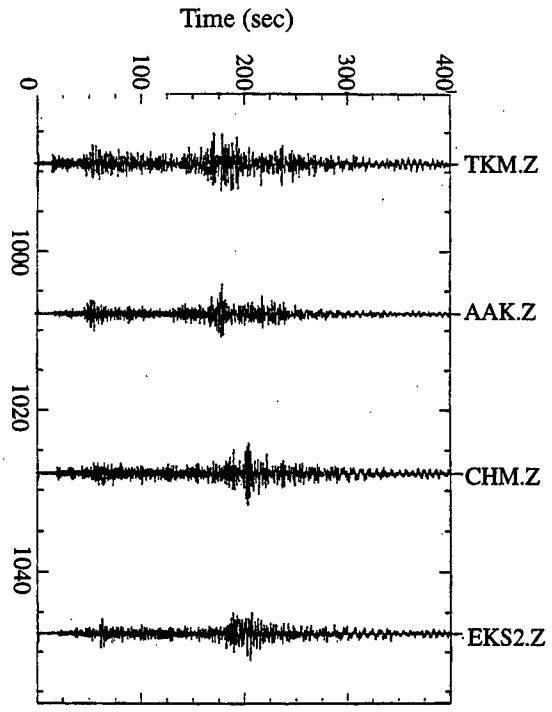


Figure 4a

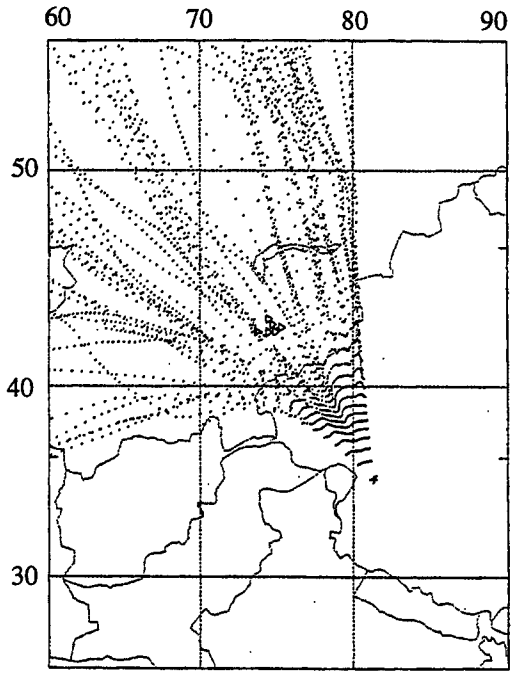


Figure 4b

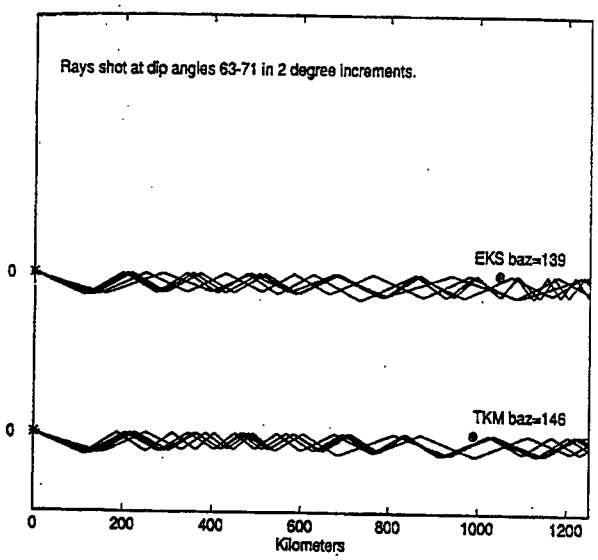


Figure 4c

Figure 4a. Record section plot of event #2, 27 June 92. Data was band passed filtered 1-5Hz. The Lg coda varies across the array.

Figure 4b. Map view of ray bounce points from Dynamic Ray Tracing for event #2. The wave front bends sharply due to steep gradients in the Moho topography.

Figure 4c. Rays for event #2 at take off angles 63-71 degrees for paths to stations TKM and EKS2. Both paths show development of unique focusing and defocusing of rays due to the gradients in the Moho topography.

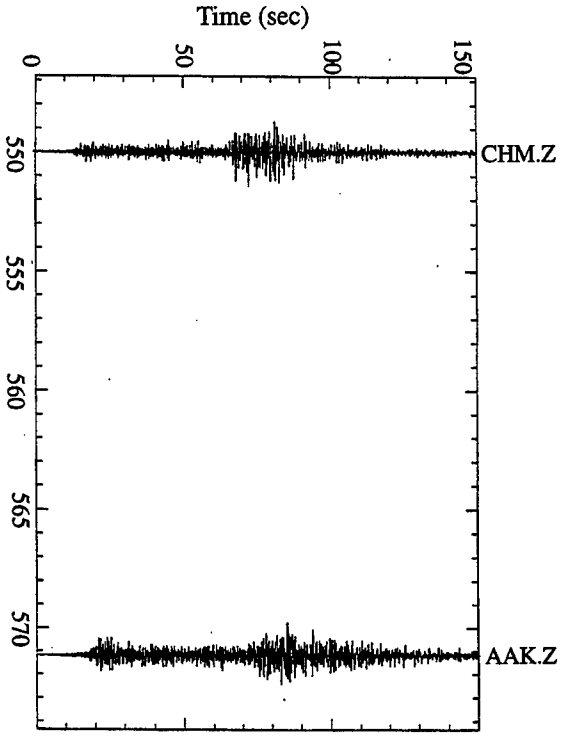


Figure 5a

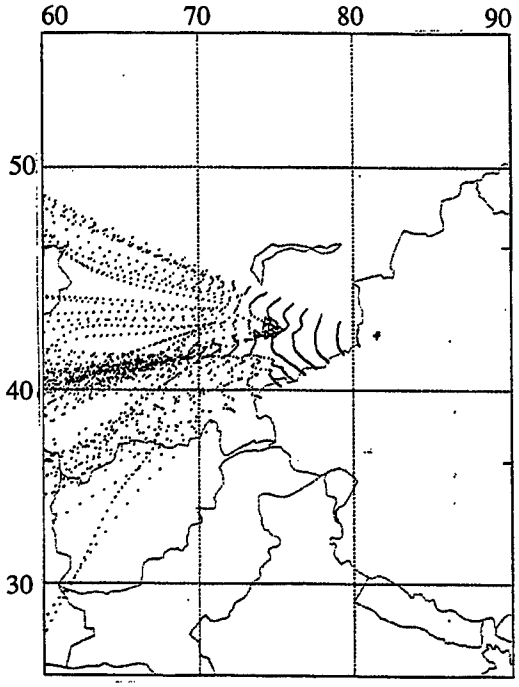


Figure 5b

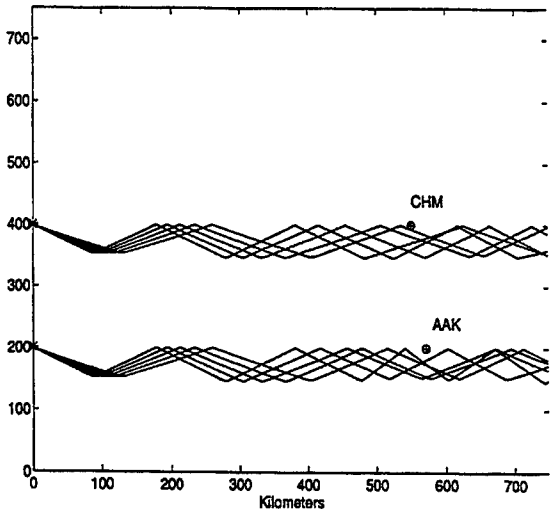


Figure 5c

Figure 5a. Record section plot of event #3, 25 March 92. Data was band passed filtered 1-5 Hz.

Figure 5b. Map view of ray bounce points from Dynamic Ray Tracing for event #3. Wave front appears to remain coherent after 10 bounces.

Figure 5c. Rays for event #3 at take off angles 61-69 degrees for paths to stations AAK and CHM. Due to the relatively flat Moho in the ray path, ray paths are similar out to 500-600km.

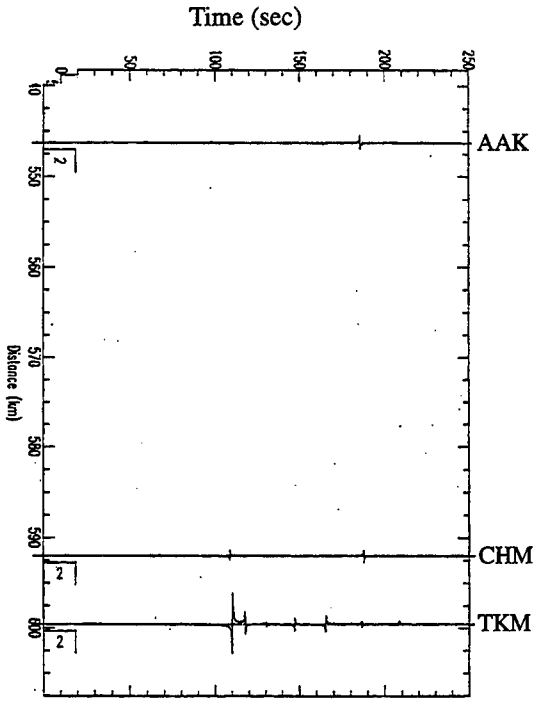


Figure 6a

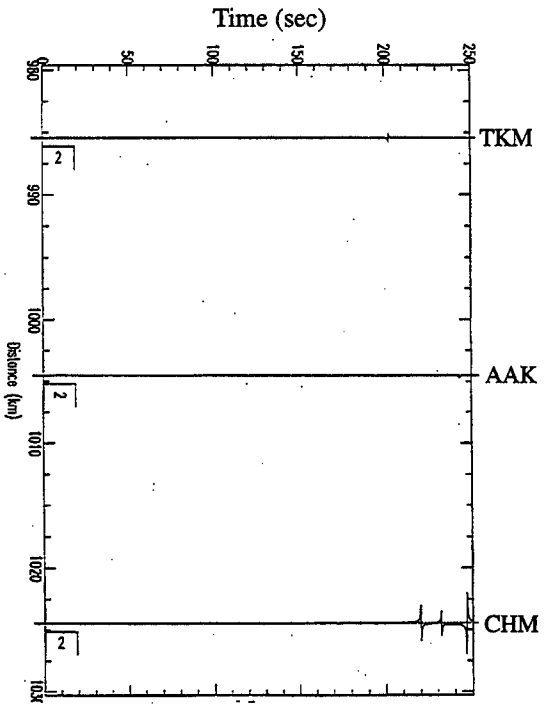


Figure 6b

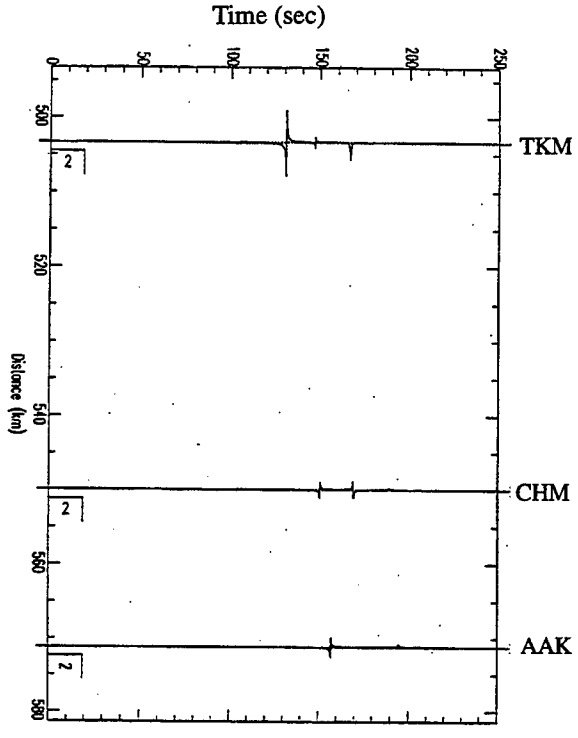


Figure 6c

Figure 6a. Preliminary ray synthetic of event 1 to stations AAK, CHM, and TKM. The synthetics for the different stations show the effects of the Moho topography associated with the Pamir Range.

Figure 6b. Preliminary ray synthetic of event 2 to stations AAK, CHM, and TKM. The synthetics for the different stations show the effects of the Moho topography associated with the Hindu Kush.

Figure 6c. Preliminary ray synthetic of event 3 to stations AAK, CHM, and TKM. The synthetics at AAK and CHM have some of the same arrivals due to the flat Moho east of the array.

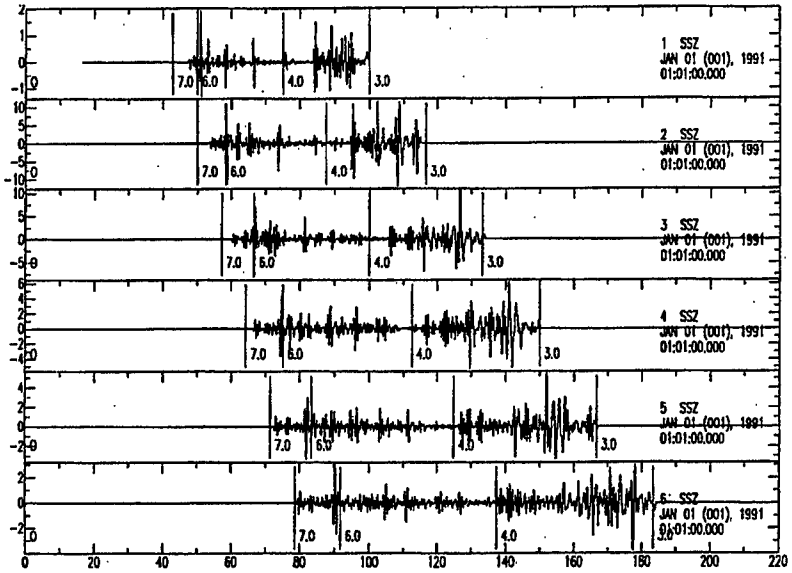


Figure 7a. Synthetic seismograms at ranges 300-550km in steps of 50. Seismograms were calculated with velocity model from the JSPC Kyrgyz Technical Manual. Qs values are 90, 130, 270, 475 for the 0-1, 1-3, 3-16, 16-50km depths.

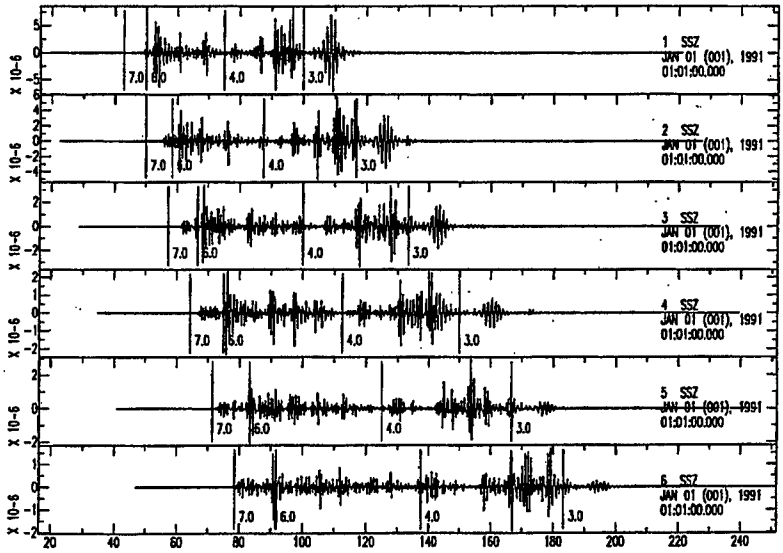


Figure 7b. Synthetic seismograms at ranges 300-550km in steps of 50. Seismograms were calculated with velocity model from the JSPC Kyrgyz Technical Manual. Qs values are 40, 95, 263, 475 for the 0-1, 1-3, 3-16, 16-50km depths.

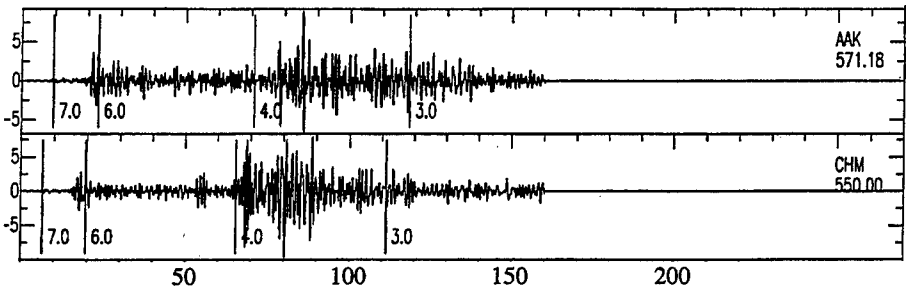


Figure 8a. Seismograms from event #3 recorded at stations AAK and CHM. Band passed 1-5Hz.

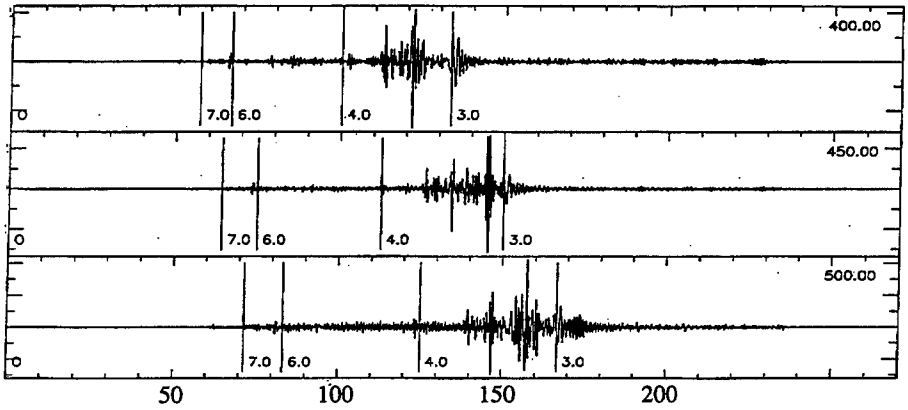


Figure 8b. Preliminary synthetic from CRFL program for event #3.

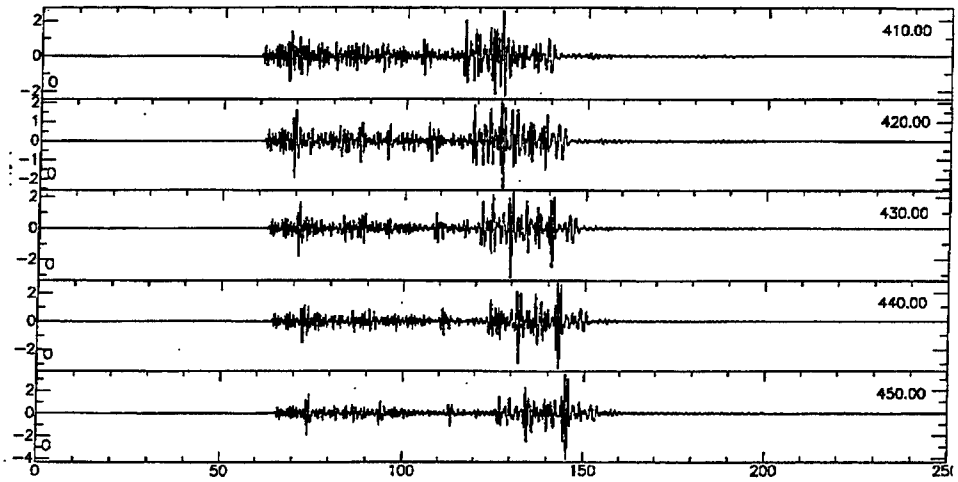


Figure 8c. At ranges of 420-430km there is a good match in coda length, shape and Lg/pn amplitude, indicating need for relocation of source location.

Depth (km)	Vp (km/s)	Qp	Vs (km/s)	Qs	ρ (g/cm^3)
1.0	5.84	140.0	3.37	50.3	2.33
3.0	5.84	140.0	3.37	70.3	2.58
16.0	6.27	400.0	3.62	190.0	2.810
44.0	6.41	400.0	3.7	190.0	2.83

Table 2

Crustal Model east of the Kyrgyz array with Q values found by comparing CRFL synthetics to Event #3.

ANTIMONOTONICITY

IN DRIVEN NONLINEAR ELECTRONIC DIODE CIRCUITS

Paul M. Bennett

Department of Mathematics and Statistics

University of New Mexico

Albuquerque, New Mexico 87131

Final Report for:

Graduate Student Research Program

Nonlinear Optics Center

Phillips Laboratory

Sponsored by:

Air Force Office of Scientific Research

Bolling Air Force Base, Washington DC

September 1995

ANTIMONOTONICITY
IN DRIVEN NONLINEAR ELECTRONIC DIODE CIRCUITS

Paul M. Bennett

Department of Mathematics and Statistics

University of New Mexico

Albuquerque, New Mexico 87131

Abstract

An investigation of antimonotonicity [S. P. Dawson, C. Grebogi, and H. Koçak, Phys. Rev. E 48 1676 (1993)] in a driven nonlinear diode circuit is presented. A geometric model describing antimonotonicity in mathematical models of dynamical systems is given. The model for the nonlinear diode circuit is presented, and it is shown that the behaviour of the model agrees very well with the behaviour of the physical circuit.

ANTIMONOTONICITY IN DRIVEN NONLINEAR ELECTRONIC DIODE CIRCUITS

Paul M. Bennett

I. INTRODUCTION

Period doubling has been studied heavily during the last 15 years as one of the routes to chaotic behaviour in dynamical systems. We can examine it in the logistic family of one-dimensional maps given by

$$f_\lambda(x) = \lambda - x^2,$$

where λ is a scalar parameter. As we vary λ upwards from, say 0, we see that the iterative maps $x_{n+1} = f_\lambda(x_n) = \lambda - x_n^2$ undergo a period-doubling cascade into chaos. See Figure 1(a). In this figure, we see only the creation of periodic orbits at the location of each bifurcation point. In a different logistic family of one-dimensional maps, we see only the destruction of periodic orbits. In this case, the pitchforks open to the left instead of the right. In either case, we are witnessing *monotonicity*. That is, the creation or else the destruction of periodic orbits.

Consider now the cubic family of one-dimensional maps defined by

$$f(x, \alpha) = x^3 + 2.0625x + \alpha.$$

This family generates the iterative maps [3]

$$x_{n+1} = f(x_n, \alpha) = x_n^3 + 2.0625x_n + \alpha.$$

If we vary α from -1.72 to 0.20, we get the diagram pictured in Figure 1(b). Notice that we have the period-doubling in both directions, corresponding to annihilation and creation

of periodic orbits infinitely often in the chaotic region. We refer to this phenomenon as *antimonotonicity* [1–4].

In section II we present a geometric model of antimonotonicity. In section III we present a model of the specific circuit we used to study antimonotonicity. In section IV we present our findings. In section V we present our conclusions.

II. ANTIMONOTONICITY

To give a geometric explanation of antimonotonicity, we refer to Figure 1(c). Let

$$x_{n+1} = f_\alpha(x_n) = \alpha[-x_n^3 + 3(.95)^2 x_n] + 0.099.$$

Critical points of f_α occur when $c_1 = -.95$ and $c_2 = .95$, for every α. In the figure, we have graphs of some of the 13^{th} iterates of f vs. x in the range $[0.945, 0.955]$. The integers k on the right hand side denote the location of selected preimage points of p. The curves labeled by α_i are graphs of $f_{\alpha_i}^{13}$. Notice that for $i = c$ and $i = 1$ we have no contact at all between $f_{\alpha_i}(x)$ and the value at the horizontal line. At $i = 2$ we have a tangential intersection. Notice that the graph of f^{13} hits one of the 5^{th} preimage points of p in two places. Then for all α in $[\alpha_2, \alpha_3]$, the graph of f^{13} takes the 5^{th} preimage value four times. Thus, α_2 is an orbit-creation value, and the tangency there is a *contact-making tangency.* At α_3 two intersection points coalesce into one point and then vanish, and for α in the interval (α_3, α_4), we have only two intersections, so that two 5^{th} preimages have been lost. Thus, α_3 is an orbit-destruction value, and the tangency is *contact-breaking.* In this way, we see the geometry behind the rapid creation and destruction of periodic orbits in the logistic family given above.

To give a more mathematical description of antimonotonicity, we first take the maps of interest to us to be those which are dependent upon a bifurcation parameter, say μ, similarly to those above. Any such map f_μ will be applied iteratively, yielding $x_{n+1} = f_\mu(x_n)$, and

must have the occurence of two critical values c_1 and c_2 within a single interval which is a chaotic attractor for f_μ. We require this situation to persist through a range of values of μ. Finally, we also require the orbits of the critical values c_1 and c_2 to be dense in the attractor with positive Lyapunov exponents.

Fix a value of the bifurcation parameter, say μ_o. It is known that in the chaotic attractor, there exists an unstable point p whose preimages form a dense subset of the attractor [1]. To see this, suppose we have exactly two distinct critical points well within the interior of the chaotic interval for μ. This ensures that the first preiterate of each point x under f consists of exactly three points. The second preiterate will then consist of nine points for each x, the third will have 27, and so on ad infinitum. In the limit, this set is infinite and dense, and is the *stable set* of p. However, the set of images of p is unstable, and belongs to the *unstable set* of p. Notice that both sets intersect at p. This intersection is of fundamental importance, for the reason that [5] under certain restrictions, it is an accumulation point of saddle-node and period-doubling bifurcation values through which periodic orbits are created or destroyed.

We now describe the restriction more fully. The requirement that the forward orbits of the critical points be dense in the attractor implies that at a value μ_1 arbitrarily close to μ_o, an iterate, say the n^{th}, of one of the critical values, call it c, will land on p. That is, $f^n(c, \mu_1) = p$. In this event, the critical point c is on both the stable and unstable sets of p [5]. The number μ_1 will be referred to as a *homoclinic tangency value.* It will be said to be *nondegenerate* if the iterate $f^n(c, \mu_1)$ crosses p with non-zero relative velocity, ie. $\partial f^n(c, \mu_1)/\partial\mu \neq 0$, and if $f^n(x, \mu_1)$ has a quadratic extremum at c. If periodic orbits are created as $f^n(c, \mu)$ crosses p, then μ_1 is referred to as an *orbit-creation* value. In this case, points are added to the preimage of p, and we refer to the tangency as *contact-making.* On the other hand, if periodic orbits are destroyed in the crossing, then μ_1 is an *orbit-annihilation* value. This corresponds to the loss of points in the preimage of p, and in this case, the tangency is called *contact-breaking.* We remark here that the creation and annihilation of

periodic orbits may be accomplished through saddle-node and period-doubling bifurcations.

In [5], it is shown that nondegenerate homoclinic tangency values are the accumulation points of saddle-node and period-doubling bifurcation values through which periodic orbits are either created or destroyed. Thus, as we increase μ, we observe the concurrent creation and destruction of periodic orbits. This phenomenon is known as *antimonotonicity.*

Referring again to Figure 1(c), we see that at α_2 we have a contact-making homoclinic-tangency, after which we have several intersections between the line and graph of $f^{13}(x)$. At this stage, we have four intersections. At α_3 we have a contact-breaking homoclinic-tangency since two intersection points coalesce into one which then disappears, so that two 5^{th} preimages have been lost.

III. THE DIODE RESONATOR REVISITED

We begin our analysis by deriving the equations for a diode resonator following the SPICE derivation modeling the diode [7]. This development follows Tim Newell's Doctoral Dissertation [6]. We then nondimensionalize the resulting system of equations.

The circuit we are studying is a standard diode-inductor-resistor circuit driven by an AC power supply. Both the diagram of the system and a model diagram are given in Fig. 2.

In developing equations governing the circuit, we will initially consider the circuit as a whole, then concentrate our attention on the diode. We will assume throughout that the use of ideal inductors. An ideal inductor assumes no resistance in the inductor of any sort. In reality, an inductor may be made of many windings of wire, and therefore may have a non-trivial resistance from the wire itself. In such a case, we may model the inductor as an ideal inductor and resistor in series. In the model below, the sole effect is to change the resitance in the calculations, and is therefore treated at once with the ideal inductor case.

We note here the use of Kirchoff's current and voltage laws, and relations governing the

various components: Across the resistor, $V_R = I_R \cdot R$, across a capacitor, $I_C = C \cdot \frac{dV_C}{dt}$, and for an inductor we have $V_L = L \cdot \frac{dI_L}{dt}$. We note also that we add parallel capacitances and serial resistances.

An application of Kirchoff's current law to the connections between the diode, inductor, and capacitor quickly yields $I_D = I_R = I_L$, where D denotes the current through the diode, L the current through the inductor, and R the current through the resistor. We will denote the common value by I.

To obtain a more detailed equation involving the diode, we now appeal to [7] for a model governing the behaviour of the diode. The diode is considered to have a small nonlinear resistance, giving the well known current-voltage characteristic of the diode, in parallel with 2 nonlinear capacitances arising due to carrier depletion and finite response time. There is also a small resistance in series with the circuit consisting of the other three components. The corresponding diagram is shown in the position of the diode in Fig. 2a.

Applying the current law shows that $I_{R_s} = I_N + I_{C_d} + I_{C_t}$, where R_s is a small resistance of the diode, and I_N a current through a nonlinear resistance in parallel with the capitances C_d and C_t. It is clear that $I_{R_s} = I$ so that $I = I_N + I_{C_d} + I_{C_t}$.

We now shift our attention to the voltages of the various components in our circuit. Applying Kirchoff's voltage law to the circuit as a whole, we find that $V_0 \sin(\omega t) = V_D + V_L + V_R$, where V_0 is the driving amplitude, V_D the voltage drop across the diode, V_L the drop across the inductor, and V_R the drop across the resistor. Concentrating on the diode itself, we see that $V_N = V_{C_d} = V_{C_t}$. We shall denote this voltage by V. It is clear that $V_D = V_{R_s} + V$.

We now form a system of coupled ordinary differential equations (ODEs) by applying the various component laws above. Thus, $V_L = L\frac{dI_L}{dt}$ and $V_R = I_R \cdot R$ imply that

$$V_0 \sin(\omega t) = V_D + IR + L \cdot \frac{dI}{dt}$$

$$
\begin{aligned}
&= V_{R_s} + V + IR + L \cdot \frac{dI}{dt} \\
&= I_{R_s} R_s + V + IR + L \cdot \frac{dI}{dt} \\
&= V + I(R + R_s) + L \cdot \frac{dI}{dt}.
\end{aligned}
$$

Looking at the diode, we see that $I_{C_d} = C_d \cdot \frac{dV_{C_d}}{dt} = C_d \cdot \frac{dV}{dt}$ and that $I_{C_t} = C_t \cdot \frac{dV_{C_t}}{dt} = C_t \cdot \frac{dV}{dt}$. Adding these two equations together yields the result

$$
I_{C_d} + I_{C_t} = C_d \cdot \frac{dV}{dt} + C_t \cdot \frac{dV}{dt} = (C_d + C_t) \cdot \frac{dV}{dt}.
$$

All that remains is to write down equations governing the nonlinear capacitances and resistance in the diode itself.

For the nonlinear resistance, we have the well known current-voltage characteristic, given by the Shockley equation $I_N = I_s \left(e^{qV/nkT} - 1 \right)$. In this equation, q is the electronic charge, and k is the Boltzmann constant, so kT/q is the thermal voltage. n is the diode's emission coefficient.

For the nonlinear capacitances C_d and C_t, we appeal to the SPICE model [7] to find that $C_d = C_0 e^{qV/nkT}$, while under large forward biasing $C_t = C_b \left(\frac{b_1 + \frac{mV}{V_J}}{b_2} \right)$, but under reverse and small forward biasing, $C_t = C_b (1 - V/V_J)^{-m}$. That is,

$$
C_t = \begin{cases} C_b (1 - V/V_J)^{-m} & \text{if } V < V_J/2 \\ C_b \left(\frac{b_1 + \frac{mV}{V_J}}{b_2} \right) & \text{if } V \geq V_J/2. \end{cases}
$$

In the first equation, C_d is the capacitance arising due to the nonzero response time of the mobile charge to the changing field. The charges build up according to the model $Q = \tau I$, where the constant τ describes the time required for the majority carriers to cross the p- or n-type material. This of course assumes the use of a pn-diode. It is the variance of the charge Q with respect to the voltage that gives the capacitance C_0. C_b is the zero voltage bias capacitance, m is a grading coefficient referring to the variation of the doping concentration across the pn-junction, and V_J is the junction potential. b_1 and b_2 are parameters ensuring

continuity of the capacitance and are given by $b_1 = (1-m)/2$ and $b_2 = (1/2)^{m+1}$. The formula for b_1 is valid as long as the forward bias is not too large.

Summarizing, we have established the equations $V_0 \sin \omega t = V + L \cdot \frac{dI}{dt} + I(R + R_s)$, $I = I_N + I_{C_d} + I_{C_t}$, $(C_d + C_t)\frac{dV}{dt} = I_{C_d} + I_{C_t}$, and $I_N = I_s(e^{qV/nkT} - 1)$. We now carry out a nondimensionalization of these equations.

The first substitution is to take $\tau = \omega t$. Then the differentiation operators in τ and t become related by $\frac{d}{dt} = \frac{d}{d\tau}\frac{d\tau}{dt} = \omega\frac{d}{d\tau}$, so that the two differential equations become $V_0 \sin(\tau) = V_d + L\omega\frac{dI}{d\tau} + I(R + R_s)$ and $(C_d + C_t)\omega\frac{dV}{d\tau} = I_{C_d} + I_{C_t}$.

The next simplification occurs in the equations for the nonlinear capacitances. Comparing their formulas gives an easy substitution to make, namely $\mathcal{V} = V_d/V_J$. We then obtain

$$C_t = \begin{cases} C_b(1-\mathcal{V})^{-m} & \text{if } \mathcal{V} < 1/2 \\ C_b(b_1 + m\mathcal{V})/b_2 & \text{if } \mathcal{V} \geq 1/2, \end{cases}$$

while for C_d we get $C_d = C_0 e^{(eV_J/nkT)\frac{V_d}{V_J}} = C_0 e^{(eV_J/nkT)\mathcal{V}}$, and for I_N we have $I_N = I_s\left[e^{(eV_J/nkT)\mathcal{V}} - 1\right]$. The differential equation for the voltage becomes $(C_d + C_t)\omega V_J\frac{d\mathcal{V}}{d\tau} = I_{C_d} + I_{C_t}$. For the current, we get the equation $V_0 \sin(\tau) = V_J\mathcal{V} + L\omega\frac{dI}{d\tau} + I(R + R_s)$.

At this point we write $eV_J/nkT = \alpha$, simplifying the exponents considerably and yielding $C_d = C_0 e^{\alpha\mathcal{V}}$ and $I_N = I_s\left(e^{\alpha\mathcal{V}} - 1\right)$.

It is now apparent that we should add and subtract I_N in the voltage equation to obtain

$$(C_d + C_t)\omega V_J\frac{d\mathcal{V}}{d\tau} = I_N + I_{C_d} + I_{C_t} - I_N = I - I_s\left(e^{\alpha\mathcal{V}} - 1\right).$$

A subsequent division of all quantities in this equation by $(C_d + C_t)\omega V_J$ yields

$$\frac{d\mathcal{V}}{d\tau} = \frac{I - I_s\left(e^{\alpha\mathcal{V}} - 1\right)}{(C_d + C_t)\omega V_J}.$$

Concentrating now on the first equation, we subtract all non-derivative terms from the RHS and write in reverse order $L\omega\frac{dI}{d\tau} = -V_J\mathcal{V} - I(R + R_s) + V_0 \sin(\tau)$. Dividing through by

V_J yields $\frac{L\omega}{V_J}\frac{dI}{d\tau} = -\mathcal{V} - \frac{R+R_s}{V_J}I + \frac{V_0}{V_J}\sin(\tau)$. Write $\mathcal{I} = \frac{L\omega}{V_J}I$. Then $\frac{d\mathcal{I}}{d\tau} = \frac{L\omega}{V_J}\frac{dI}{d\tau}$, so we find that

$$\frac{d\mathcal{I}}{d\tau} = -\mathcal{V} - \frac{R+R_s}{L\omega}\frac{L\omega}{V_J}I + \frac{V_0}{V_J}\sin(\tau) = -\mathcal{V} - \frac{R+R_s}{L\omega}\mathcal{I} + \frac{V_0}{V_J}\sin(\tau).$$

Simultaneously, the voltage equation becomes

$$\frac{d\mathcal{V}}{d\tau} = \frac{I - I_s\left(e^{\alpha\mathcal{V}} - 1\right)}{(C_d + C_t)\omega V_J} = \frac{\frac{\omega L}{V_J}\left[I - I_s\left(e^{\alpha\mathcal{V}} - 1\right)\right]}{\omega L \cdot \omega(C_d + C_t)} = \frac{\mathcal{I} - \frac{\omega L I_s}{V_J}\left(e^{\alpha\mathcal{V}} - 1\right)}{\omega L \cdot \omega(C_d + C_t)}.$$

The expression for the current is somewhat complicated in its apparent dependence upon parameters, so we will substitute $V_1 = V_0/V_J$ and $\beta = L\omega/(R + R_s)$ to obtain

$$\frac{d\mathcal{I}}{d\tau} = -\mathcal{V} - \frac{\mathcal{I}}{\beta} + V_1\sin(\tau).$$

This is the dimensionless equation we will use for the current.

At the same time, however, we must be sure to check the voltage equation for any such substitutions. We have

$$\frac{d\mathcal{V}}{d\tau} = \frac{\mathcal{I} - \frac{\omega L}{R+R_s}\frac{I_s(R+R_s)}{V_J}\left(e^{\alpha\mathcal{V}} - 1\right)}{\frac{\omega L}{R+R_s}\omega(R + R_s)(C_d + C_t)} = \frac{\mathcal{I} - \frac{\beta I_s(R+R_s)}{V_J}\left(e^{\alpha\mathcal{V}} - 1\right)}{\beta\omega(R + R_s)(C_d + C_t)}.$$

We can simplify this expression by letting $\gamma = \beta I_s(R + R_s)/V_J$, and then simplify the voltage equation further by writing in the equations for the capacitances in the denominator. We have $C_d + C_t = C_0 e^{\alpha\mathcal{V}} + C_b \cdot var = C_0\left(e^{\alpha\mathcal{V}} + C_b/C_0 \cdot var\right)$, where

$$var = \begin{cases} (1 - \mathcal{V})^{-m} & \text{if } \mathcal{V} < 1/2 \\ (b_1 + m\mathcal{V})/b_2 & \text{if } \mathcal{V} \geq 1/2. \end{cases}$$

Write $C_b/C_0 = C_2$ so that $C_d + C_t = C_0\left(e^{\alpha\mathcal{V}} + c_2 \cdot var\right)$. Then

$$\frac{d\mathcal{V}}{d\tau} = \frac{\mathcal{I} - \gamma\left(e^{\alpha\mathcal{V}} - 1\right)}{\frac{\omega L}{R+R_s}\omega(R + R_s)(C_d + C_t)} = \frac{\mathcal{I} - \frac{\beta I_s(R+R_s)}{V_J}\left(e^{\alpha\mathcal{V}} - 1\right)}{\beta\omega(R + R_s)(C_d + C_t)}.$$

Then

$$\frac{d\mathcal{V}}{d\tau} = \frac{\mathcal{I} - \gamma\left(e^{\alpha\mathcal{V}} - 1\right)}{\beta\omega(R + R_s)C_0\left(e^{\alpha\mathcal{V}} + c_2 \cdot var\right)} = \frac{\mathcal{I} - \gamma\left(e^{\alpha\mathcal{V}} - 1\right)}{c_1\left(e^{\alpha\mathcal{V}} + c_2 \cdot var\right)},$$

so that we may write

$$\frac{d\mathcal{V}}{d\tau} = G(\mathcal{I},\mathcal{V}) = \begin{cases} \frac{\mathcal{I}-\gamma\left(e^{\alpha\mathcal{V}}-1\right)}{c_1\left(e^{\alpha\mathcal{V}}+c_2(1-\mathcal{V})^{-m}\right)} & \text{if } \mathcal{V} < 1/2 \\ \frac{\mathcal{I}-\gamma\left(e^{\alpha\mathcal{V}}-1\right)}{c_1\left(e^{\alpha\mathcal{V}}+c_2\left(\frac{b_2+m\mathcal{V}}{b_1}\right)\right)} & \text{if } \mathcal{V} \geq 1/2 \end{cases}.$$

This equation is the dimensionless equation we will use for the voltage.

We have then the following system of dimensionless equations:

$$\frac{d\mathcal{I}}{d\tau} = \mathcal{V} - \frac{\mathcal{I}}{\beta} + V_1 \sin(\tau)$$

$$\frac{d\mathcal{V}}{d\tau} = \begin{cases} \frac{\mathcal{I}-\gamma\left(e^{\alpha\mathcal{V}}-1\right)}{c_1\left(e^{\alpha\mathcal{V}}+c_2(1-\mathcal{V})^{-m}\right)} & \text{if } \mathcal{V} < 1/2 \\ \frac{\mathcal{I}-\gamma\left(e^{\alpha\mathcal{V}}-1\right)}{c_1\left(e^{\alpha\mathcal{V}}+c_2\left(\frac{b_2+m\mathcal{V}}{b_1}\right)\right)} & \text{if } \mathcal{V} \geq 1/2 \end{cases}$$

Notice that these equations represent an oscillator with nonlinear damping and nonlinear driving. To see this, note that we can differentiate the current equation, then insert the voltage equation and obtain the form of the driven oscillator with nonlinear damping.

In the above analysis, we have assumed the driving to be given by $V_0 \sin(\omega t)$. That is, the driving has DC-offset equal to zero. We can model a non-zero DC-offset by writing the driving function as $W + V_0 \sin(\omega t)$, where W is the offset voltage. We then have a division by V_J in the nondimensionalization, which will affect the new driving function by

$$\frac{W}{V_J} + \frac{V_0}{V_J}\sin(\tau) = V_2 + V_1 \sin(\tau),$$

where it is apparent that we have a new parameter to vary, namely V_2.

IV. RESULTS

We present now the results of experiments run physically and numerically during the latter part of June, July, and early August, 1995. We used a 1N4004 diode for the experi-

ments.

TABLE 1: ASSIGNED VALUES			
R	330Ω	m	0.38
R_L	368Ω	n	1.7
R_s	0.0422Ω	α	7.692
L	47mH	β	46.54
C_b	52.1pF	γ	0.2752
C_0	668pF	b_1	0.38422
V_J	0.34V	b_2	0.31
I_s	5.86μA	c_1	5.16
ω	$2\pi \cdot 110$kHz	c_2	0.078

TABLE 2: SCALED PARAMETERS	
PARAMETER	RANGE
V_1	0 to 20
V_2	−10 to 10
T	1 to 350K
f	5 to 300kHz

Notice in Figure 3 that we have a period doubling cascade into chaos in the filled bubbles, then a rapid transition back out to a period two orbit, then a subsequent descent into chaos, a large period 3 window, and finally a descent out of chaos back to a limit cycle. The rapid transition around -2.3 is actually a sudden switch to a coexisting period-2 attractor that subsequent numerical investigations revealed. In particular, notice the shapes of the two lobes are strikingly similar to the shapes of the lobes of Figure 1(b).

Figure 4 shows first return maps calculated for the diagram in Figure 3. The return maps have very good agreement with the return maps in the corresponding regions in the experimental data. Notice the unusual shape of the chaotic attractor taken in the region of the filled bubbles. Notice also the foliations in all of the first return maps, revealing the existence of higher dimensional attractors.

Figure 5 is a bifurcation diagram made experimentally. Notice the close agreement between this figure and the numerical experiment pictured in Fig. 3 using the model of Section III.

V. CONCLUSIONS

We have studied the deceptively simple driven nonlinear diode oscillator and found that it exhibits some very complicated nonlinear phenomena, such as period-doubling cascades into and out of chaos, coexisting attractors, and the subtle appearance of antimonotonicity. The numerical work strongly mirrors the experimental data, validating the model given in [7] and developed by Newell [6]. We use this as justification in searching for antimonotonicity numerically in detail small enough to be obscured by experimental noise.

VI. ACKNOWLEDGEMENTS

P.M.B. wishes to thank the Air Force Office of Scientific Research for a summer fellowship, and Dr. Vasilios Kovanis and Dr. Timothy Newell for their valuable assistance.

REFERENCES

[1] S. Dawson, C. Grebogi, and H. Kocak, *Geometric mechanism for antimonotonicity in scalar maps with two critical points*, Phys. Rev. E., **48**, (1993) pp 1676-1682.

[2] I. Kan, H. Kocak, and J. Yorke, *Antimonotonicity: Concurrent creation and annihilation of periodic orbits*, Ann. Math., **136** (1992), 219-252.

[3] S. Dawson and C. Grebogi, *Cubic Maps as Models of Two-dimensional Antimonotonicity*, Chaos, Solitons and Fractals, **1**, pp. 137-144, 1991.

[4] S. Dawson, C. Grebogi, J. Yorke, I. Kan, and H. Koçak, *Antimonotonicity: inevitable reversals of period-doubling cascades*, Phys. Lett. A, **162** (1992) pp 249-254.

[5] R. Devaney, *An Introduction to Chaotic Dynamical Systems*, (Benjamin/Cummings, Menlo Park, CA 1986), pg 121.

[6] T. Newell, *Experimental Synchronization of Chaotic Attractors Using Control*, Doctoral Thesis, University of North Texas, Denton, TX, 1994.

[7] ed. by A. Antognetti and G. Massobrio, *Semiconductor Device Modelling with SPICE*, (McGraw-Hill, Inc., 1988), Ch. 1.

FIGURES

FIG. 1. (a). Bifurcation diagram for the family $f_\lambda(x) = \lambda - x^2$. Notice the pitchforks open to the right in every case. See [2]. (b). Bifurcation diagram for the cubic family $x_{n+1} = x_n^3 - 3x_c^2 x_n - \alpha$, with $3x_c^2 = 2.0625$ and $\alpha \in [-1.72, 0.20]$. See [3]. (c). Graphs of several 13^{th} iterates of $f_\alpha(x) = \alpha(-x^3 + 3(0.95)^2 x) + 0.099$ in a neighborhood of the critical point $c = 0.95$. Here $\alpha_c = 1.0194140$, $\alpha_1 = 1.0194050$, $\alpha_2 = 1.1094110$, $\alpha_3 = 1.0194140$, and $\alpha_4 = 1.0194160$. See [1].

FIG. 2. The diagram representing the chosen circuit. Note the replacement of the diode symbol with the array of a nonlinear resistance in parallel with two nonlinear capacitances. The array is in series with a small resistance. See [6,7].

FIG. 3. Bifurcation diagram of the above differential system. It was run in floating-point precision on a SPARC 10. There is a very close similarity to the bifurcation diagrams obtained experimentally. The bifurcation parameter is the offset voltage in dimensionless units, which was varied downward from -0.2 to -4.0 in 210 steps. At each step, we collected 40 peaks in a fourth-order Runge-Kutta integrator, collecting data after the 3750^{th} timestep, for 250 timesteps. We broke up the 4000 timesteps into 32768 intervals, each of which contained 20 Runge-Kutta iterations.

FIG. 4. First return maps for the data graphed in Fig. 3. To obtain these maps, we collected peaks between timesteps 2000 and 10000. We used a Runge-Kutta integration routine, breaking the 10000 timesteps up into 32768 intervals, with 20 integrations per interval. (a). The dimensionless offset voltage was held fixed at -2.4. (b). The offset voltage was held at -1.0.

FIG. 5. Bifurcation diagram made experimentally by Dr. Timothy Newell and Rebecca Kalmus by sweeping the offset voltage through the range from -0.2 to roughly 0.35. Notice the twin lobes joined together between -0.1 and 0.0, the rapid jumps to the right of the lobes, then another period-doubling cascade into chaos, a period-3 window, and another cascade, eventually reaching the limit cycle at about 0.3. Compare this behaviour to the behaviour exhibitted in Fig. 3.

Figure 1(a).

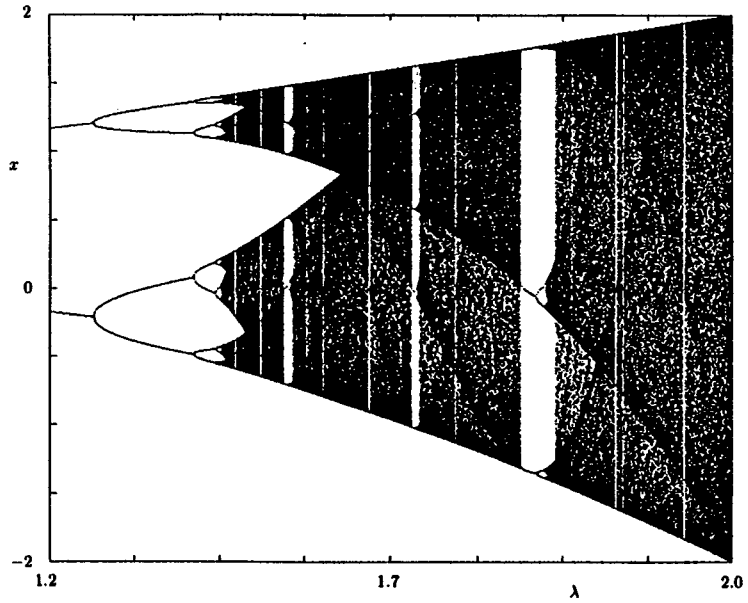


Figure 1(b).

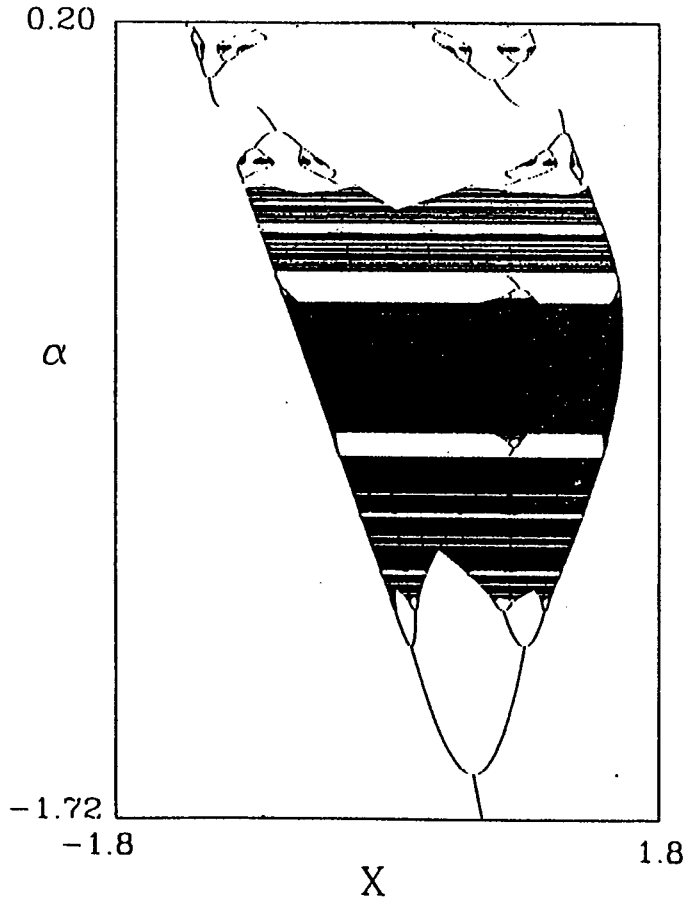


Figure 1(c).

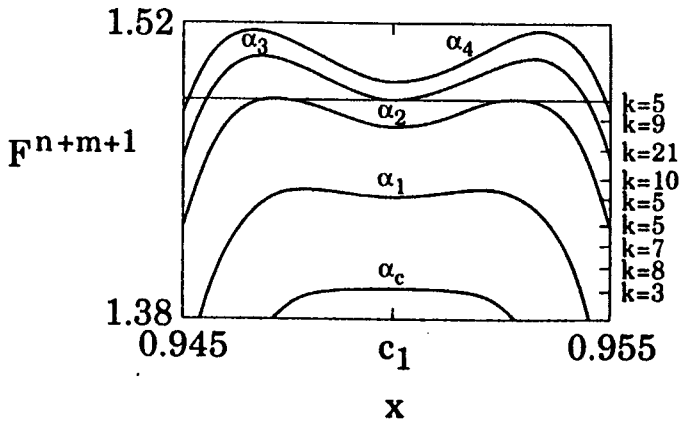
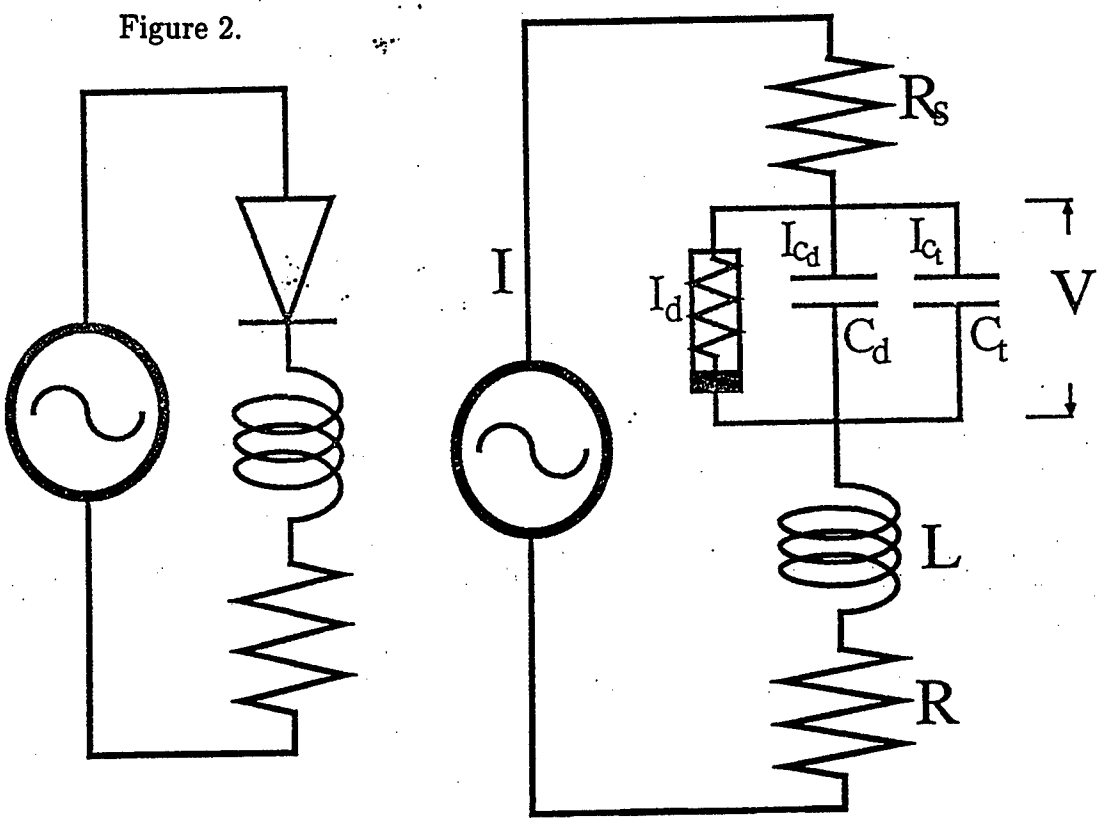


Figure 2(a).

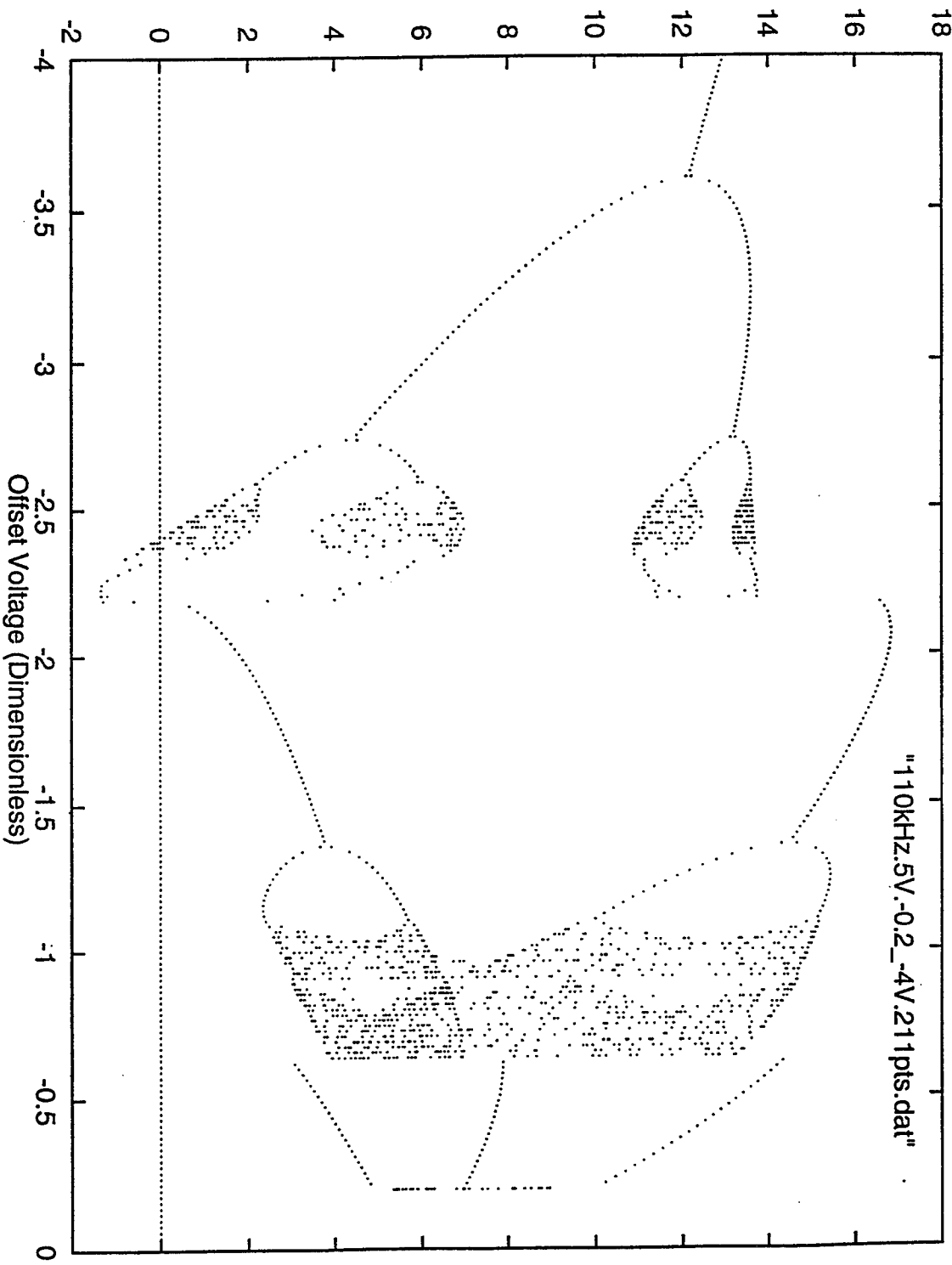
Figure 2.



Diode Resonator

Modelled Diode Resonator

Figure 3.


Voltage Peaks vs. Offset Voltage
"110kHz.5V.-0.2_-4V.211pts.dat"
Offset Voltage (Dimensionless)
18
16
14
12
10
8
6
4
2
0
-2
-4
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0

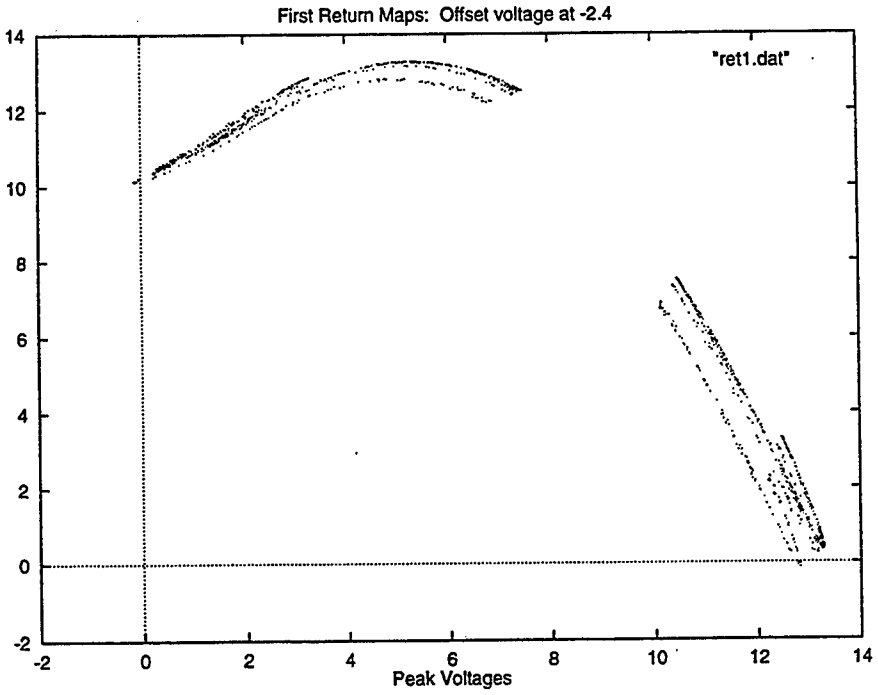


Figure 4(a).

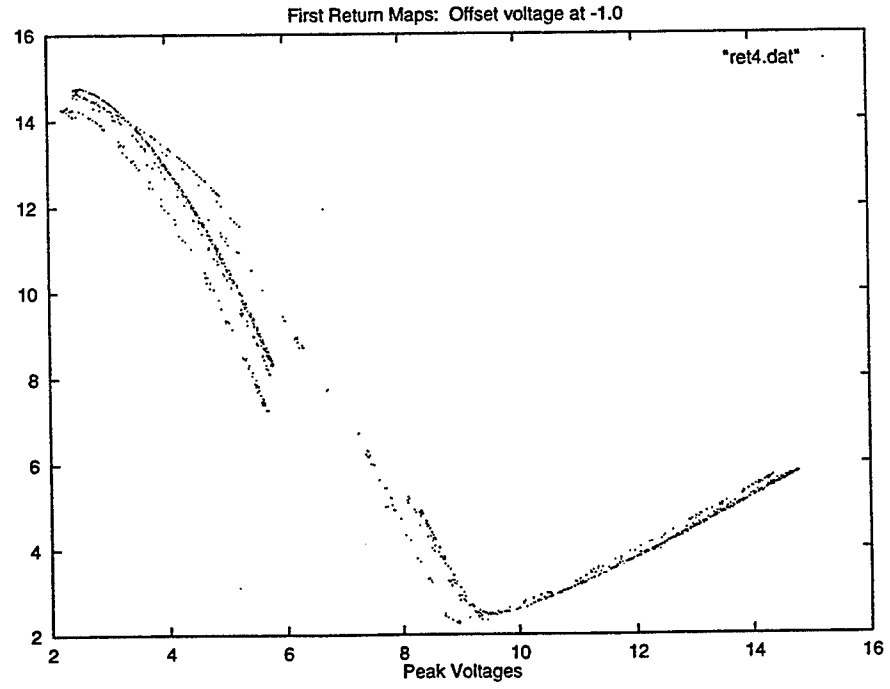
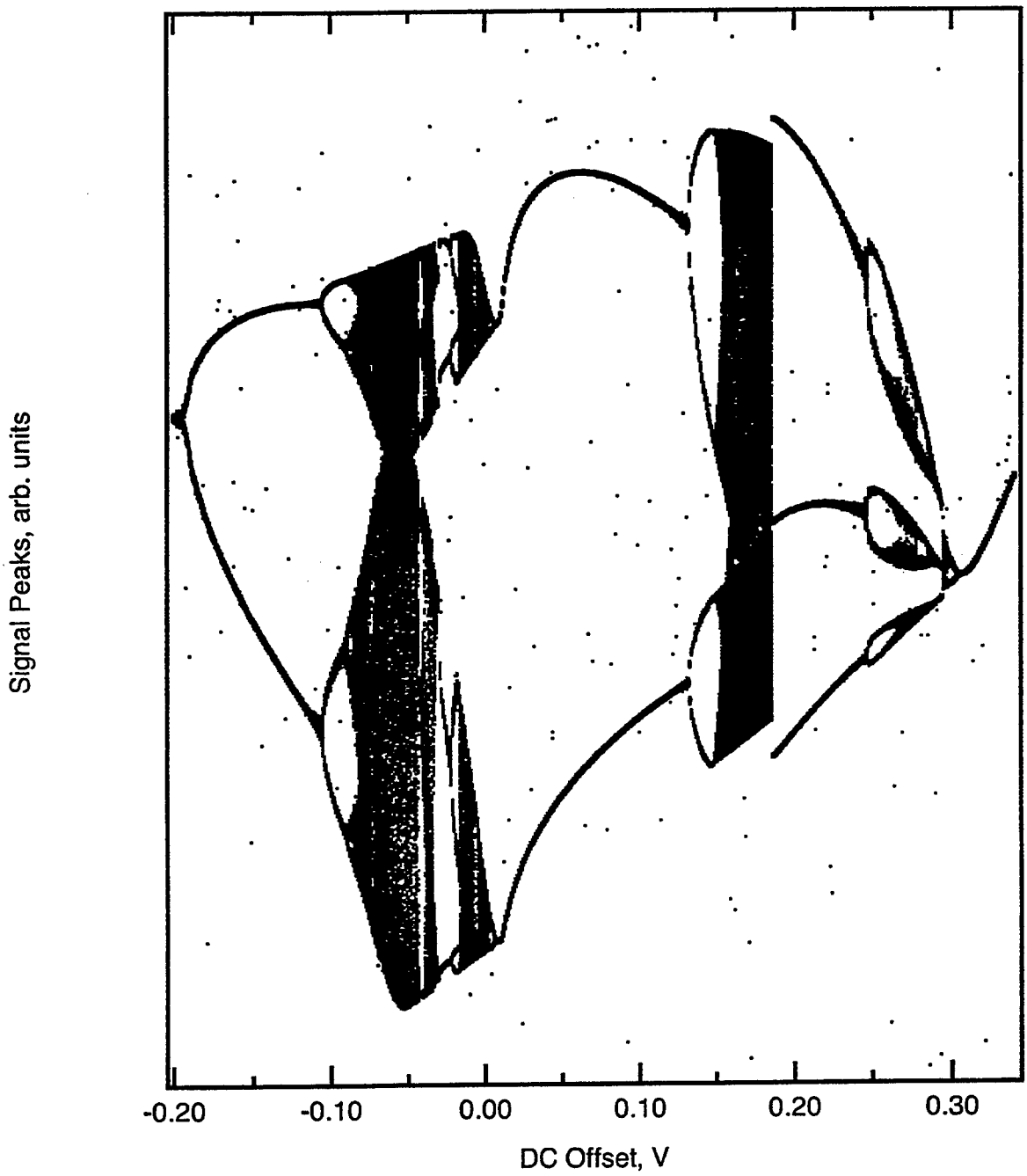


Figure 4(b).

Figure 5.



Associate did not participate in program.

Associate did not participate in program.

Analysis To Determine The Quality Factor Of A Celestron-8 Telescope

Ronald R. DeLyser
Assistant Professor of Electrical Engineering
Department of Engineering

Peyman Ensaf
Graduate Student
Department of Engineering

University of Denver
2390 S. York Street
Denver, CO 80208-0177

Final Report for:
Summer Faculty Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.
And
Phillips Laboratory, WSA
Kirtland AFB, N.M.

September, 1995

Analysis To Determine The Quality Factor Of A Celestron-8 Telescope

Ronald R. DeLyser
Assistant Professor of Electrical Engineering
Peyman Ensaf
Graduate Student
Department of Engineering
University of Denver

Abstract

Due to geometrical similarities between the Celestron-8 telescope and some of the space telescopes, calculating near fields within the Celestron-8 telescope was important for survivability/vulnerability assessments of space telescopes. In this analysis our goal is to use the software package CARLOS-3D [2], to calculate the near fields within the cavity of the Celestron-8 telescope and compare the results obtained using two different discretizations. This report provides some of the formulations and approaches used for this type of assessment along with the errors in results and some suggested improvements.

Overview

Our initial contact with this project was on February 1, 1995, resulting from AFOSR contract F49620-C-0063. The objective of this contract was to perform numerical analysis on a Celestron-8 telescope (Figure 1). We were asked to continue these analysis for the duration of our summer 1995 tour at the Satellite Assessment Center (SAC), Phillips Laboratory.

Our initial approach was to model the telescope's outer most shell as an open ended cylinder with an end cap at the top. At the top portion of the cylinder, there is an aperture for the eyepiece. Considering the aperture to be very small compared to the larger wavelengths, we decided to model the top end of the telescope as an end cap. We modeled the resulting open ended cylinder using Mathematica [1], and analyzed that model using CARLOS-3D [2]. As a starting point, we analyzed the open cylinder model over a frequency range of 0.7 GHz - 3.0 GHz in order to avoid a computationally extensive problem. Once the data for the open cylinder model was analyzed, we proceeded with creating the model which included the interior geometry of the telescope. The interior geometry, consisted of two interior cylinders. These interior cylinders (Figure 1) are at the center of the aperture, where a convex reflecting mirror resides (the short plastic cylinder) and at the eyepiece (the long metal cylinder). The two reflecting mirrors are not shown in Figure 1. The CARLOS model (Figure 2) was analyzed for a frequency range of 0.5 GHz up to 3.0 GHz. Each of the models, the open cylinder and that shown in Figure 2, are discretized with triangular cells, each having sides $\lambda/10$ in length corresponding to a frequency of 1.0 GHz. The models are illuminated at incidence angles θ and ϕ (see figure 2) both equaling 45°.

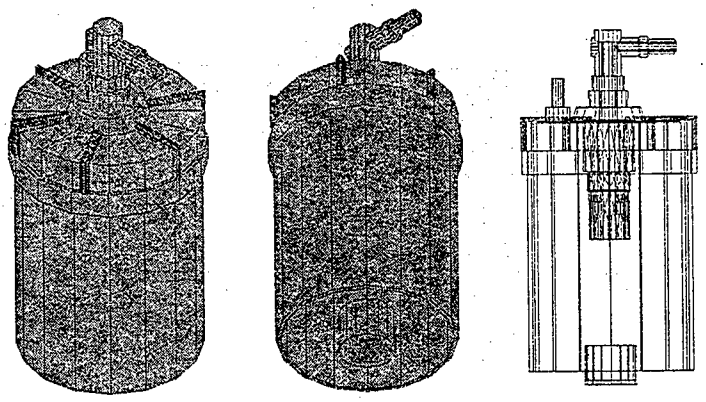


Figure 1. Celestron-8 telescope.

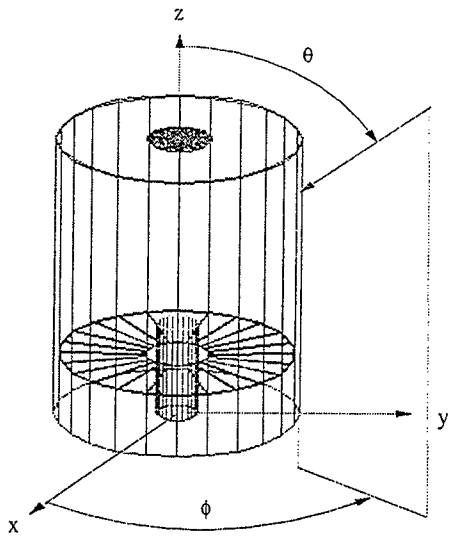


Figure 2. CARLOS model for the telescope.

Parallel to this analysis using CARLOS, we used another software package called GEMACS [3]. GEMACS offers three techniques which can be used to analyze a geometry, Geometrical Theory of Diffraction (GTD) and Method of Moments (MOM) to analyze the exterior of a geometry and the Finite Difference method (FD) to analyze the interior regions. We are interested in comparing the near field results obtained using GEMACS's MOM to the near field results obtained using MOM implemented in CARLOS. Furthermore, using the FD technique in GEMACS, we are trying to obtain cavity field results which should prove to be numerically more accurate. Due to the problems encountered with CARLOS we were unable to compare results with GEMACS. This is still a topic for further study.

Also of interest is the verification and validation of the results using different discretization techniques. We, therefore, discretized the model in Figure 2, to $\lambda/5$ at 1.0 GHz. We analyzed this model for a frequency range of 0.5 GHz-3.0 GHz. Comparing the results of the Quality Factor (Q), total energy, and exiting power for the $\lambda/5$ discretized model to those of $\lambda/10$ model, we noticed pronounced differences in the spectrums. These results along with other problems encountered with CARLOS, and methods for rectification are discussed in the following.

Preliminary Observations

The model of Figure 2 was discretized into triangles having sides no longer than $\lambda/10$ in length at 1.0 GHz frequency. The interior of this geometry was divided into cubes of side length 1.5 cm ($\lambda/10$ at 2.0 GHz). A total of 3717 grid points were created. These grid points are the location of the near field points calculated by

CARLOS. The model in Figure 2 was analyzed for a frequency range of 0.5 GHz up to 3.0 GHz. Using the fundamental equations describing the electric and magnetic energies along with power exiting the cavity, we calculated Q [4, 5] given by

$$Q = \omega \frac{U}{P} \tag{1}$$

where ω is the operating angular frequency, U is the total energy in the cavity and P is the power exiting the cavity. Figure 3 shows Quality Factor vs. Frequency for the model of Figure 2. Resonances for various frequencies are evident in Figure 3, along with the fact that the ϕ polarization dominates the θ polarization. From Figure 3, it is evident that the θ and ϕ polarizations follow a similar trend throughout the spectra.

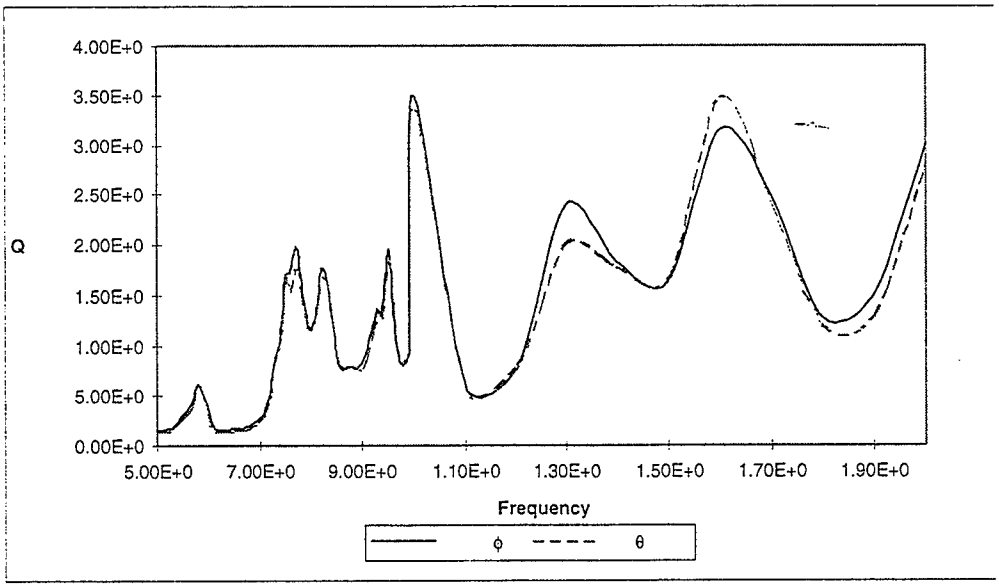


Figure 3. Quality factor vs. frequency, θ and ϕ polarization with $\lambda/10$ discretization.

The next step was to investigate mesh sizes. We, therefore, discretized the model of Figure 2 to $\lambda/5$ at 1.0 GHz, and analyzed that model over frequency range of 0.5 GHz up to 3.0 GHz. Figure 4 shows the spectrum for the $\lambda/5$ and $\lambda/10$ discretization. It is evident from Figure 4 that the two spectra are profoundly different. These discrepancies led us to further investigate meshes using only the open cylinder model, using two different discretization schemes, $\lambda/10$ and $\lambda/20$. We analyzed these models for a frequency range of 0.1 GHz up to 2.0 GHz. Figure 5 shows these results. Obviously the two discretization techniques provide different results.

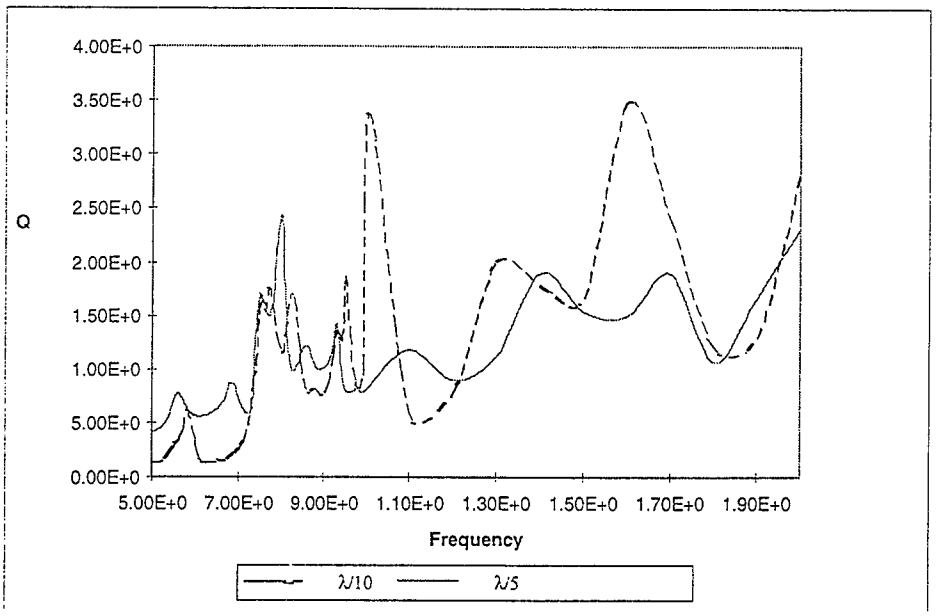


Figure 4. Comparison of quality factors, θ Polarization, λ/5 and λ/10 discretizations.

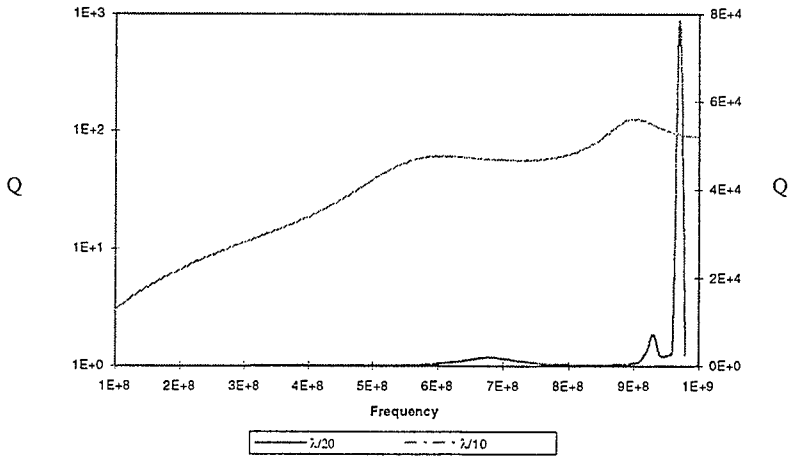


Figure 5. Comparison of the quality factors, ϕ polarization, λ/10 and λ/20 discretizations.

Field Calculation method and results

We first approached the meshing problem by comparing the field quantities for the models (λ/10 and λ/20). Doing this, we realized that there is a large difference in the field values near the cylinder walls than those field values more towards the center of the cylinder. CARLOS calculates the near field quantities using the current at the centroid of each triangular cell. In this scenario, each current source will now act as an infinitesimal dipole. Therefore, the field quantities are calculated using [6]

$$E_z(x,y,z) \propto \sum_{n=1}^{N} \frac{e^{-jkr_n}}{kr_n^3} \cos(\theta) J_n \qquad (2)$$

where k is the propagation constant, J_n is the n^{th} surface current, and r_n is the distance from the n^{th} surface current to the field point. However, the current is distributed across the surface. Therefore the distance r from the field point varies according to where the current is calculated on the surface. The distance r is a very critical quantity since for small r, its variations will have a large effect on the field calculations. Figures 6, 7 and 8 show the percent difference, between the $\lambda/10$ and $\lambda/18$ discretizations, in the electric field quantities at the end cap, the center, and at the aperture. Iterative refinement (see succeeding sections) was implemented for creating the Figures 6, 7 and 8.

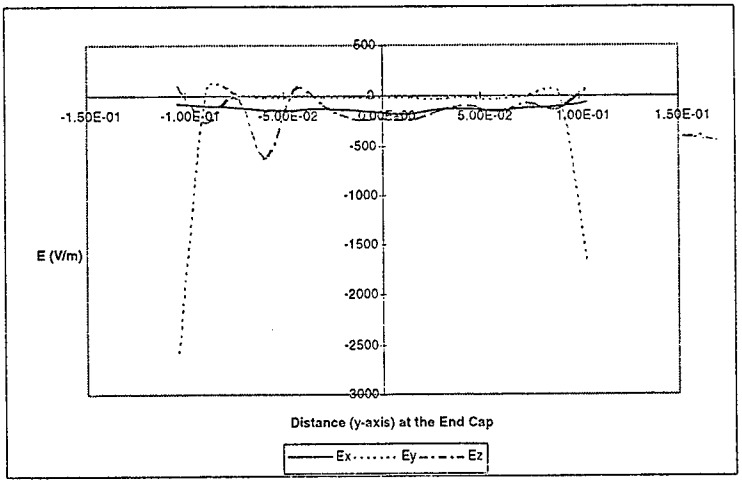


Figure 6. Percent difference in the electric field quantities, ϕ Polarization, θ=90, ϕ=0, 500 MHz, end cap.

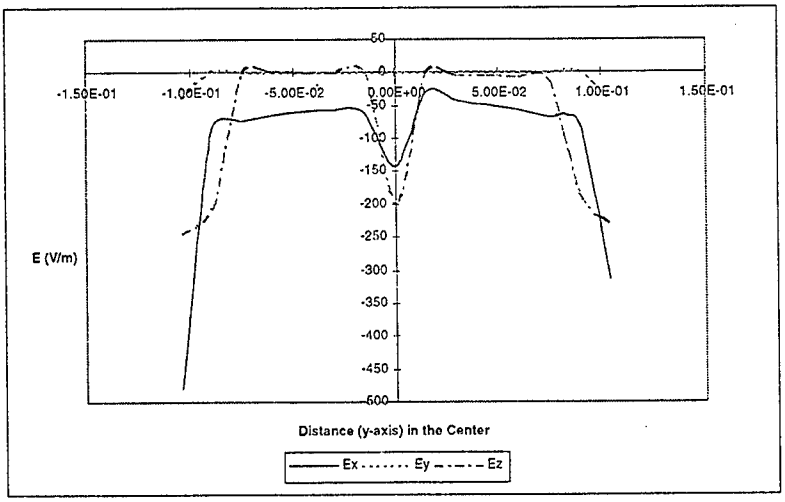


Figure 7. Percent difference in the electric field quantities, ϕ Polarization, θ=90, ϕ=0, 500 MHz, center.

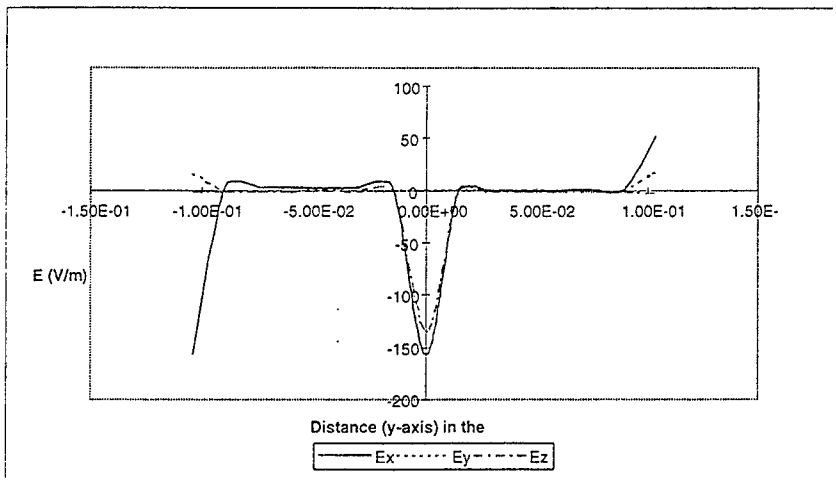


Figure 8. Percent difference in the electric field quantities, ϕ Polarization, θ=90, ϕ=0, 500 MHz, aperture.

It is evident in Figures 6, 7 and 8, that the biggest differences are close to the cavity-boundary. This error is caused by the crude approximation used to calculate the near fields. A more suitable approach would be to use a solution technique which considers the changes in r along with changes in J. This solution is given by [7]

$$E^s(r) = -j\omega\mu \int_{\partial V_s} J_s(r')g(r,r')ds + \nabla \int_{\partial V_s} \frac{1}{j\omega\varepsilon} \nabla' \cdot J_s(r')g(r,r')ds' \tag{3}$$

where

$$g(r,r') = \frac{e^{-jk|r-r'|}}{4\pi|r-r'|} \tag{4}$$

ω is the angular operating frequency, ε and μ are the permittivity and permeability of free space, respectively.

Meshing Methods And Results

In the neighborhood of a conducting edge, electric and magnetic field components behave in a particular manner dependent on the wedge angle ϕ_o (see Figure 9). From [8] the behavior of the surface current density in the vicinity of the edge is

$$\mathbf{E} \propto \rho^{\nu} \sin(\nu\phi) \qquad \mathbf{J} \propto \rho^{\nu-1} \tag{5}$$

where

$$\nu = \frac{\pi}{\phi_o} \tag{6}$$

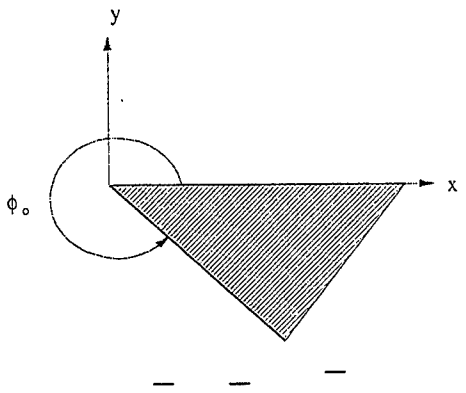


Figure 9. Wedge angle ϕ_o.

where for the top edge of the Figure 2, $\phi_o=2\pi$, and for the bottom edge $\phi_o=3\pi/2$. For our geometry then, at the top edge, $J_\phi \propto \rho^{-1/2}$ and at the bottom edge, $J_\phi \propto \rho^{-2/3}$.

The edge condition is an additional constraint that is needed for a unique solution whenever a geometric singularity, such as a sharp edge, exists. The edge condition states that the energy density in the vicinity of an edge, or any geometrical singularity, must be integrable, that is

$$\mathrm{Re}\int_S \mathbf{E}\times\mathbf{H}^*\cdot d\mathbf{S} \to 0 \tag{7}$$

as the surface S enclosing the edge shrinks to the edge. The edge condition dictates that the edge shall not radiate any energy because it is not a source. In order to accurately model the edge surface currents we used a meshing scheme as shown in Figure 10. Figure 11 shows the z, ρ and ϕ directed surface currents for the meshing scheme of Figure 10.

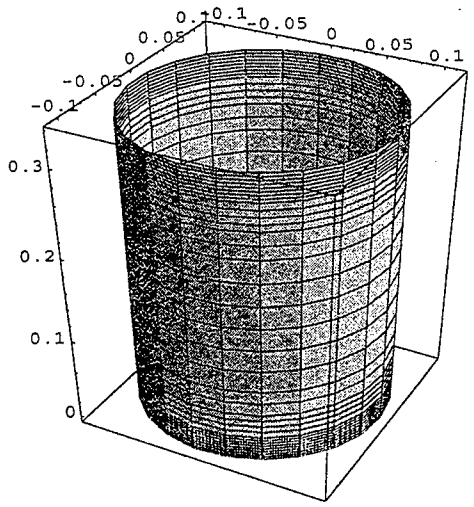


Figure 10. Non-uniform discretization.

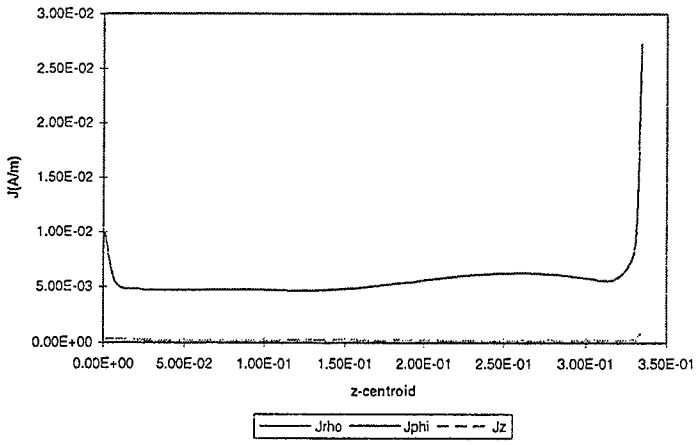


Figure 11. Surface currents for figure 9 at ϕ=10°, 500 MHz, θ=90°, ϕ=0°.

In addition to the accurate modeling of edge currents, we also looked at the problem of generating a mesh for a circular disk. For problems of circular symmetry, the cells near the center of circles become exceedingly narrow. A simple transition to a non-symmetric griding, as shown in Figure 12 would be desirable to maintain reasonable cell sizes. Dr. Pat McDaniel, wrote a program to accomplish this task, and also give a finer meshing for cells closer to the edges. Figure 13 shows some results for this meshing scheme.

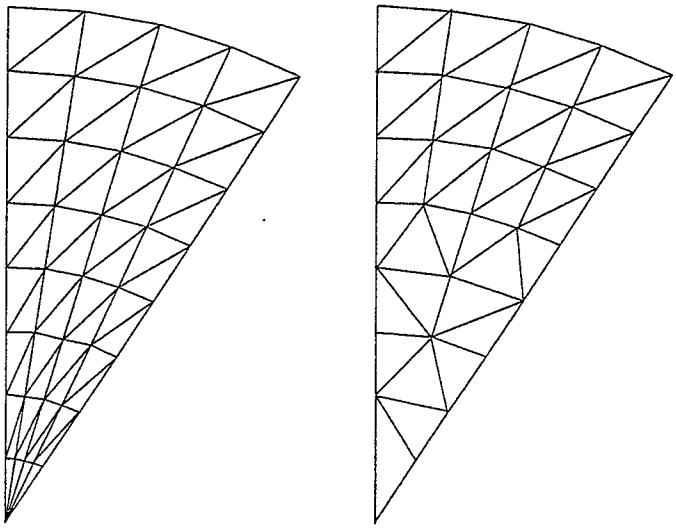


Figure 12. Alternate griding scheme for circular geometry.

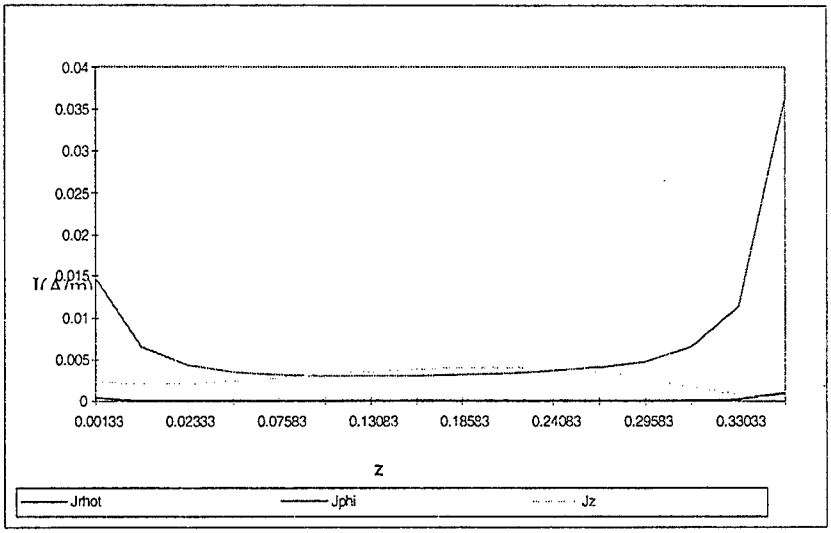


Figure 13. Surface currents using Dr. McDaniel's meshing scheme at ϕ=13°, ϕ Polarization, ϕ=45°, θ=135°

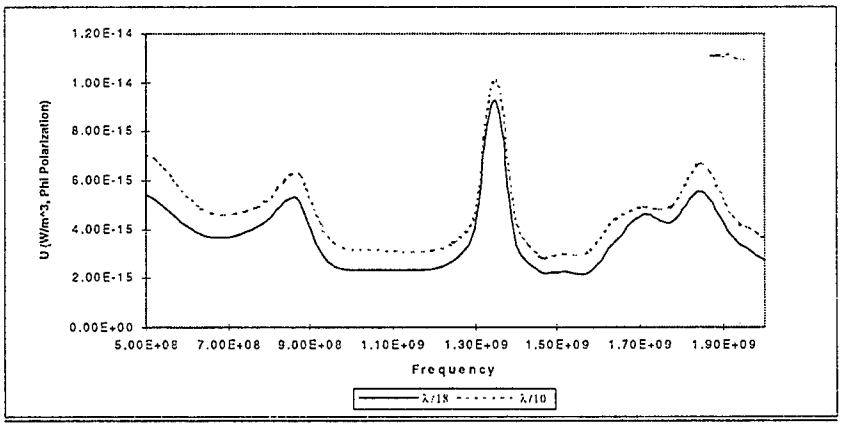


Figure 14. Energy for the two discretizations using Dr. McDaniels meshing scheme, 500 MHz, θ=90°, ϕ=0°.

It is evident from Figure 14 that using the new meshing scheme has made the energy distribution much more compatible for the two discretization. However, there is no improvement in the results for power. This could be due to the fact that we are summing over a much smaller number of grid points in the aperture and thereby not obtaining comparable accuracy in the results. Also, since the fields vary rapidly as the conducting side wall is approached, perhaps a finer field mesh is needed in the aperture.

Accuracy and results for high condition number

Other numerical difficulties resulting from using CARLOS, may be due to the large condition number of the matrix for large models. The condition number of the matrix may be defined as its largest element value divided by its smallest [9]. A matrix is typically considered ill-conditioned if its corresponding inverse matrix is unstable numerically. Taking another point of view, a zero, or near zero determinant indicates that the matrix has, in effect, two rows that are identical or approximately identical. In a system of linear algebraic equations,

a zero determinant (two identical rows) means that two of these equations are either redundant or inconsistent In an approximate-solution method such as MOM, the redundancy and inconsistency in a system of linear equations manifest themselves as numerical instabilities. A large condition number indicates that the matrix may be ill-conditioned. The condition numbers are not always reliable in predicting the stability of a matrix. We have to know how large the condition number must be before we declare the matrix ill-conditioned. A practical rule of thumb is that the matrix is ill-conditioned if the inverse of the condition number approaches the computer's floating-point precision. For an ill-conditioned matrix equation, the iterative methods are generally more suitable. An iterative refinement technique is implemented by Dr. Pat McDaniel and is discussed here.

It is not always easy to obtain greater precision for the solution of a linear set than the precision of your computer's floating-point. In direct methods of solution, round-off errors accumulate, and they are magnified to the extent that the matrix is close to singular. Under these conditions, matrices, which were far from singular, can lose two or three significant figures. Iterative refinement (improvement) is an approach to restore full machine precision [10]. Suppose that a vector $\mathbf{x}$ is the exact solution of a linear set

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \tag{8}$$

However, the exact solution x is not known. What is known, in this case, is a slightly wrong solution $\mathbf{x}+\delta\mathbf{x}$, where $\delta\mathbf{x}$ is the unknown error. Therefore, from equation (8), it is evident that when multiplied by the matrix **A**, the slightly wrong solution gives a product slightly discrepant from the desired right-hand side **b**, such that

$$\mathbf{A} \cdot (\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b} \tag{9}$$

Now, subtracting equation (8) from (9) gives

$$\mathbf{A} \cdot \delta\mathbf{x} = \delta\mathbf{b} \tag{10}$$

Using equation (10) to solve for $\delta\mathbf{b}$ and then substituting that into equation (9) gives

$$\mathbf{A} \cdot \delta\mathbf{x} = \mathbf{A} \cdot (\mathbf{x} + \delta\mathbf{x}) - \mathbf{b} \tag{11}$$

The right side of equation (11) is known, since $\mathbf{x}+\delta\mathbf{x}$ is the wrong solution that requires refinement. Therefore solving equation (11) for $\delta\mathbf{x}$, subtracting it from the wrong solution will provide an improved solution.

Numerical Integration Of The Matrix elements

The integral of a function is approximated by the sum of its functional values at a set of equally spaced points, multiplied by certain aptly chosen weighting coefficients. The idea of Gaussian Quadrature is to give the user the freedom to choose not only the weighting coefficients, but also the location of the abscissas at which the function is to be evaluated, hence they will no longer be equally spaced. Using Gaussian Quadrature formulas, one can arrange the choice of weights and abscissas to make the integral exact for a class of integrands.

Therefore, the weighting function W(x) can be chosen to remove integrable singularities from the desired integral. Hence, one can find a set of weights w_i and abscissas x_i such that the approximation

$$\int_a^b W(x)f(x)dx \approx \sum_{i=1}^{N} w_i f(x_i) \tag{12}$$

is exact if f(x) is a polynomial. CARLOS is a MOM code and hence its primary calculation is surface currents. Figures 15 and 16 present the differences in ϕ directed currents and z directed currents along the wall of the cavity. The primary purpose of these figures is to show the difference in the results using iterative refinement and iterative refinement with improved Gaussian Quadrature.

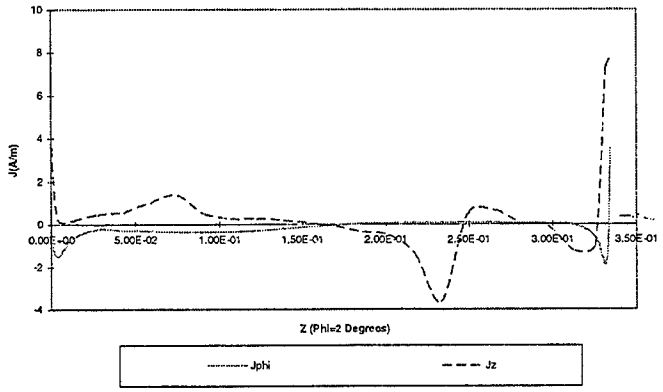


Figure 15. % difference of currents using iterative refinement vs. iterative refinement improved Gaussian. 500 MHz, θ=90°,ϕ=0°.

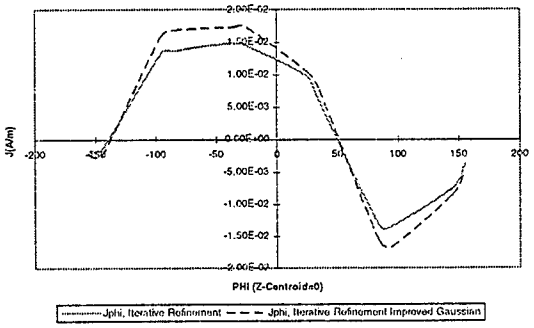


Figure 16. % difference of currents using iterative refinement vs. iterative refinement improved Gaussian, 500 MHz, θ=90°,ϕ=0°.

From Figures 15 and 16, it is evident that using improved Gaussian Quadrature formulas does make an improvement in the results obtained. Figure 15 shows a very small difference in the surface currents for the two discretizations, however, this small difference may reflect a large difference in the near field calculations.

HFSS description and results

The Hewlett Packard (HP) High Frequency Structure Simulator (HFSS) is a full three dimensional finite element simulator. In the simulator, the geometric model is automatically divided into a large number of tetrahedrons and is called the finite element mesh. Within each tetrahedron, the vector field components perpendicular to the edges and tangential to the edges are mathematically modeled using polynomials. The finer the mesh, the more accurate the solution. Of course, the finer the mesh, the greater the demands on computational facilities in terms of memory, disk space and solution time. In order to find a good solution to the problem and minimize computer resources, HFSS uses "an iterative process in which the mesh is automatically refined in critical regions." The simulator first solves the problem using a course initial mesh, refines the mesh in areas of high error density, generates a new solution, etc. until selected S-parameters converge to within a desired limit, or until all computer resources have been exhausted.

The computer that we use to run HFSS is an IBM RS6000 model 570 Powerstation with 128 MBytes of RAM and a 2 GByte hard drive. We have access to all of the RAM, about 200 MBytes of swap space, and (at the time of this study) 340 MBytes of disk space for user files. During the solution process, a convergence menu can be monitored containing memory and file size values. Two matrices are written to disk (user files), the A matrix is the "full matrix" and the U is the "upper triangular matrix". All other memory size values indicated on the convergence menu are RAM usage. As stated above, the iterative process of mesh refinement is continued until convergence, or until the computer resources are exhausted. If the disk space is exhausted (which was the limiting resource for our models), the solution did not converge.

A total of eight projects were created HFSS. The most important ones described here are:

1. telescop.pjt - A full model of the telescope with the exception that the large mirror is flat. We could not get the 2-D cross-section to close using the true radius of curvature. Since the 2D cross-section would not close, the 3D mirror could not be revolved about the z-axis to be created. Thus, a flat mirror with approximate tilt was created. The center of the mirror is at z = 72 mm and the outside is at z = 73.8 mm. This model is excited by a circular waveguide port located at the eye piece.
2. telehf.pjt - A model of the telescope excluding the volume behind the large mirror and with the radiation boundary located 17 mm from the large cylinder boundary. This model was specifically created at a frequency of 4.5 GHz in an attempt to have a model at a frequency higher than the cutoff frequency of the inner cylinder (4.4 GHz.)
3. cyltest2.pjt - An open cylinder excited by a square coaxial probe at the center of the cylindrical shell with the probe oriented along the x axis. The square cross-section probe was used to alleviate the problem of a high mesh density at the small cylindrical probe. The inner conductor (probe) of the coaxial feed is 4 mm wide, the outer conductor has a 9.2 mm inside width. The radiation boundary was chosen to accommodate a 0.75 GHz - 1.5 GHz frequency range.

4. cyltest3.pjt - This project is the same as cyltest2.pjt with the exception that the radiation boundary is closer to accommodate a 1.5 GHz - 2 GHz frequency range.

Once again, we ran up against memory (RAM) and disk storage limitations. Table 1 shows the details of the various projects and models at the given frequencies along with the memory requirements (disk space and RAM) of each. We were unable to converge to a 2 GHz model for the telescop.pjt. Also, the telehf.pjt model did not converge. Even though a good model was not generated at 4.5 GHz telehf.pjt, we looked at the radiation results for the best model generated. When compared to the radiation plot of telescop.pjt at 1.18 GHz, it is clear that once we get above the cutoff frequency of the inner cylinder, the radiation increases substantially. It is interesting to note that even though the frequency of excitation

Table 1. Summary of HFSS projects and resource requirements.

Name of Project	Frequency	A matrix Disk Space (K)	U matrix Disk Space (K)	csolve RAM (K)	Total Time
telescop.pjt	1 GHz	14,130	77,376	66,332	17:13
	1.5 GHz	23,122	164,935	103,132	41:48
	2 GHz	---	>340,000	---	---
telehf.pjt	4.5 GHz	---	>340,000	---	---
cyltest2.pjt	1.5 GHz	34,083	211,128	107,996	2:40:00
cyltest3.pjt	2 GHz	39,834	243,013	114,396	3:03:00

for telescop.pjt (1.18 GHz) is below the cutoff frequency of the inner cylinder waveguide, field plots of the electric fields show that there are calculable fields at the end of the inner cylinder and at the aperture of the large cylinder. The magnetic field plots also show significant fields between the large mirror and the outer cylinder.

The most successful models were cyltest2.pjt and cyltest3.pjt. HFSS results for S_{11} are shown in Figure 17.

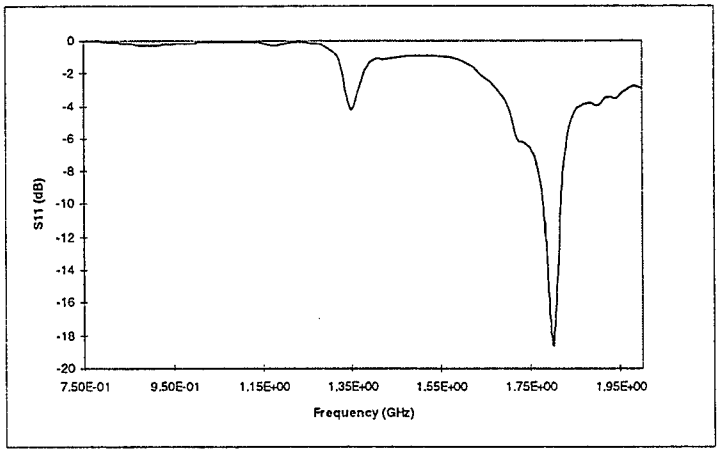


Figure 17. HFSS results for the probe excited open cylinder (20 MHz resolution.)

Figure 18 shows results for a test of the Celestron 8 Telescope with a coaxial probe positioned as for cyltest2.pjt. As a reference, the TE_{211} and TE_{311} modes for a closed cylindrical cavity of the dimensions of our open cavity resonate at 1.36 GHz and 1.822 GHz. The electric field plots at 1.35 GHz and 1.8 GHz for these HFSS models clearly indicate that the transverse modes are TE_{21} and TE_{31} respectively. So our simple model of an open cylinder accurately models the telescope for these low frequency resonances.

Figures 18 amd 19 represent the physical test results using a probe excitation. In one scheme, a 3 cm probe was placed at the center of the most outer shell and a set of S_{11} measurements were taken. This measurement was done once having a conducting plate at the aperture of the telescope and then same measurements were done having no cover plate at the aperture. In both cases a receiver horn (probe) antenna was placed at the eyepiece of the telescope. The probe antenna was 0.88 inches in full length with aperture dimensions of 0.49 inches in length and 0.40 inches in height. Figure 18 shows the results having no cover plate, and Figure 19 shows the same measurements with a conducting cover plate at the aperture. It is evident from Figure 18 that this test result follow the same trend as the result obtained using HFSS shown in Figure 17.

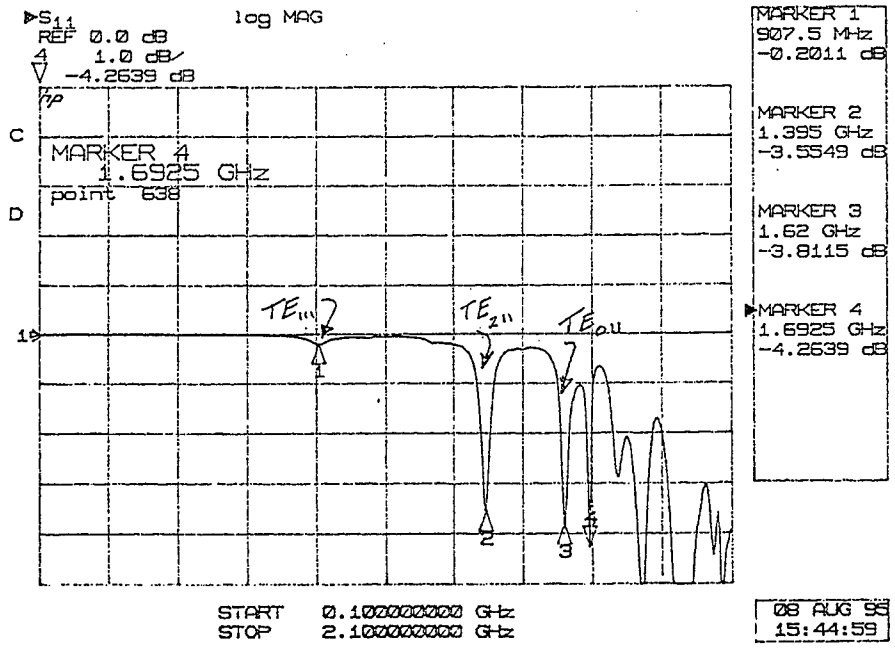


Figure 18. S_{11} measurements of the Celestron-8 telescope using a center mounted 3 cm probe - no cover plate.

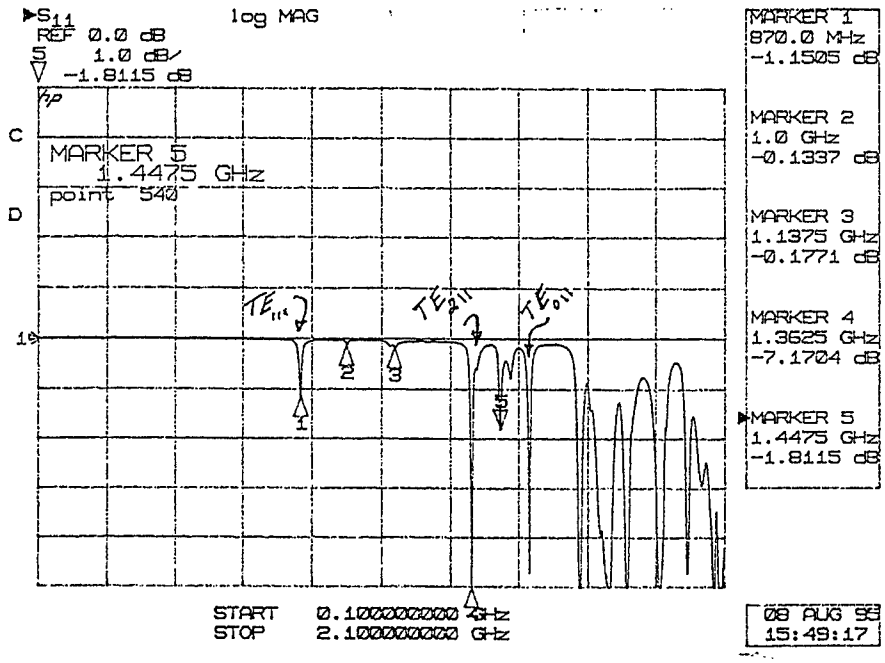


Figure 19. S_{11} measurements of the Celestron-8 telescope using a center mounted 3 cm probe - with cover plate.

Test description and results

Testing was done by Hugh Pohle over a frequency range of 1 GHz - 40 GHz. For these measurements, the telescope was excited by a transmitting antenna 40.0 inches in front of the aperture. A probe antenna was mounted at the eyepiece. The transmitting horn antenna was 3.02 inches in full length with aperture dimensions of 1.60 inches by 1.21 inches. The probe horn antenna was 0.88 inches in full length with aperture dimensions of 0.49 inches by 0.40 inches. Figure 20 shows one series of these measurements.

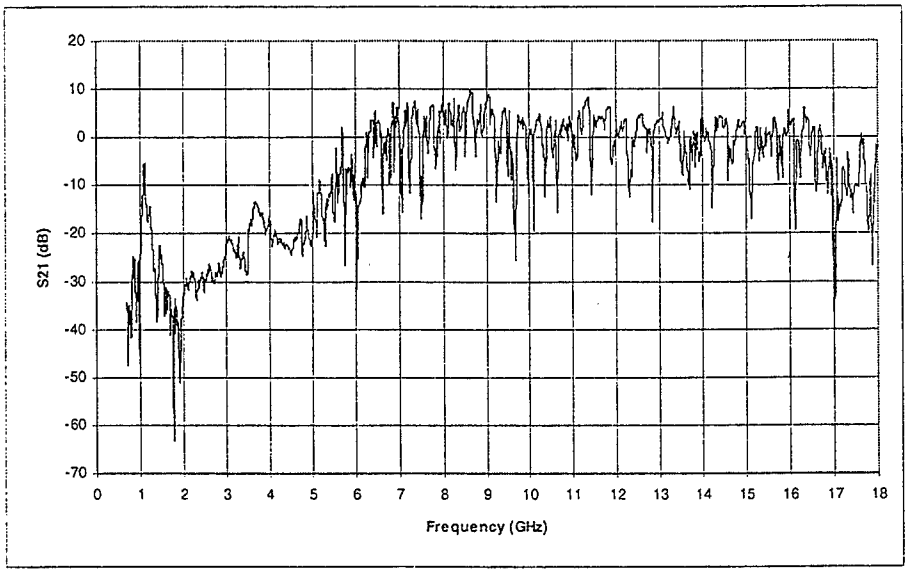


Figure 20. S_{21} Measurements of the telescope.

New approach suggested by Dr. Coffey [11]

It is known from experiment that the resonances of the cavities under consideration do not vary much from the closed empty cavity case. This knowledge is used to extend the traditional eigenvalue approach to a more general approach that is suitable for open-ended cavities. The method includes the following:

1) *Solve the eigenvalue problem for Dirichlet and Neumann boundary conditions on the openings and Dirichlet boundary conditions elsewhere.* This is the classic eigenvalue/eigenvector problem. From this step we can obtain estimates of eigenvalues and eigenvectors for the open-cylinder case. In a transmission line analogy, the Dirichlet conditions correspond to short-circuit loads, while the Neumann conditions correspond to open-circuits loads.

2) *Use the eigenvectors obtained in step 1 as basis functions to generate an equivalent impedance matrix* $Z_{int}(r,r')$ *that fully describes the interior problem at any frequency.*

3) *Use GEMACS (General Electromagnetic Model for the Analysis of Complex Systems) to obtain an equivalent impedance matrix* $Z_{out}(r,r')$ *that fully describes the exterior problem at any frequency.* We know that if the FD/FE (Finite Element) lattice extends to the far-field sphere, the impedance matrix is diagonal with terms equal to 376.7 ohms per square. The impedance matrix at the cavity opening(s) is not diagonal but dense, and the matrix terms vary extensively. However, by not having to work a FD/FE problem in the exterior, great computational savings can be made. GEMACS can be used to "bring the infinite boundary to the opening." This is a MOM/UTD (Unified Theory of Diffraction) hybrid problem (not an FD/FE problem. In the transmission line approach, this might be seen as referring a load impedance to another point along the transmission line via equations or smith chart.

At the same time, GEMACS may bring a source or sources from their exterior locations to equivalent sources at the opening. The combination of impedance matrix and equivalent excitation completely describes the exterior physics to the interior problem.

4) *Utilize Householder's method of modified matrices to combine the matrices together, enforcing E and H continuity at the open boundary.* This is the technique used in GEMACS, by Harrington and Mautz, and by K. K. Mei. In the transmission line (TL) analogy it is like connecting two (or more) lines together at a single point. The act of "connection" ensures the transmission line voltage is the same on either side of the junction (analogous to E-field continuity) and the transmission line currents are equal and opposite on either side of the junction (analogous to H-field continuity).

5) *Find the stationary points of the combined problem. At these points will be the complex resonant frequencies. From the complex resonant frequencies, obtain the resonances and Q's for each mode.* Depending on how rapidly the exterior equivalent impedance varies with frequency, this may require more than a matrix eigenvalue computation.

6) *Utilize the eignevectors of step 1 as basis functions for the coupling problem.* Since these are not orthonormal for the open-cavity case, the Rayleigh-Ritz-Galerkin procedure will not yield a diagonal matrix.

However, only the eigenvectors corresponding to eigenvalues near the operating frequency are required. Thus, the matrix problem to be solved for field coupling in the open-cavity case is quite small, perhaps no more than 20x20.

Conclusion

We have observed that using the method of iterative refinement did not make noticeable improvements in the near field calculations. However, all the methods in this report were applied to an empty open-ended cylindrical cavity. Therefore, we may or may not see different responses to the method of iterative refinement related to changes in the internal geometry. Using suitable Gaussian Quadrature did make substantial difference in the results. However, an adaptive method of integration is necessary in order to avoid comprehensive analysis trying to find the appropriate Gaussian Quadrature.

Using non-uniform meshing schemes which provide a model for the edge currents, did make a substantial improvement in the energy calculations. A denser field mesh in the aperture will undoubtedly improve the power calculations. However, until the appropriate formulation for calculating the near fields, to include the variations of current and of distance to the field point over individual cells, is implemented in CARLOS adequate accuracy of near fields can not be obtained.

CARLOS calculations in this report were only done for an empty open-ended cylindrical cavity. We ran up against memory (RAM) and disk storage limitations beyond a $\lambda/9$ discretization at 2 GHz. Once the interior geometry of the telescope is implemented, it would be necessary to increase these computational capabilities in order to continue with these calculations.

If Dr. Coffey's approach is successful, it will enable us to use other more computationally intensive codes over a more limited frequency range. It seems that once a high enough frequency is reached, computational resources will be strained. If we can limit use of those resources to frequency ranges of interest, our analytical problem will be substantially reduced. Of course, much research and validation needs to be done in order to obtain confidence in such a method.

Finally, the concept of determining energy at a particular point or area in a system by doing the inverse (antenna) problem is an attractive one. The vast amount of information contained in the HFSS solution can be used to determine optimum incidence angles to couple the energy into the system, field quantities anywhere in the system, system resonances, quality factor for the system, etc. HFSS was used in this study to demonstrate the concept; perhaps other types of codes (finite difference - time domain, MOM, etc.) could also be used to solve the "antenna" problem. CARLOS would be a good candidate to do the antenna problem because of its

ability to handle dielectrics as well as conductors. We encourage its modification to include port and wire excitations and post processing to give radiation patterns and radiation efficiency.

References

[1] Wolfram, S., Mathematica, Addison-Wesley, 1991.

[2] J.M. Putnam, L.N. Madgyesi-Mitschang and M.B Gedera, CARLOS-3D™ Three Dimensional Method Of Moments Code, McDonnell Douglas Aerospace-East, December 1992.

[3] General Electromagnetic Model for the Analysis of Complex Systems (GEMACS) - Version 5, Rome Air Development Center, AFSC, Griffiss Air Force Base, NY, December 1990.

[4] R. R. DeLyser and P. Ensaf, "Analysis to Determine the Quality Factor of a Complex Cavity," Final Report for the Summer Faculty Research Program at the Phillips Laboratory, Sponsored by the Air Force Office of Scientific Research, September, 1994.

[5] R. R. DeLyser, "An Analysis Approach to Determine Quality Factors of Large , Complex Cavities," Final Report for the Summer Faculty Research Program at the Phillips Laboratory, Sponsored by the Air Force Office of Scientific Research, September, 1993.

[6] Balanis, C.A., Antenna Theory, John Wiley & Sons, New York, pp. 100-109, 1989.

[7] Johnson J.H. Wang, Generalized Moment Method In Electromagnetics, John Wiley & Sons, New York, pp. 74-75, 1991.

[8] Ishimaru, Akira, Electromagnetic Wave Propagation, Radiation, and Scattering, Prentice Hall, New Jersey, pp. 578-580, 1991.

[9] Canning, Francis X. "Singular Value Decomposition of Integral Equations of EM and Applications to the Cavity Resonance Problem". IEEE Trans, Antennas Propagat., vol. AP-37, no. 9, pp. 1156-1163, 1988.

[10] William H. Press, Brian P. Flannery, Saul A. Teukolsky, William T. Vetterling, Numerical Recipes, Cambridge University Press, New York, pp. 47-48, 1986

[11] Edgar L. Coffey, III, "Electromagnetic Coupling to Open-Ended Cavities", Proposal Report to Phillips laboratory, August, 1995.

Average Power Effects in $AgGaS_2$ Optical Parametric Oscillator

Juan C. Garcia
Graduate Student
Department of Electrical Engineering

University of California, Los Angeles
405 Hilgard Ave.
Los Angeles, CA 90024

Final Report for
Graduate Student Research Program
Phillips Laboratory

Sponsored By;
Air Force Office of Scientific research
Bolling Air force Base, DC

and

Phillips Laboratory

September 1995

Average Power Effects in $AgGaS_2$ Optical Parametric Oscillator Title

Juan C. Garcia
Graduate Student
Department of Electrical Engineering
University of California, Los Angeles

Abstract

The behavior of a silver thiogallate ($AgGaS_2$) optical parametric oscillator(OPO) was studied as a function of the average pump power. The singly resonant OPO operates at a signal wavelength near 1.319 μm and a idler wavelength near 5.505 μm. The OPO was synchronously pumped by a cw actively mode-locked Nd:YAG laser at 1.064 μm. The output of the pump was 18W and for the experiment the average power was reduced by a mechanical chopper to 0.47W, 0.95 W, 1.95 W, 3.8 W and 7.8 W. Results indicated that the output of the signal of the OPO was linear up to 3.8W. At a 7.8W thermal effects began to affect the performance of the OPO.

Average Power Effects in $AgGaS_2$ Optical Parametric Oscillator

Juan C. Garcia

Introduction

Optical Parametric Oscillators(OPO) are a convenient way to generate tunable wavelengths not covered by conventional lasers sources. As more high power applications of OPO's are found, the high average power requirements will begin to affect the performance of OPO's [1]. Effects such a thermal lensing and phase mismatch become relevant due to absorption of the pump and signal by the nonlinear crystal[1].

In this experiment average power measurements were performed on a silver thiogallate($AgGaS_2$), singly resonant OPO. The OPO was pumped by a cw actively mode-locked Nd:YAG laser. The $AgGaS_2$ crystal was cut at θ=48 and ϕ=0 for type II phase matching. The signal wavelength is near 1.319 μm corresponding to an idler wavelength near 5.505 μm[2]. The average power of the Nd:YAG laser was 18W and was reduced by a chopper to 0.47 W, 0.95 W, 1.95W, 3.8 W and 7.8 W. At each power level, the spectrum and temporal pulse width were measured and also a scan of the OPO cavity was performed.

Methodology

A schematic of the experimental setup is shown in page7-6. There are four output beams coming out of the Brewster plates. One of the beams was used to monitor the signal power using a calibrated Ge detector and also the transmitted pump was observed by putting a 1.064 μm reflector and sending the transmitted pump to a Si photodiode. The other two were used to look at the spectral and temporal profiles of the signal and the temporal pulse shape of the pump. The spectrum of the signal was observed by a supercavity spectrum

analyzer and for the pulses a fast photodetector was used. The calculate the FWHM of the spectrum, the data was calibrated by using the spacing of the transverse modes which was 28 GHz

Results

In Page 7-7 a plot of the signal vs. the average power pumped into the OPO is given. The signal output maintained a linear relationship until the pump power reaches 7.8W . The same can be said about the pulse width which is plotted in page 7-8. And the Spectral width is plotted in page 7-9 and this shows a increasing width as the average power is increased. The spectrum for the 7.8 W case is not given since the spectrum could not be calibrated using the transverse mode spacing. The gradual broadening of the spectrum is expected due to variations of the indices of the refraction as the average power is increased.

At 7.8 W thermal effects begin to affect the performance of the OPO. The thermal gradients prevents maximum phase matching conditions over the entire crystal. Also thermal lensing decreases the performance, but in this case it depends on the change of one index of refraction as opposed to the difference in the change of the ordinary and extraordinary indices for phase mismatch[1].

It is interesting to note that for noncollinear phase matching at 7.8 W, the OPO provided a signal output of 985 mW which is more a long the lines of the expected value.

Conclusion

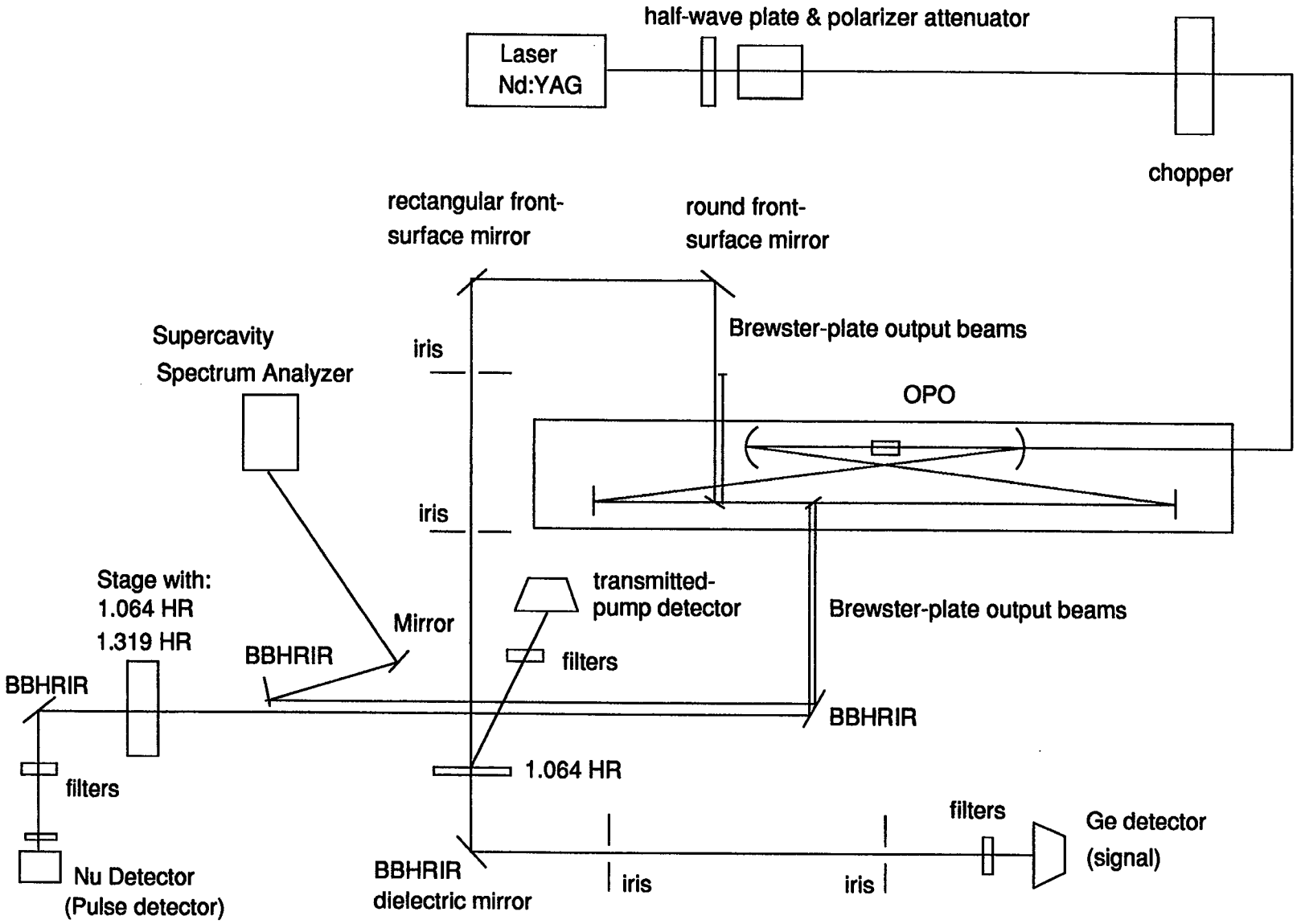
With this information a the operation of the OPO can be maximized. The absorption coefficients along with other parameters will be calculated to optimize the operation of the $AgGaS_2$ OPO.

REFERENCES

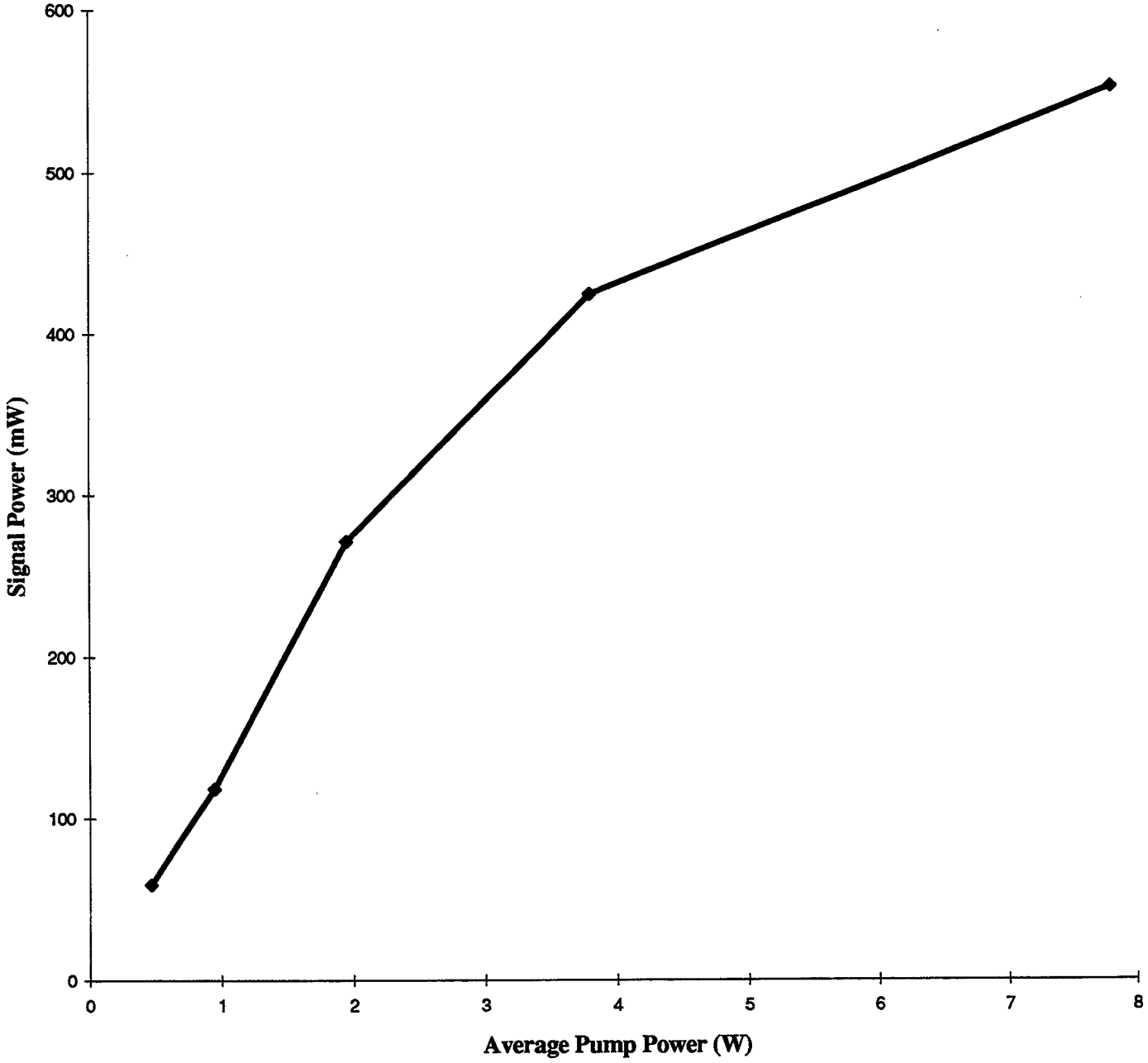
1 N.P. Barnes, J.A. Williams-Byrd, "Average Power effects in Parametric Oscillators and Amplifiers" Journal of the Optical society of America B, Vol. 12, No. 1, 124-131 (1995)

2 E.C. Cheung, K Koch and Gerald T. Moore, "Silver thiogallate, singly resonant parametric oscillator pumped by a continuous-wave mode-locked Nd:YAG laser", Optics letter, Vol. 19, No. 9, 631-633 (1994)

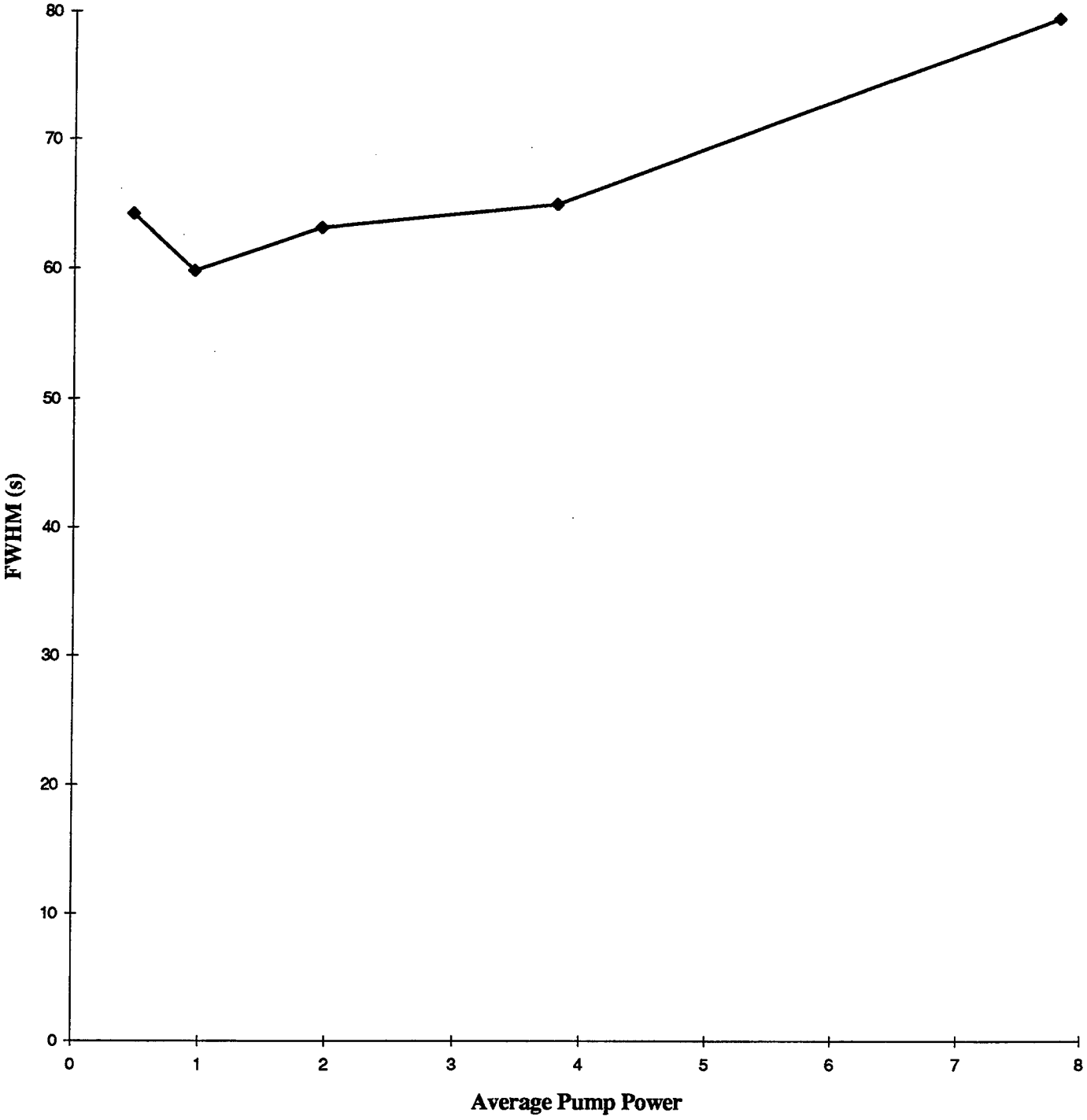
Power Experimental Setup



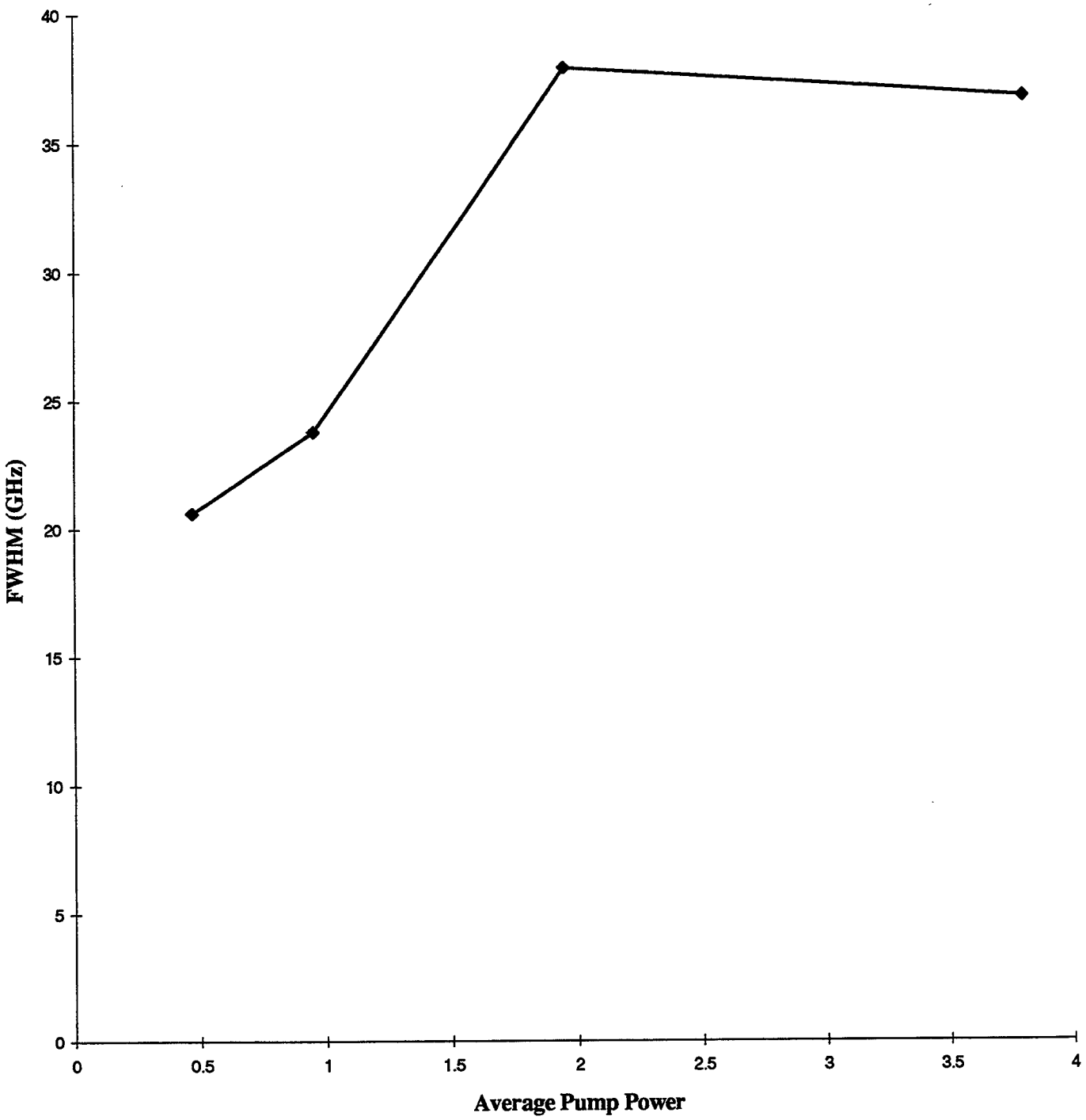
Signal Power



Temporal FWHM



Spectral FWHM



SPACE WORTHINESS OF MEMs DEVICES: INITIAL STUDIES

Winn Hong
Graduate Research
Department of Mechanical Aerospace Engineering
Center for Micro Systems

University of California, Los Angeles
405 Hilgard Ave.
Los Angeles, CA 90095

Final Report for:
Graduate Summer Research Program
Philips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Kirkland Air Force Base, NM

and

Philips Laboratory

September 1995

SPACE WORTHINESS OF MEMs DEVICES: INITIAL STUDIES

Winn Hong
Graduate Research
Department of Mechanical and Aerospace Engineering:
Center for Micro Systems

Abstract

The feasibility of using micro-electro-mechanical system (MEMs) devices for possible space applications is investigated. Specifically, the Analog Device's ADXL-05 and ADXL-02 tri-axis micro-accelerometer units will be used in hope of monitoring the changes in satellite acceleration due to tumbling, release shock, and orbiting. Ground tests using the shaker table for characterization before and after the total dose radiation affects in a Cesium source irradiation chamber is also performed. Temperature affects and long term dc drift are also investigated.

SPACE WORTHINESS OF MEMs DEVICES: INITIAL STUDIES

Winn Hong

1.0 Introduction

The following is a Microsystems And Packaging for Low-power Electronics (MAPLE) project for the Applied Microelectronics Branch of the Philips Laboratory (PL/VTEE). MAPLE consists of an assembly of five to eight subexperiments, each encased in its own space qualified aluminum frame. Of the assembly, one of the test frames will be the Micro-Electro-Mechanical system (MEMs) tri-axis accelerometer.

Micromachine technology has been in existence for the past 40 years. Since then it's development has grown exponentially, a growth very similar to what IC industry has done in the past. In fact, Micromachine technology is the derivative of the IC processing technique; by layering, patterning, etching, etc... below the surface of the substrate. The major difference being that Micromachine technology took this IC process above the surface of the substrate, commonly known as "surface micromachining." By building structures both above and below the surface, these micro structures now have the degrees of freedom in movement at submicron precision that is only limited by its designer.

In 1990, the total funding for MEMs was estimated to be about $500 million. By the year 2000 ARPA-Advanced Research Programs Agency, has estimated this amount to be about $15 billion dollars in total federal and commercial funding. As of today, there has been many fundamental, as well as applications, research and development in MEMs. However, the actual commercial use of MEMs devices have been limited to a small fraction of its capabilities, such as air bag deployment, NOX, and air to fuel mixture sensors. It is unfortunate that such an impressive technology-with its extremely low cost, small size, low power consumption, high sensitivity, negligibly light weight, and batch fabrication features, should have such misuse, or lack of use; especially in space and satellite applications where size, weight, and power consumption are at a premium. Which brings to the main point of this paper of space worthiness of MEMs devices. But, before any further discussion of the space worthiness of MEMs an introduction to the basic operation of the MEMs device is due. Note, at the time ground tests were performed the tri-axis

accelerometer was not yet available. Instead a single axis unit was tested in its place. This, of course, is not what was intended. However, since the tri-axis unit is composed of 3 individual single axis units the results from the single axis can be correlated to the tri-axis sensor.

2.0 Accelerometer Operations Theory

The particular sensor used is an Analog Devices' ADXL-05 Monolithic Accelerometer. The tri-axis accelerometer, including packaging and electronics, weighs less then 2 ounce and has the dimensions 2"x1"x .75" cubed. The accelerometer senses by means of 42 differential capacitive sensor units. Each consists of a pair fixed outer plates with a center plate that is fixed to a common moving, "floating," central plate which is anchored to the silicon substrate, see figure 1. The displacement between the two fixed outer plates and the center plate as a result of a change in relative motion is immediately corrected by a force balance technique which counters any impeding deflection due to the acceleration and forces the sensor back to its 0g position [Analog Devices Accelerometer Date Sheets and Application Notes].

Figure 2 shows the sensor responding to an applied acceleration. When this occurs the floating beam moves closer to either one of the fixed capacitive plates while moving further away from the other. The sensor's fixed capacitor plates are driven deferentially by a 1MHz square wave: the two square wave amplitudes are equal but are 180° out of phase. When at rest, the values of the two capacitors are the same and therefore, the voltage output at their electrical center (i.e. at the center plate) is zero. Once subjected to an acceleration, its capacitive sensor begins to move, creating a momentary output signal. This signal is conditioned and amplified through the demodulator and pre-amp circuits.

When a signal is present the demodulator the same 1MHz oscillator is used to rectify any voltage that is in sync with the clock signal. If the signal is in sync and in phase with the clock then the signal would register a positive acceleration. Synchronous and 180° out of phase signal is registered as negative acceleration. Otherwise, all other signals are rejected. During 0 g output the sensor is referenced at +1.8 Vdc, with an uncommitted buffer amplifier to adjust the scale factor and 0g offset, figure 3.

For further discussion of the accelerometer and its applications please refer to the Analog Devices Data Sheets and Application Notes for the ADXL-50. Also note that the basic structures are the same for all ADXL accelerometer series.

3.0 Driving Circuit

The ADXL units come with an internal signal conditioning demodulator, a preamplifier to isolate the demodulator from the amplification stage and a buffer amp to allow scale factor and 0g offset adjustments. All of the internal circuitry and the monolithic accelerometer are hermetically sealed in a TO-100 can. By itself, the accelerometer is ready for use. However, to be able to trim the 0g offset, i.e. 0g offset at +2.5Vdc with a ±2V output swing, such that the sensor does not saturate nor hit the supply rails a simple external circuit is implemented.

This simple circuit, as shown in figure 4, allows us to select several operational frequencies as well as several full-scale ranges. The full-scale range of ±10g, with the maximum sensitivity of 200mV/g and a DC-100Hz 3dB bandwidth frequency range are arbitrarily selected. The corresponding circuit components are used to build the driving circuit with the 0g offset trimmed at +2.5Vdc. Note that as the full-scale range is decreased, the buffer amplifier gain is increase, and the noise contribution as a percentage of full scale will also increase. This signal-to -noise ratio can be improved by reducing the circuit bandwidth by either increasing the demodulator capacitance, C1, or by adding a post filter using the buffer amplifier [ADI accelerometer data sheet and application notes].

Two schemes of filtering were employed: a two pole filter which could not eliminate much of the jitter. As a result, a three pole post filter scheme was selected. This showed a much improved filter for relatively low frequency noise and at the same time showed a high level of output stability.

It must be mentioned that the op-amp AD820, as suggested in the schematic, was not available during the initial bread-boarding. Instead, the AD820 was replaced with a TL082 dual bifet op-amp. The TL082 is a dual supply op-amp which was used in a single supply fashion. As a result, when the trim potentiometer was adjusted to both the minimum and maximum dc output these values were not enough to cover the ±2 V output swing of the sensor, i.e. range of dc output using the TL082 is only +2.02 Vdc ±.29V rather than the +2.5 Vdc with ±2 V output swing.

When the AD820 op-amp was finally placed into the circuit things did not go as expected. The dc output mysteriously drifted uncontrollably. Trouble shooting showed that this random drift also occurred at the output of the sensor before even going through the op-amp. One speculation as to the cause of this anomaly is that the RC oscillator in the sensor might be laser trimmed such that it might have high frequency phase jitters that are superposed with the 1MHz pulse. Consequently, it was finally decided that

an AC coupled circuit was necessary to keep the output stable enough to do any meaningful ground testing [figure 5]. For reason that the operating frequency range is 1Hz-100Hz, the circuit was built to have an output of 100mV/g.

4.0 Ground Testing

There are several issues which must be tested on the ground, or ground tests, in an attempt to simulate and characterize the accelerometer for orbit. First, the affects on the functionality of the sensor due to total dose radiation onto the accelerometer is monitored. Second, the dc drift and sensitivity of the sensor due to temperature change is verified. Finally, the long term dc drift of the circuit is checked.

The long term dc drift monitor was performed on the first circuit shown in figure 4 with a three pole post filter circuit with a gain of 10. This circuit contained the TL082 op-amp trimmed at 0g offset at +2.68Vdc with a ±15 mV drift over a period of 16 days. It was mentioned earlier in the report that the TL082 op-amp causes saturation of the output signal. However, since the dc drift voltage is well within the rail limits of the TL082 it is safe to assume that there is relatively low drift. It is also safe to assume, based on the fact that the output did not change significantly over a 16 day period, that the dc drift of the sensor is acceptable over the life time of the mission of one year.

The temperature test is not at all extensive such that the unit can be characterized. It is only a quick check to see in which direction will an increase in temperature affect the output. The output dc drift due to temperature change was verified on the same board as the above. By placing the circuit into a heat source at 20°C higher than atmosphere, the output changed from +2.68V to +2.835V. According to the specification sheet for the accelerometer unit the drift should only be about 4mV. This, of course, does not agree with the specification. When the unit is touched with the finger the same trend and similar increase in dc output. From this it is safe to assume that the thermal coefficient of resistance of the electronics is not the cause of this drift. This would mean that these units must be temperature compensated somehow, by either a constant temperature quartz oven or an AD590 chip as shown in the applications manual.

A more extensive test was performed for the radiation affects. To get the initial characterization of the accelerometer's response to acceleration, it was first taken to a shaker calibration table. The accelerometer was mounted onto the shaker table with the sensitive axis in the vertical direction. Leading

from the output pins of the can is a 12" ribbon cable which connects the accelerometer to the circuit board. The sensor was subjected to an array of frequencies and accelerations.

After the shaker test the can was put back into the DIP socket and submerged into a Cesium source for an irradiation baking during a static run. Just before irradiation the sensor output, reference voltage, and the current drain were measured at 2.495V, 3.8V, and 10.6 mA respectively. The Cesium chamber is a 20 Curry source and at 28cm from the source the amount of radiation was measured at 45.5 Rankin/hr, or 39.403 Rad-Si/hr. The static run was monitored daily and the output voltage recorded for a total of eight days, bringing the total dose radiation to 7.6 K rad-si. This amount well exceeds the total Low Earth Orbit exposure for a one year period.

Once the irradiation is finished the sensor was again taken to the calibration shaker table for a post-radiation dynamic test. This, of course, isn't enough testing to fully understand the sensor. However, it will have to conclude here since the summer research term has ended. Further testing will have to be continued once an extension is granted.

5.0 Results and Discussion

In the test for long term dc drift it was found that the output drift is insignificant, even with an op-amp which would cause saturation had the sensor been put to its designed full-scale range of ±10g. However, the drift due to temperature effects is substantially large. A probable cause is likely to be the thermal coefficient of resistance of the electronics inside the can which is amplified by the op-amp having a gain of 10.

A pre-rad acceleration test was performed by varying the frequency and acceleration of the calibration shaker table. As the frequencies were varied from 100Hz down to 10Hz something very peculiar appeared. At the frequency of 23Hz the sensor output was very similar to the factory specs of 100mV/g. It seemed almost as if the frequency response of the sensor took on a Gaussian distribution, with 23Hz as the peak. This "Gaussian" distribution was later explained by Analog Devices technician as caused by the resonance of the ribbon cable which connects the sensor to the circuit board. For lack of time the rest of the data were taken at 23Hz while the amplitude of acceleration varied from .25g to 3g, see figure 6a-g. The results were linear as a function of acceleration as anticipated, see figure 8.

Once the pre-rad shake test was finished the can was placed back into the DIP socket and submerged into the Cesium radiation chamber. The test duration of the irradiation lasted for eight days, during which the can was at a powered on static stage. Data were recorded daily and was found to have no affect on its static dc output [figure 8]. When the board and the can were powered off for removal it was quickly powered on by a portable source to avoid annealing of the trapped charges.

A post-rad acceleration test was performed to see the damages, if any, due to radiation. Surprisingly, the results from the post-rad shake showed an improved sensitivity of 10% of the sensor compared with the pre-rad results. It is uncertain at this stage as to why the sensor behaved this way.

Conclusion

Like IC components, Micro-Electro-Mechanical system (MEMs) devices, too, will soon become an indispensable part of everyday gadgets from homes to offices to space and satellites. Because of its many wonderful features such as its small size, extremely light weight, low cost, batch fabrication, low power consumption, and high sensitivity one of its many possible applications come to mind: Space and satellite systems. Where else would one less pound in weight would save the investor hundreds of thousands of dollars?

One cannot expect MEMs to be flown in multi-million dollar satellites without absolute assurance that these devices will function properly and reliably. To prove the worthiness of MEMs devices in space application is, of course, no easy task and it must start at the very basic and simple level. Which justifies the present research and experiments.

To begin, a long term output drift is measured and was found to have insignificant affect on the performance of the sensor. This is important for devices which must operate for long duration without the benefits of constant monitoring and adjusting, such as a satellite or hard to reach components to have as little an affect due to time degradation.

Temperature affect is a great concern for satellites which have large temperature changes during its orbit cycles. Although the temperature affects on the micro accelerometer is large it can easily be compensated using a constant temperature quartz oven or other electronic compensators as suggested in the Applications Manual.

The last, but certainly not the least important, factor to consider during orbit is radiation affects on electronic components. The test accelerometer was subjected to a total dose irradiation, using a Cesium irradiation chamber, of 7.6K Rad-Si for an eight day duration. This is more than three times the total exposure for low earth orbit in an enclosed bay. Not only did the radiation not degrade the sensor it showed an improved response of 10% in sensitivity. However, this is only one set of irradiation experiment. Repeating the radiation affects test was not possible for lack of time.

Even though these were only preliminary tests one might begin to see the benefits and the worthiness of MEMs devices in space and satellite applications. Now it is only a matter of actually putting these devices in orbit to test to see if they can really replace the existing larger and much more expensive units.

Acknowledgment

The author is grateful for the much needed help from Mr. Dennise Adamson and Mr. Todd Goforth for their help and explanations on the circuitry. The author would also like to express his gratitude to Captain James Lyke for the opportunity to work on this project and his valuable advises. Finally, the author would like to thank the Air Force Office of Scientific Research for providing this program to graduate students.

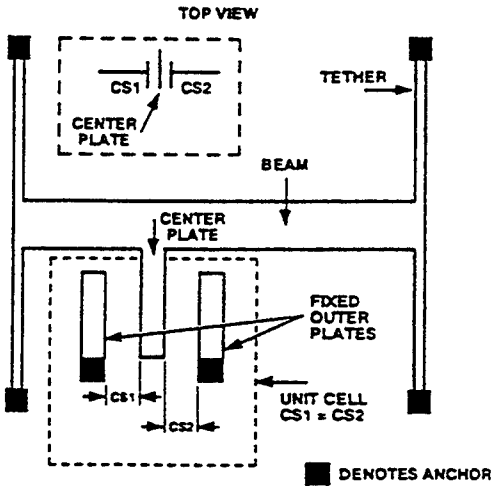


Fig. 1. A Simplified Diagram of the ADXL50 Sensor at Rest.

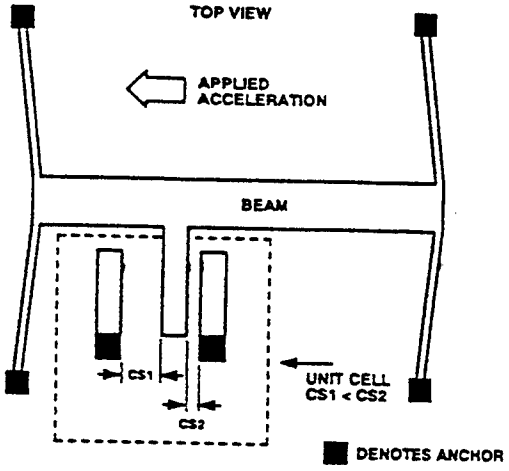


Fig. 2. The ADXL50 Sensor Momentarily Responding to an Externally Applied Acceleration.

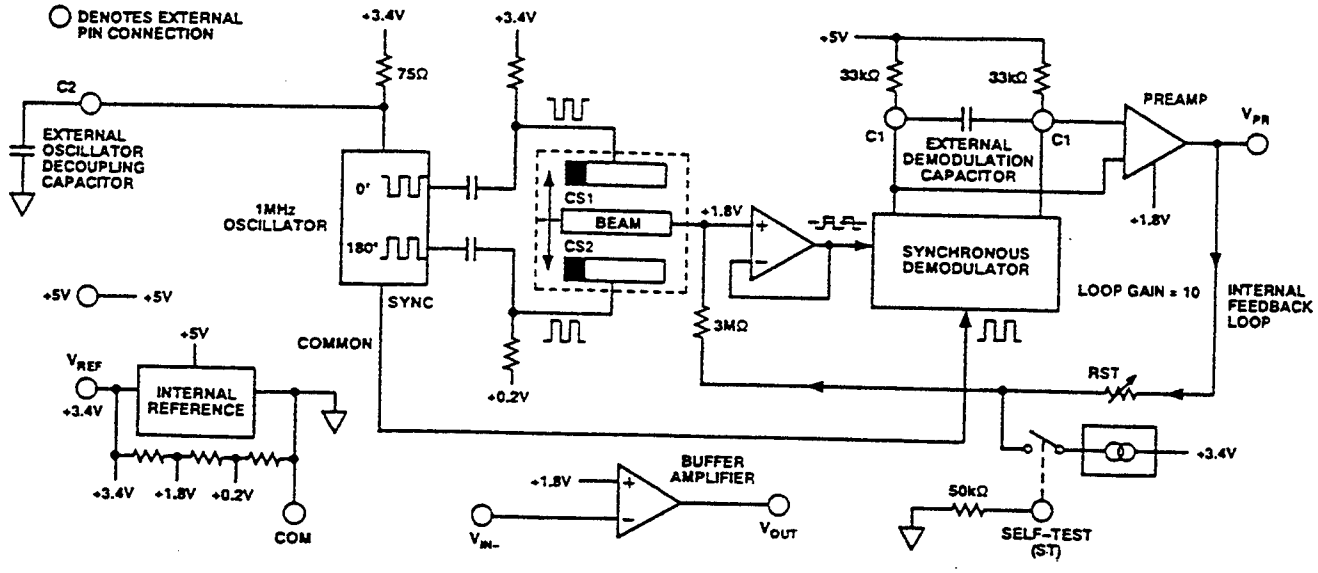


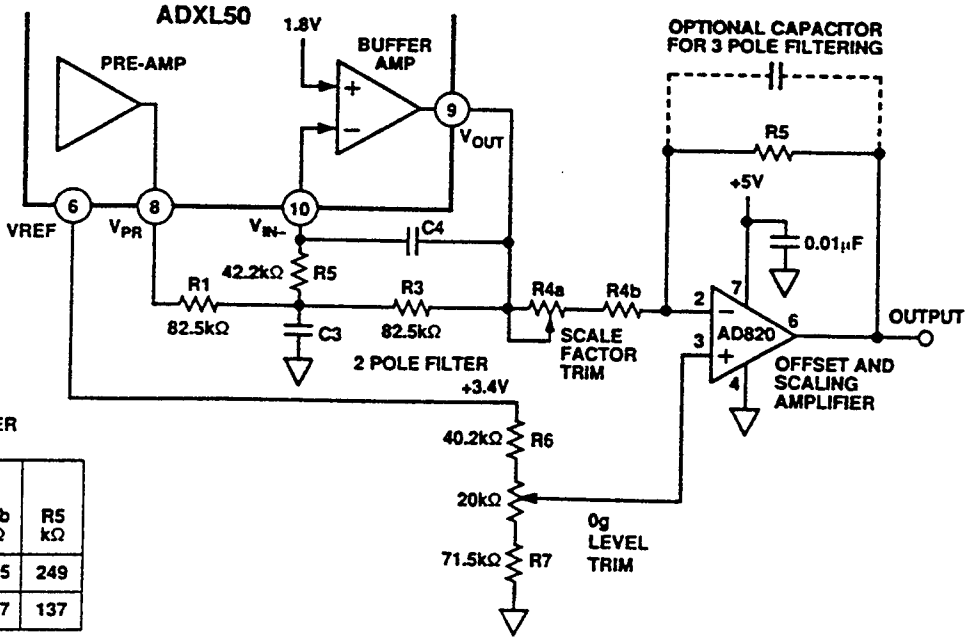
Figure 3. Simplified Block Diagram of Accelerometer Operation.

2 POLE FILTER COMPONENT VALUES

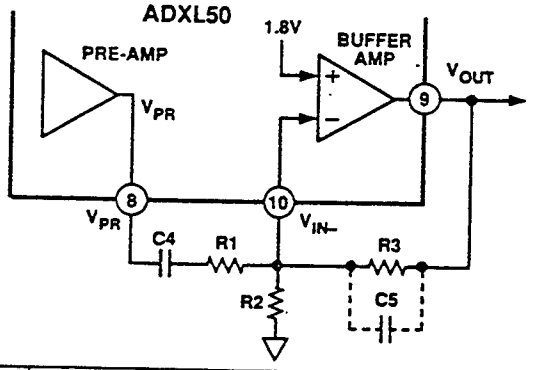
3dB BW (Hz)	C3μF	C4μF
300	0.027	0.0033
100	0.082	0.01
30	0.27	0.033
10	0.82	0.1

OFFSET AND SCALING AMPLIFIER COMPONENT VALUES

SCALE FACTOR IN mV/g	MAX INPUT	GAIN	R4a kΩ	R4b kΩ	R5 kΩ
200	±10g	10.53	5	21.5	249
100	±20g	5.26	5	23.7	137



Firure 4. Two-Pole Filtering Circuit with Gain and 0g Offset Adjustment.



SCALE FACTOR IN mV/g	DESIRED LOW FREQUENCY LIMIT, F_L	R1 USED	CLOSEST C4 VALUE	DESIRED HIGH FREQUENCY LIMIT, FH	R3 kΩ	CLOSEST C5 VALUE
200	30	24kΩ	0.22μF	300	249	0.002μF
100	10	24kΩ	0.68μF	300	127	0.0039μF
200	3	24kΩ	2.2μF	100	249	0.0068μF
100	1	24kΩ	6.8μF	100	127	0.01μF
200	0.1	24kΩ	68μF	10	249	0.068μF

Figure 5a. AC Coupling the Vpr Output to the Buffer Input.

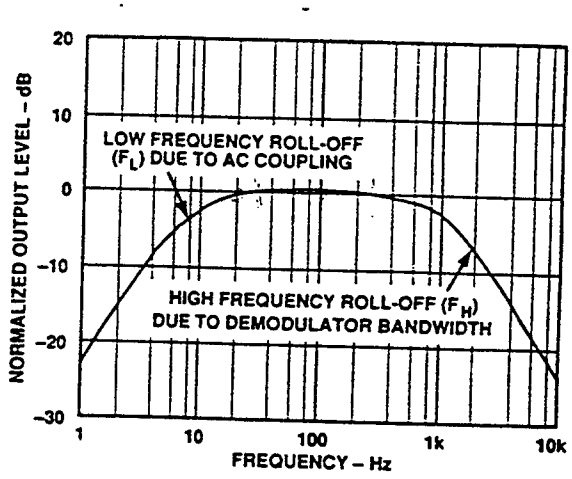


Figure 5b. Typical Output vs. Frequency Curve when AC Coupling Vpr to the Buffer.

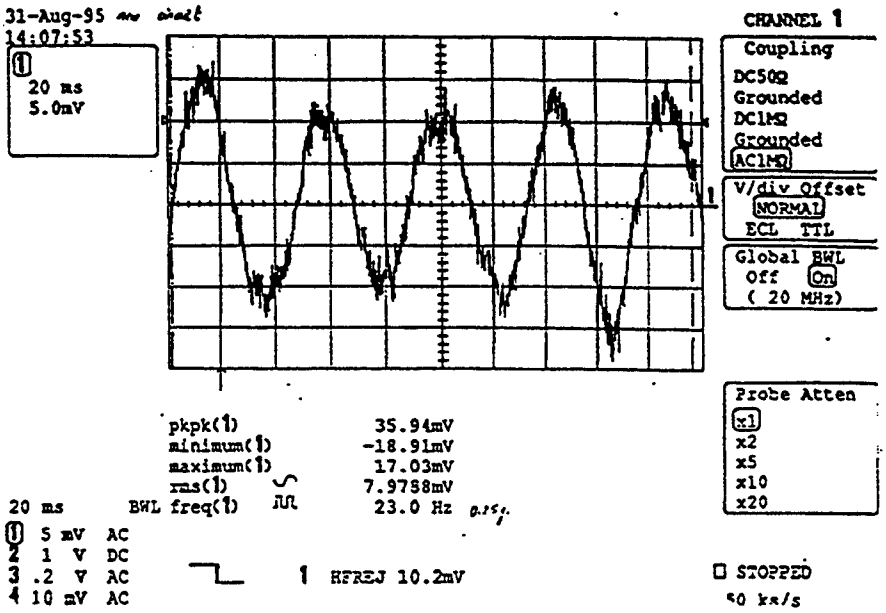


Figure 6a. Pre-rad Acceleration Response of Sensor at 23Hz and .25g Acceleration.

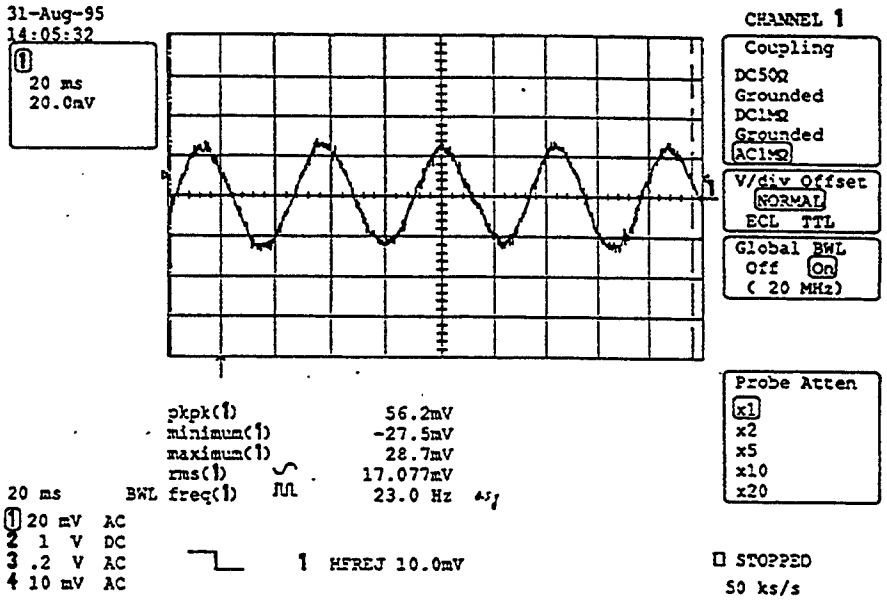


Figure 6b. 23Hz and 0.50g Acceleration.

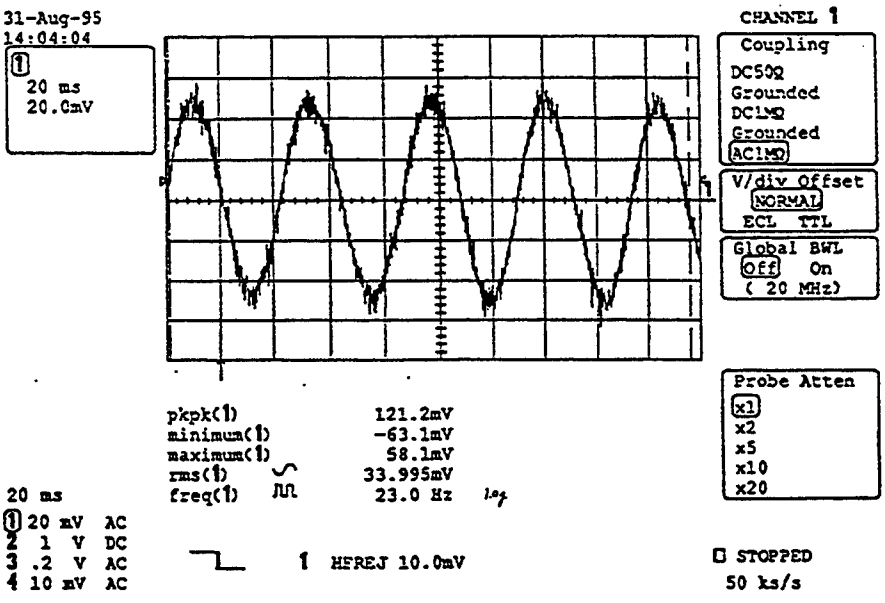


Figure 6c. 23Hz and 1.0g Acceleration.

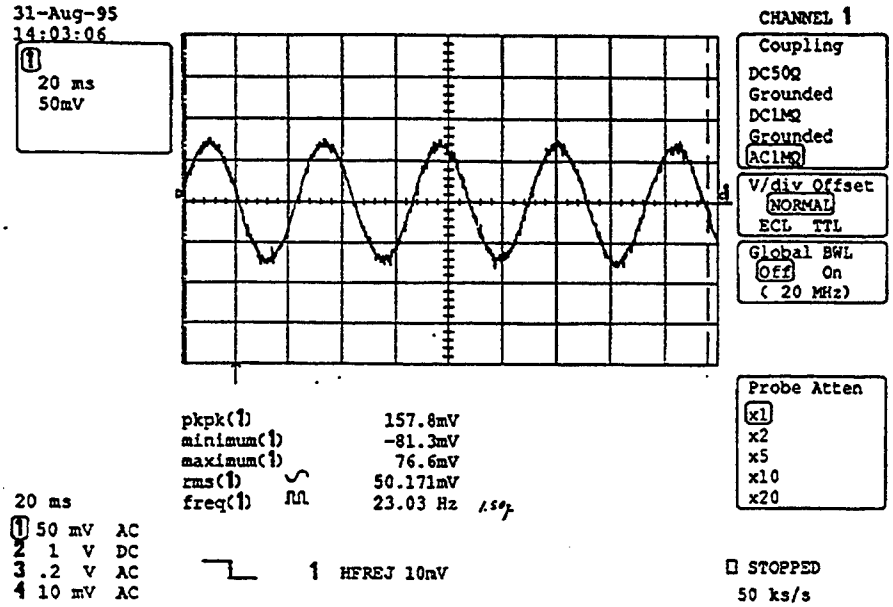


Figure 6d. 23Hz and 1.5g Acceleration.

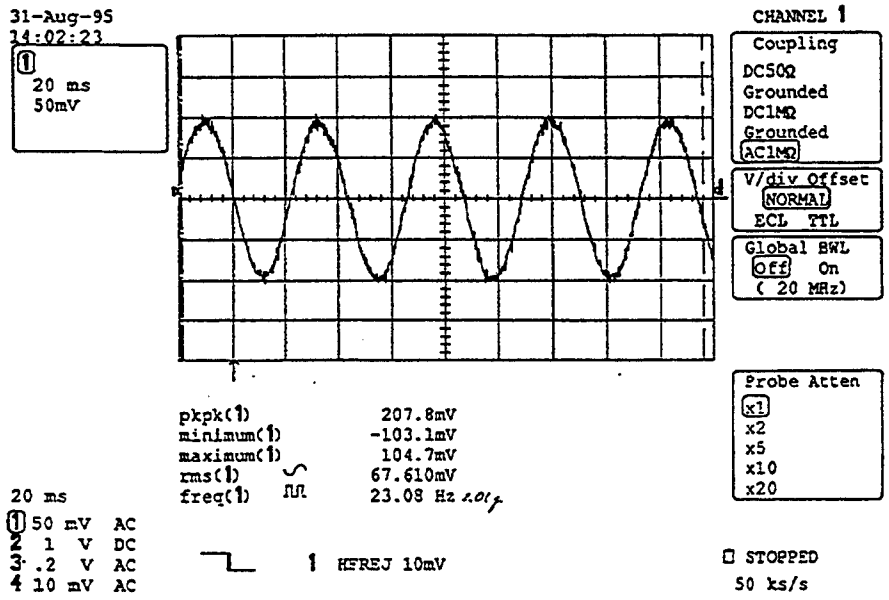


Figure 6e. 23Hz and 2.0g Acceleration.

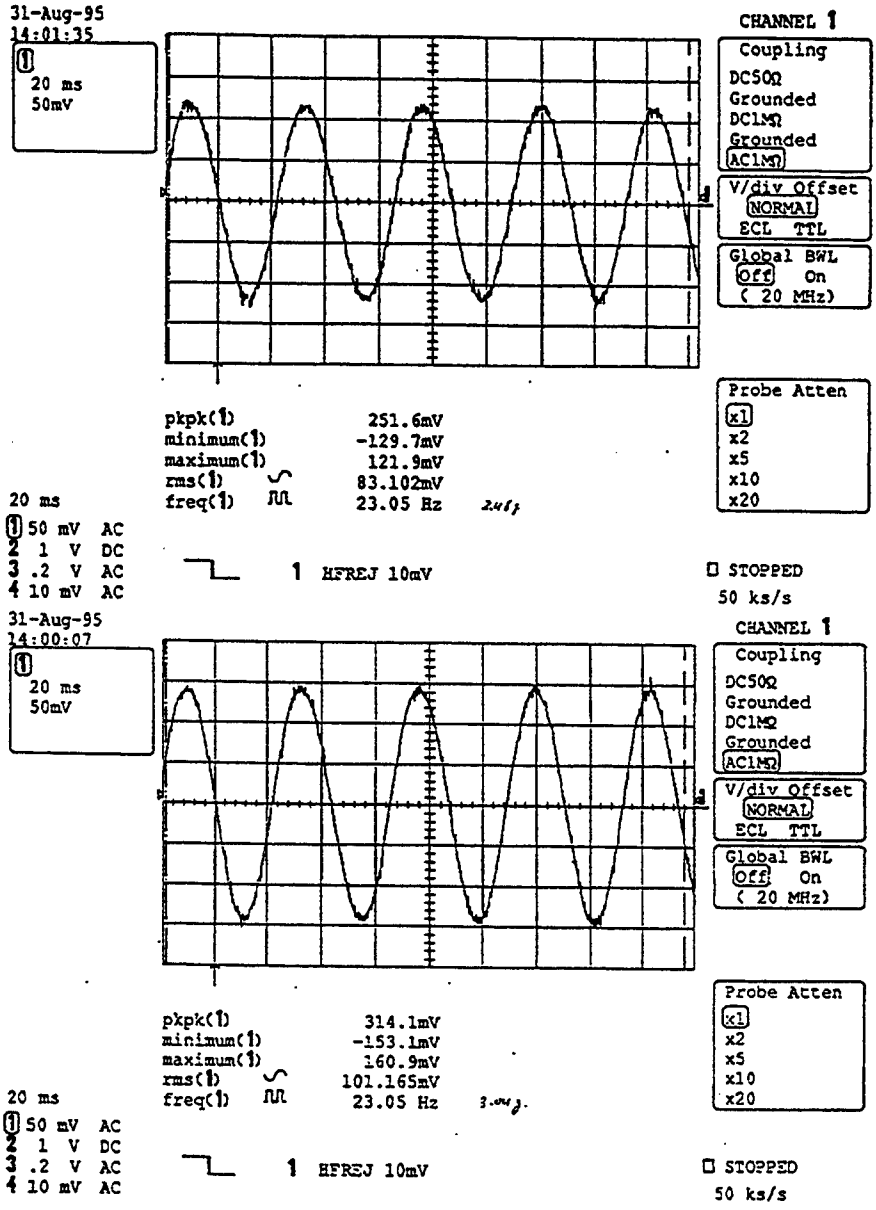


Figure 6f. 23Hz and 2.5g Acceleration.

Figure 6g. 23Hz and 3.0g Acceleration.

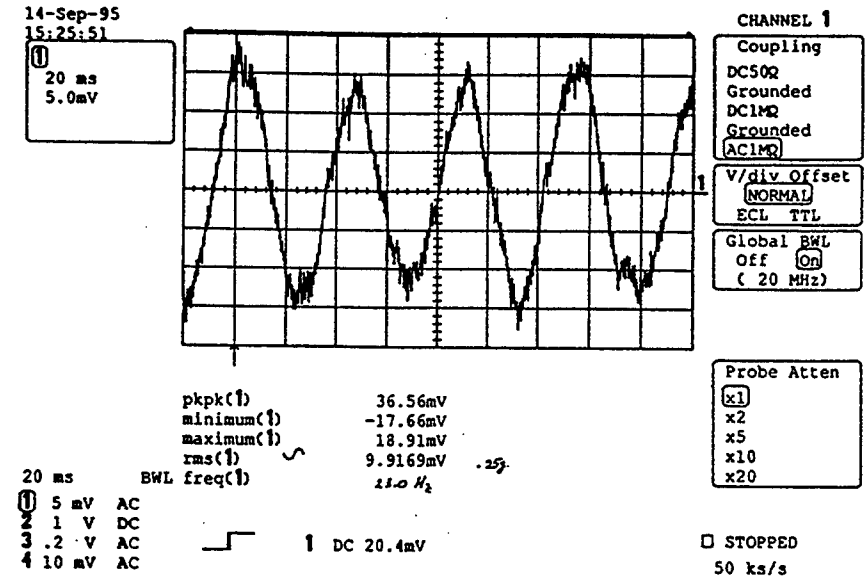


Figure 7a. Post-Rad Acceleration Response at 23Hz and 0.25g Acceleration.

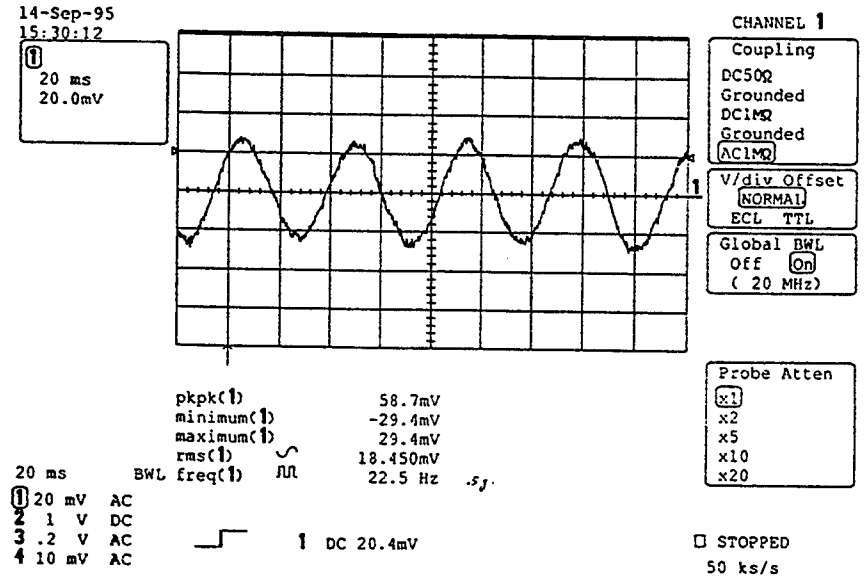


Figure 7b. 23Hz and 0.5g Acceleration Response.

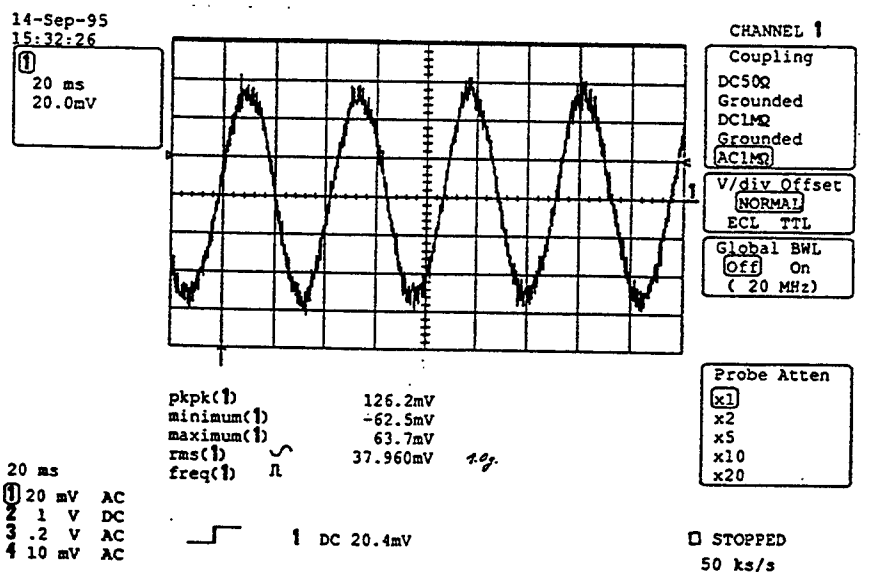


Figure 7c. 23Hz and 1.0g Acceleration Response.

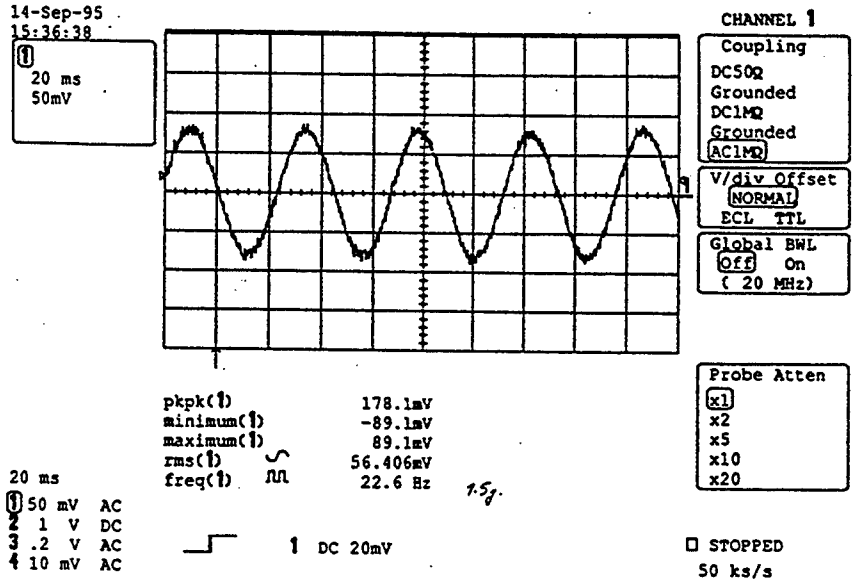


Figure 7d. 23Hz and 1.5g Acceleration Response.

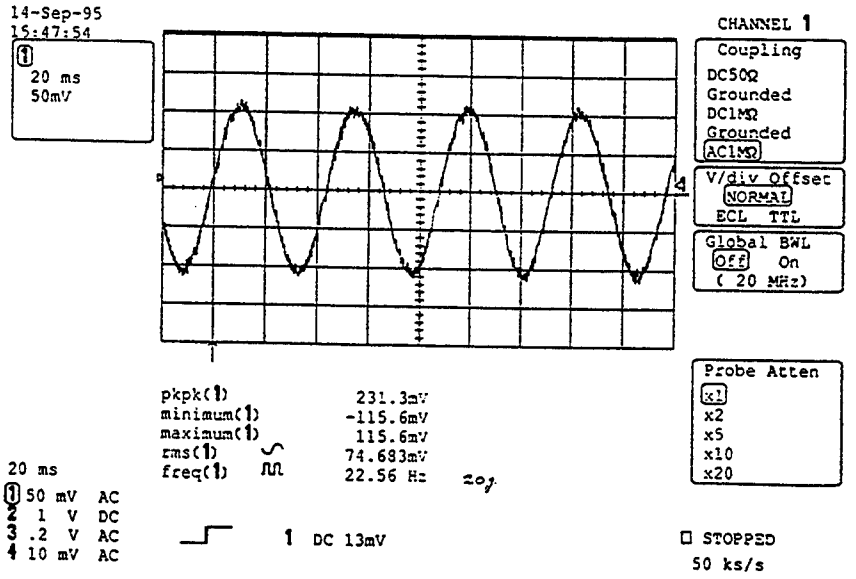


Figure 7e. 23Hz and 2.0g Acceleration Response.

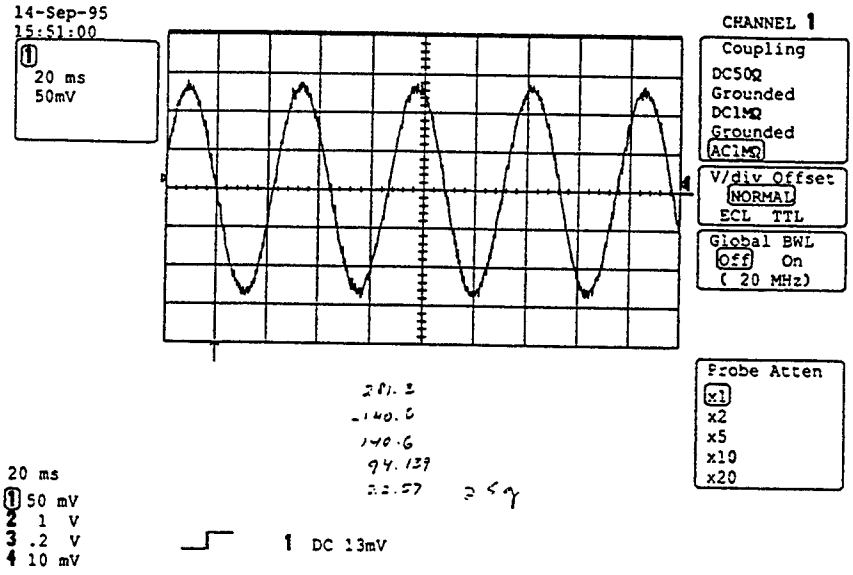


Figure 7f. 23Hz and 2.5g Acceleration Response.

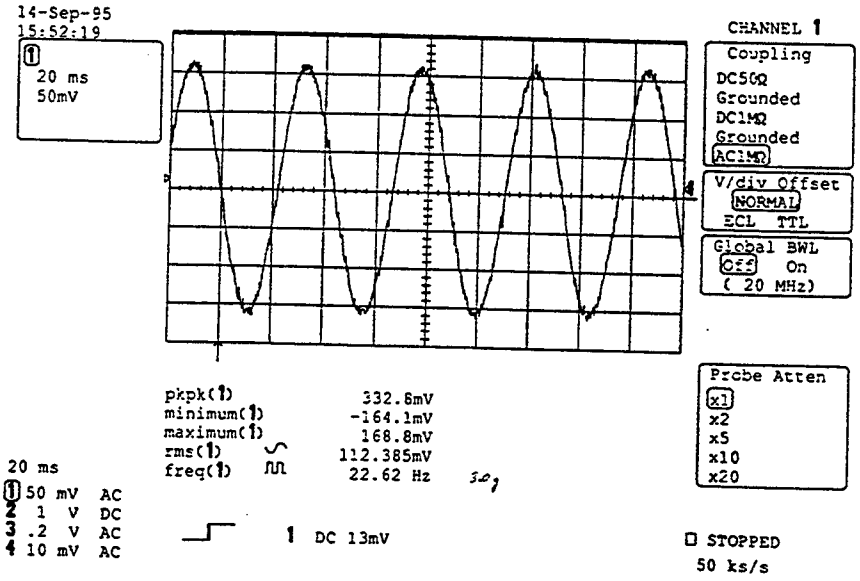


Figure 7g. 23Hz and 3.0g Acceleration Response.

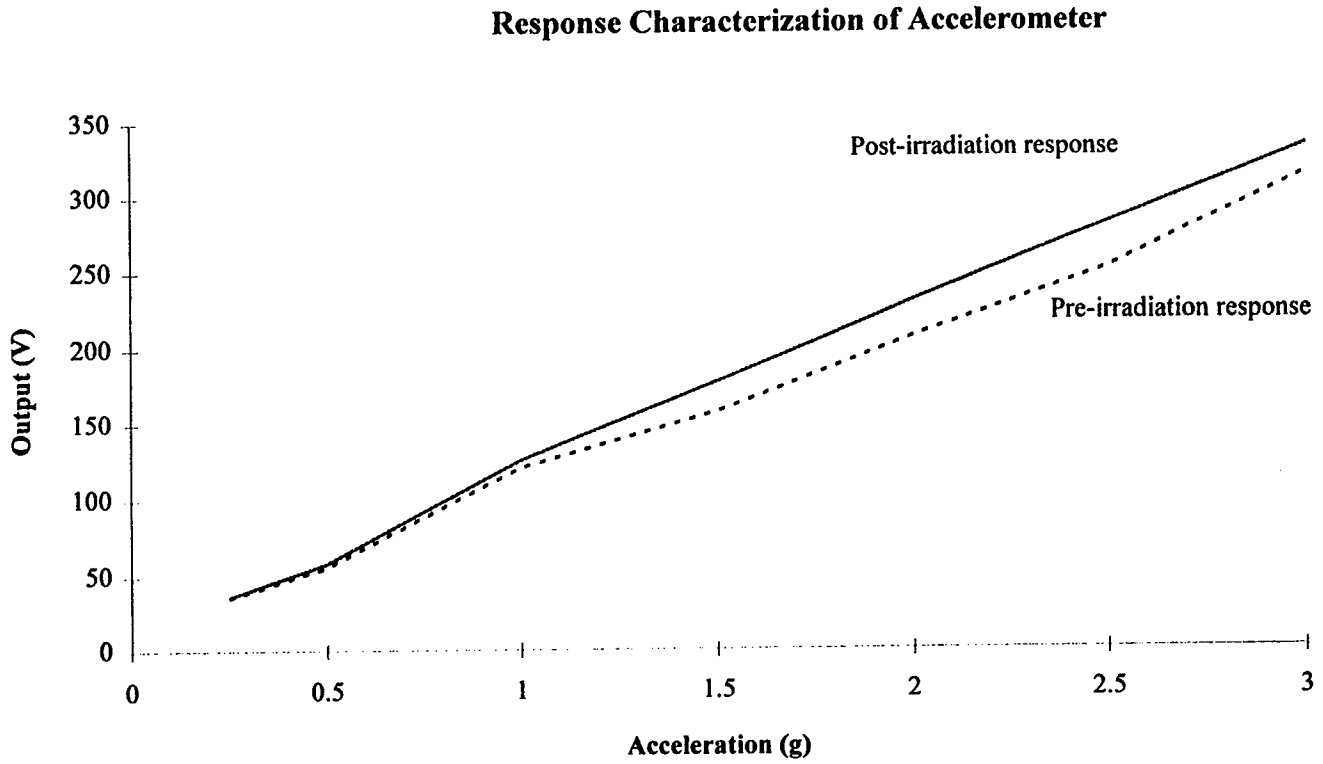


Figure 8. Accelerometer Response Before and After Being Irradiated.

TECHNICAL SOFTWARE DOCUMENTATION FOR PRFECT

Ann M. Kaul
PhD Student
Department of Mathematics and Statistics

University of New Mexico
Albuquerque, NM 87131

Final Report for:
Graduate Student Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Laboratory

August 1995

TECHNICAL SOFTWARE DOCUMENTATION FOR PRFECT

Ann M. Kaul
PhD Student
Department of Mathematics and Statistics
University of New Mexico

Abstract

The Predictive Radio Frequency Effects Coupling Tool (PRFECT) is a computer model currently being developed that will simulate the effects of a ground-based radio frequency (RF) source engaging a space system. PRFECT will provide an estimate of the probability of failure, degradation or upset of a receiver's performance by an RF source. One of the necessary components for validation and verification of a computer model such as PRFECT is accurate documentation of the processes which the model simulates. User's Manual for PRFECT - Volume II presents the statistical methodology used in the model.

TECHNICAL SOFTWARE DOCUMENTATION FOR PRFECT

Ann M. Kaul

Introduction

The purpose of User's Manual for PRFECT - Volume II is to provide users of the Predictive Radio Frequency Effects Coupling Tool (PRFECT) computer model with an explanation of the underlying physical and statistical principles applied by the model. Knowledge of these principles should enable users to conduct more effective and precise simulations of scenarios of interest. In addition, the manual provides analysts with a reference for the statistics and mathematics used in the PRFECT code.

Description of Model

PRFECT is a data-driven model that depends on user inputs to define various process control parameters. These inputs are provided through an engagement scenario that may include three primary aspects: a ground-based high power microwave (HPM) or jamming threat, a target satellite receiver system, and an uplink transmitter. The code processes these inputs through a chain of modules, each link of which uses physics-based algorithms with simplifications appropriate to an engineering code.

Where there are known or estimated variabilities or incomplete knowledge of the relevant inputs, a probabilistic description is given and carried through the chain of modules. The model then calculates the estimated effect and provides associated uncertainty bounds. The user may employ a combination of data, including empirical and subjective/engineering judgment, to describe the probabilistic behavior of these inputs.

PRFECT can address three types of problems caused by RF sources: component failure, component degradation and receiver bit-error-rate (BER) performance. The first two problems assess permanent effects on the receiver. The last problem assesses receiver performance and the effects of permanent degradation or

temporary interference. The type of problem being addressed in the engagement determines which user inputs are required. Definitions of the target and the threat are needed for all problem assessments. Definition of the uplink is required only for the receiver performance assessment.

User input controls the flow of model operations. PRFECT begins the simulation by calculating from the user-supplied information the incident power flux density (Watts/cm^2) from the ground-based threat(s) at the exterior of the space system. Next, PRFECT determines the probability density function (PDF) for source stressing power at a particular point in the satellite receiver system. For probability of failure and degradation assessments, PRFECT calculates the stressing power at the user-specified critical component. For receiver BER performance assessment, PRFECT calculates the stressing power at the point where the receiver attempts to recognize the signal, that is, at the demodulator. As part of the stressing power calculation, PRFECT also determines the power from the uplink command signal at the demodulator if a receiver BER performance assessment is being conducted. At this point, PRFECT can begin the assessment of the threat's ability to affect the space system.

The type of problem of interest to the user determines the next step in the modeling process. To assess probability of failure or degradation, estimates of the parameters of a strength function for the critical receiver component are calculated off-line and supplied to PRFECT. The parameters of the strength function for the critical component are based on user-supplied test data of the component's failure or degradation rate with respect to power or energy. The model then generates the PDF for failure effect (P_e) or degradation (d_D) by comparing the appropriate strength function to the stressing power. If the problem is receiver performance assessment, PRFECT generates the PDF for the system bit error rate (P_{BER}) by calculating the ratio of received energy per bit to noise density. P_{BER} can be calculated with or without taking into account degradation, interference or both.

Methods Used in PRFECT

PRFECT calculates the stressing power as the product of random variables and constants. The product may be calculated using Monte Carlo techniques or by analytical means using the method of moments.

For a probability of failure effects assessment, a weighted nonlinear least squares program fits a cumulative distribution function (CDF) to experimental failure effects data. The CDF is then weighted with a stress PDF, which provides a measure of the probability of failure effect for the problem. In PRFECT there are three options to evaluate this probability: a Monte Carlo technique, an analytical/numerical technique using method of moments, and maximum likelihood integration.

The gain or efficiency of RF components may also be degraded at lower stresses than those that cause failure. Given a linear fit to degradation data from a linear least squares program, PRFECT can calculate a PDF for probability of degradation. A probability of degradation can also be calculated for receiver designs that are different from the test structure from which the degradation data were taken, as long as the critical component is the same. PRFECT currently uses only a Monte Carlo technique to determine probability of degradation. A maximum likelihood integration technique is being developed.

PRFECT allows the user to investigate receiver performance in addition to failure and degradation. A PDF for carrier power and thus energy per bit is calculated. By determining a PDF for the noise power density, the model calculates a density function for the ratio of these two random variables. This ratio provides an estimate of the bit error rate for a particular modulation scheme. The results of a previous calculation of the PDF for degradation due to an HPM threat pulse may be used to evaluate the impact on BER. Similarly, a noise jamming signal may be introduced to evaluate its impact on BER. Calculations for probability of bit error rate can be done using either a Monte Carlo technique or a method of moments technique.

Conclusion

The goal of the PRFECT development effort is to provide a useful tool for determining the front-door susceptibility of space systems. The tool first must be based on general physical principles and physics models. Because of the uncertainties in environments, threat designs and system/subsystem/component variations, a statistical approach is an efficient way of handling multiple random variables and the interaction

with systems. An accurate description of the techniques used is necessary for PRFECT to become accepted as a valid simulation model.

My task this summer was to compile the information available into a rigorous mathematical and statistical description of the PRFECT simulation model. The resulting document provides the scientist, engineer or mathematician with a comprehensive and sound reference of the techniques used by the model.

In an effort to enhance the readability of the document, the manual was organized to follow the calculation flow of the model. The manual begins with a general description of the need for and use of the PRFECT simulation model, definitions of the probabilistic and statistical terminology used in the manual and an example scenario. This is followed by a description of the methods used by PRFECT to develop the stressing power PDF at the receiver, the sources of experimental data for failure and degradation of a critical component and the methods used to model the critical component's strength curves, and the three methods used to determine the probabilities of failure effect, degradation and bit error rate. Each section concludes with an application of the methods described in that section to an example scenario. The final section contains a brief summary of the PRFECT model.

Among the resources available for rewriting the manual were a draft of the manual and several documents expanding on the methods used in the model. After reading the available documents, I developed a plan of attack with Dr. William J. Zimmer and Dr. Richard R. Prairie of The University of New Mexico and Javon Evanoff of Ball Systems Engineering Operations. Each method was researched to develop a description of its use in the PRFECT model, then checked by the UNM/Ball team for accuracy. The draft manual was then edited to include these descriptions and to improve the flow of the document. Each section was reviewed by the UNM/Ball team as it was developed. I extend my sincere gratitude for their assistance to Dr. Zimmer, Dr. Prairie and Mr. Evanoff.

References

1. A. J. Alexander, G. E. Fauss, R. M. Clark and J. D. Black, "User's Manual for PRFECT, Volume II - Statistical Methodology," SSESS-1-MRC-004 (MRC/ABQ-R-1682), Mission Research Corporation, February 1995 (Draft).

2. J. C. Krainak and T. W. Larkin, "High-Powered RF Effects on Electronic Systems: A Statistical Model for Probability of Effect Characterizations," SSESS-4-META-014 (MRC/META-R-107), Mission Research Corporation, June 1994 (Draft).

3. T. W. Larkin, "Bayes' Theorem and the Probability of Effect," Metatech Corporation, 14 March 1995.

4. W. J. Zimmer, R. R. Prairie, and J. Evanoff, "A Discussion of Probability of Effect and Confidence with Respect to PRFECT Predictions," University of New Mexico and Ball Systems Engineering Operations, 17 January 1995.

5. D. J. Finney, *Probit Analysis*, 3rd ed., Cambridge University Press, 1971.

6. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, 1993.

7. Kalos, M. H., and P. A. Whitlock, *Monte Carlo Methods, Volume I: Basics*, John Wiley and Sons, New York, NY, 1986.

8. Shreider, Yu, *Methods of Statistical Testing: Monte Carlo Method*, Elsevier Publishing Company, Amsterdam, 1964.

AN ADVANCED PULSED PLASMA THRUSTER DESIGN STUDY USING ONE-DIMENSIONAL SLUG MODELING

Robert John Leiweke

Graduate Research Associate

Department of Aerospace Engineering,

Applied Mechanics, and Aviation

The Ohio State University

2036 Neil Avenue Mall, Room 326

Columbus, Ohio 43210

Final Report for:

AFOSR Summer Research Program

Phillips Laboratory

Sponsored by:

Air Force Office of Scientific Research

Edwards Air Force Base, CA

November, 1995

AN ADVANCED PULSED PLASMA THRUSTER DESIGN STUDY USING ONE-DIMENSIONAL SLUG MODELING

Robert John Leiweke
Graduate Research Associate
Department of Aerospace Engineering,
Applied Mechanics, and Aviation
The Ohio State University

Abstract

Design goals for an Advanced Pulsed Plasma Thruster (APPT) are established with respect to the current state-of-the-art LES-8/9 PPT. One main goal is to achieve a thrust efficiency of at least 16% and I_{sp} greater than 1500s using a 70J discharge. Historical data suggests that between 70μg to 130μg of solid Teflon fuel may be ablated to give an $I_{bit} \approx 1500\mu N\text{-}s$. It is known that quasi-steady operation may increase the efficiency of pulsed electromagnetic thrusters if the dynamic impedance is matched with that of the power supply. A circuit utilizing an inductive energy store to release an exponentially decaying quasi-steady current pulse of characteristic time L/R has been selected for one-dimensional numerical slug mass modeling. Preliminary results indicate that I_{sp} and overall efficiency decreases with increasing store inductance and slug mass. The one-dimensional model shows that total thrust efficiencies of 33%, 25%, and 22% are achievable at the 70J level for specific fuel consumption rates of 1.0μg/J, 1.57μg/J, and 1.86μg/J, while I_{sp} values vary between 1000s and 2300s. Both time (μs) and position (cm) of the maximum slug de-acceleration scales almost 1:1 with the current e-fold time. This suggests that short pulse times with relatively low storage inductances ($\frac{1}{2}\mu H$ to 1μH) are favorable for high efficiencies.

AN ADVANCED PULSED PLASMA THRUSTER DESIGN STUDY USING ONE-DIMENSIONAL SLUG MODELING

Robert John Leiweke

Design Goals

Design goals for an Advanced Pulsed Plasma Thruster (APPT) are established with respect to the current state-of-the-art LES-8/9 device. As shown in Table 1, two main design criteria for the APPT are thrust efficiencies at least twice that of the LES-8/9 while dropping the total "wet" mass in half. The specific impulse should be between 1000s and 2000s, while the average power requirement between 2 W and 200 W. The total impulse must lie within the $3x10^4$-$5x10^4$ N-s range. The energy per discharge, E_0, based upon LEO stationkeeping requirements, is about 70 J.

	LES-8/9	Advanced PPT
Overall η_T, %	8	$\geq$ 16
I_{sp}, s	1000	1000 - 2000
Average Power, W	25 - 50	5 - 200
Total I_{bit}, N-s for N $\equiv$ $20x10^6$ shots	7,320	$3x10^4$ - $5x10^4$
Wet Mass, kg	6.6	$\leq$ 3.5
Energy per Discharge, J	20	70

Table 1

The storage capacitor chosen for use in The Ohio State University prototype APPT has a 10μF capacitance, 70 nH - 90 nH parasitic inductance, and is rated for 5000 V at 25 kA peak current (pulsed) at a maximum voltage reversal of 20%. Operation at 3,750 Volts (70J), gives a predicted lifetime of $3.2x10^6$ shots.

Impulse Bit and Ablation Mass

Figure 1 shows a trend in impulse bit as a function of discharge energy for breech-feed, Teflon fuel laboratory and flight qualified PPT's.[1] A least squares fit of the data shows the trend to be linear. From this plot, it seems reasonable to expect a PPT impulse bit from a 70 J discharge to be within the 1000 μN-s to 2000 μN-s range, as shown in Figure 1.

A somewhat linear trend between mass loss and discharge energy is shown in Figure 2. In Figure 2, the slope of a linear fit between the LES-6 and LES-8/9 (76.5J) corresponds to a specific fuel consumption (sfc) of 1.14μg/J. Note the mass loss varience at the 4J level (Chinese study).[2] This is due to parametric changes in both fuel face surface areas and electrode geometries. In that study, it was shown that the ablated mass increased with increasing fuel face area and aspect ratio. However, at the 23J level, the mass loss varience became less sensitive to these parametric changes. Based upon the trend shown in Figure 2, it is reasonable to study APPT performance using mass values in the range 60μg to 130μg. It is worth mentioning that the LES-8/9 PPT operating at 76.5J ablated 98μg/shot with $I_{bit} \approx 1,347 \mu N\text{-}s$.[3]

PPT Slug Modeling

Simple PPT "slug" models assume that all of the mass to be accelerated is ablated and fully ionized during the initial breakdown event and that no mass is lost or gained as the plasma moves down the electrode channel.[4] The arc is assumed to be a moving conductor. Figure 3 shows the general features of a "lumped" parameter LRC circuit for a rectangular geometry, breech-feed PPT.[5] Here, the capacitor has some small amount of internal resistance and inductance, R_c and L_p. The plasma is assumed to have lumped properties giving rise to a time

varying resistance and inductance, L(t) and R(t) including an initial inductance, L_0, due to the backstrap. The power delivered by the capacitor discharge is[4]

$$P = I^2R + \frac{d}{dt}\left(\frac{1}{2}LI^2\right) + \frac{1}{2}I^2\frac{dL}{dt} \quad , \tag{1}$$

where R represents the equivalent resistance of the LRC circuit and time varying plasma resistance. The three terms from left to right represent (1) the rate of resistive heat generation, (2) the rate of change of energy stored in the magnetic flux, and (3) the rate at which work is done on the moving arc-slug.[4]

The one-dimensional force acting on the slug in the x-direction is

$$F(t) = m\,\frac{du}{dt} \tag{2}$$

where t is time, m is the slug mass, and u=dx/dt is the slug speed. The rate at which work is done on the moving arc-slug is F·u, so from Equation 2

$$P_{slug} = F\,u \equiv \frac{1}{2}I^2\frac{dL}{dt} \quad . \tag{3}$$

Note that the dynamic impedance, dL/dt, may be written as

$$\frac{dL}{dt} = \frac{dL}{dx}\,\frac{dx}{dt} = \frac{dL}{dx}\,u(t) \quad , \tag{4}$$

so by combining Equations 2 and 3 the slug acceleration is

$$\frac{du}{dt} = \frac{I^2}{2m}\frac{dL}{dx} \quad . \tag{5}$$

The current is obtained by solving the nonlinear ODE

$$(L_p + L)\,\ddot{I} + (2\dot{L} + R)\,\dot{I} + \left(\ddot{L} + \frac{1}{C}\right)I = 0 \tag{6}$$

numerically (4^{th} Order Runge-Kutta Method) with initial conditions

$$I(0) \equiv 0$$

$$\dot{I}(0) = \frac{V_o}{L_o + L_p} \quad . \tag{7}$$

The maximum current is given approximately by

$$I_o \propto \frac{V_o e^{-\frac{\alpha\pi}{4\omega}}}{\omega\,(L_o + L_p)} \tag{8}$$

where ω is the angular frequency (from the linear ODE) and $\alpha \equiv (L_0 + L_p)/R$. The time derivative of Equation 4 gives

$$\ddot{L} = \dot{u}\frac{dL}{dx} + u^2\frac{d^2L}{dx^2} \quad . \tag{9}$$

PPT Efficiency

Energy losses due to Teflon fuel vaporization, dissociation, and ionization also effect PPT efficiency. Figure 5a shows the results of a PPT exhaust gas analysis using a mass spectrometer.[6] The PPT used in that study ablated 10μg of Teflon by discharging a 2μF capacitor at 1.6kV (2.5J) within a 4.0m³ vacuum chamber pumped down to $1.0\text{x}10^{-6}$ Torr.[6] In Figure 5a, particles with a mass-to-charge ratio of 12, 19, 31, 50, and 69 correspond to singly ionized C, F, CF, CF_2, and CF_3.[6] However, it is believed that up to 60% of the ablated fuel is composed of neutral Teflon monomers, C_2F_4, because observed impulse bits are lower than expected for pure electromagnetic acceleration.[3,8] Other experiments[7] suggest that doubly and triply ionized C and F with varying lifetimes are created during the discharge.[7] With the total energy lost to vaporization, dissociation, and ionization be defined as

$$E_{loss} \equiv E_{vap} + E_{dis} + E_{ion} \quad , \tag{10}$$

A "worst-case" value for these energy losses may be computed by first assuming that after vaporization of a given Teflon mass, m, (211 kJ/mol) all of the monomer double carbon bonds are broken (614 kJ/mol) so that only CF_2 remains. Adhering to mass conservation and the spectrometer data, these molecules are then broken into the mole-fraction distribution of species shown in Figure 5b (C-F bonds: 485 kJ/mol, F-F bonds: 154 kJ/mol). All of these particles are then assumed to be singly ionized ($\phi \approx 11$ eV/molecule). Using the value of E_{loss} obtained in this manner, the instantaneous thrust efficiency is then defined as

$$\eta_T \equiv \frac{\frac{1}{2}mu^2}{E_{loss} + \frac{1}{2}CV_o^2} \quad . \tag{11}$$

The Inductance Models

The inductance of two planar electrodes under the condition that field fringing effects are negligable ("Aspect Ratio" AR ≡ h/w ≪ 1) is[4]

$$L(x) = \mu_o \frac{h}{w} x \tag{12}$$

where h and w are the electrode separation distance and width. From Equation 11, the one-dimensional inductance gradient, dL/dx, is $\mu_0 \cdot h/w$. In practice, however, all breech-feed PPT's with rectangular electrode geometries have $1 \leq AR \leq 4$, making fringe fields non-negligable. It is important to obtain realistic inductance values because the slug acceleration is proportional to dL/dx (see Equation 5). An alternate model is shown in Figure 4, where the electrodes and slug-arc are considered to be a rectangular strip of rectangular conductor of thickness δ on the order of the magnetic field diffusion depth based on the pulse time.[9] The inductance in μH is[9]

$$L = A\left[(h + x)\log_{10}\frac{2hx}{w + \delta} - \log_{10}[(h + g)^h (x + g)^x]\right] + B\left[2g - \frac{x + h}{2} + C(\delta + w)\right], \quad (13)$$

where A≡0.02339, B≡0.01016, C≡0.447, δ=1.0 mm, and g≡$\sqrt{[x^2+h^2]}$. Hereafter, Equation 11 will be referred to as the "*simple model*" and Equation 12 as the "*fringe model*". Comparison shows that the typical fringe model varies linearly with electrode position for AR ≥ 1, although the gradient is less than half that of the "simple" model. An estimate of fringe effects can be found in literature for the LES-6 PPT.[10] Here, a correction factor, f, to the one-dimensional inductance is obtained by best fit with experimental data.[10] Defined as the ratio of the actual inductance gradient to that of Equation 11, the best fit gives f≈0.253.[10]

LRC Slug Model Results

Two PPT's from literature were used as a bench test for an LRC slug code; the LES-6 and the LES-8/9.[3,5] For both cases current/voltage waveforms, internal resistances and inductances, and electrode geometries were used as guidelines to reproduce quoted performance values.

The LES-6 PPT has parallel electrodes spaced 3 cm apart, 1 cm wide, and 0.6cm long (AR=3).[3] It has a 2μF capacitor charged to 1,360 V (1.85J), and ablates 10μg/shot giving an I_{sp} of about 312 s.[3] The external circuit inductance and resistance (plasma resistance not included) were measured to be about 34nH and 30 mΩ.[3] Experimentally derived current and voltage waveforms are shown in Figure 6a. For this electrode geometry, simple and fringe model inductance gradients are 3.77μH/m and 1.11μH/m, respectively. Recall that a fringe field correction factor for the LES-6 PPT was f≈0.253.[10] Here, the ratio of fringe-to-

simple model gradients is 0.294 (within 16%). Using the "simple" inductance model, the total resistance and inductance producing a best fit with the experimental waveforms shown in Figure 6a are 80mΩ (50mΩ plasma) and 45nH. The resistance value correlates well with the measured time varying circuit impedance data.[3] In Figure 6b, note that the slug reaches the electrode end at 2.0μs, which corresponds to I_{sp}≈400s. This I_{sp} is is larger than the experimental value probably because of the "re-strike" hitch in the current waveform at 0.25μs. In order to aquire the proper I_{sp} value using the slug code "fringe" model, the total equivalent circuit resistance had to be lowered to about 30mΩ (effectively zero plasma resistance). However, the current and voltage waveforms were severely underdamped and did not resemble any of the waveforms in Figure 6a.

The LES-8/9 PPT has parallel electrodes 2.54 cm apart, 2.54 cm wide, and 2.54 cm long (AR = 1).[5] It has a single 17μF capacitor charged to 1,544V (20J) and ablates 28.5μg/shot, giving an I_{sp}≈1000s.[5] Additional experimental data was also taken at 76.5J; 95μg/shot gives an I_{sp}≈1,440 s.[5,11] The total discharge inductance and resistance was measured to be about 35nH and 45mΩ, respectively.[11] The "simple" and "fringe" models give dL/dx as 1.25μH/m and 0.57μH/m. Figure 7a shows the experimentally derived voltage waveform (a current trace was unavailable) along with "simple" model numerical results for E_0=20J. Figure 7b shows that the slug attains an I_{sp}≈1010s at the electrode end (2.54cm) in 3.8μs. In the 76.5J case, the LRC code using the "simple model" yields an I_{sp}≈1,350s, which is 6% lower than the experimental value. Note in Figure 7b that the uncoupled waveform is nearly identical to the coupled, indicating that the dynamic impedance (≤12mΩ) is negligable compared to the external circuit impedance (≈65mΩ).

The Advanced PPT Power Source

To date, the external power source of most flight qualified PPT's have been LRC circuits. The resulting current waveforms tend to be oscilliatory pulses that decay with e-fold times of 5μs to 15μs. The dynamic portion of the circuit impedance attributed to the moving arc-slug ($Z_{slug}=u \cdot dL/dx$) must be closely matched to the output impedance of the external power supply, $\sqrt{[L_0/C + R^2]}$, in order to achieve an effective energy transfer from source to plasma.[4] This tends to be difficult using an LRC circuit because the arc-slug speed is not constant.

An additional problem with the LRC driving circuit is the tendancy for secondary "crowbar" breakdowns to occur back at the fuel face at a time slightly prior to circuit current reversal.[4] This phenomena occurs when residual propellant material near the fuel face attains sufficiently high electrical conductivity to provide an arc path with lower impedance than the slug position.[9] Elimination of the oscilliatory behavior is a practical way to preclude this phenomena.[4] Efficient PPT operation may be realized by use of a non-inverting, *quasi-steady* current waveform.[4,12]

Quasi-steady operation using a pulse-forming LRC network proves to be difficult because of large parasitic inductances that discourage impedance matching. Additionally, the energy densities of off-the-shelf capacitors lead to excessive power supply specific mass. Inductive energy storage as an alternative to the LRC circuit has been proposed for use with the APPT.[12] Relatively high energy densities of 500 - 1000 J/kg and high Q-factors (L/R_{coil} » pulse time) of 10 to 50 are achievable. The inductive store yields an exponentially decaying current pulse of e-fold time $\tau_{LR}=L_0/R$, where L_0 and R are the storage inductance and total resistance (including the plasma resistance), respectively. For example, to achieve a 100μs quasi-steady pulse with a 20mΩ

circuit resistance, $L_0 \approx 100\mu s \cdot 20m\Omega \approx 2\mu H$. Since the energy stored within the inductor is $E_s = \frac{1}{2} L_0 \cdot I_{max}^2 \equiv 70J$, the maximum current is $I_{max} \approx \sqrt{[2 \cdot 70J \div 2\mu H]} \approx 8.4kA$. Note that this value is one third of the 25 kA limit on the capacitor chosen for use in The Ohio State University APPT prototype.

The APPT Circuit Model

The inductive energy store converts a significant portion of magnetic flux into directed kinetic energy as long as $dL/dt \geq R$. When the energy within the main inductor is maximum ($\frac{1}{2} L I_{max}^2$), some means of switching the resulting high current is necessary. Use of a plasma as the conducting switch media not only allows for high current but also complies with a desire for mass savings.[12] This plasma closing switch may also have the additional benefit of providing thrust if it is also a PPT.[12] Figures 8a-c show one such arrangement of dual PPT's, each represented as lumped-parameter circuits connected together to form "loop 1" and "loop 2". PPT_1 has inductance $L_1(t) \equiv L_{1o} + L_{d1}(t)$, where L_{1o} is the static inductance of the main toroidal store and backstrap, while $L_{d1}(t)$ is the time varying portion. PPT_2 has inductance $L_2(t) \equiv L_{2o} + L_{d2}(t)$ where L_{2o} is the initial static inductance due to the backstrap. The lumped resistance and capacitor parasitic inductance, denoted as R_1, R_2, and L_p, are considered to be constant.

During the inductor "charging" phase shown in Figure 8a, PPT_1 is ignited using a typical spark-gap discharge system. The loop 1 current I_1 reaches a maximum given approximately by Equation 8 at time $t \approx \pi/(2 \cdot \omega)$. Ignoring the small inductance change of L_1, the voltage across the main inductor is approximately $L_1 \cdot dI_1/dt$ while that of the current is $-C \cdot dV_c/dt$, where V_c is the capacitor voltage. Numerical integration of Equations 6 and 7 gives the exact solution.

As shown in Figure 8b, the "switching" phase occurs by discharging PPT_2

(closing the loop 2 switch) at a time when the capacitor voltage decreases to a small negative value ($V_c \approx -100V$). This tends to drive a negative current in both loops since $\Delta I_2 \approx (-|V_c| \cdot \Delta t)/L_2$ and $\Delta I_1 \approx -|V_c| \cdot \Delta t \cdot (1/L_1 + 1/L_2)$, where Δt is a small incremental time-step after the switch. The slug mass within PPT_1 accelerates due to the (positive) net current $I_N \equiv I_1 - I_2$. Since loop 1 now provides a higher impedance path than loop 2, V_c and I_1 quickly decay so that the current through both PPT's is $\pm I_2$, an exponentially decaying pulse of e-fold time $\tau_{LR} \approx L_1/(R_1+R_2)$ shown in Figure 8c. The loop currents I_1 and I_2 are determined by simultaneous numerical integration of

$$(L_p+L_1)\ddot{I}_1 + (2\dot{L}_1+R_1)\dot{I}_1 + (\frac{1}{C}+\ddot{L}_1)I_1 = L_1\ddot{I}_2 + (2\dot{L}_1+R_1)\dot{I}_2 + \ddot{L}_1 I_2$$

and (14)

$$L_2\ddot{I}_2 + (2\dot{L}_2+R_2)\dot{I}_2 + \ddot{L}_2 I_2 = -L_p\ddot{I}_1 - \frac{1}{C}I_1$$

with switch conditions

$$I_2|_{t_s} \equiv 0$$

$$\dot{I}_1|_{t_s} = \left.\frac{V_c(L_1+L_2) - L_2(\dot{L}_1+R_1)I_1}{L_p(L_1+L_2)+L_1L_2}\right|_{t_s} \qquad (15)$$

$$\dot{I}_2|_{t_s} = \left.\frac{L_1\dot{I}_1+(\dot{L}_1+R_1)I_1}{(L_1+L_2)}\right|_{t_s}$$

and Equation 5 for both PPT's. In the case that switching occurs shortly *before* capacitor voltage reversal (i.e., at +100V), then I_2 rapidly becomes much larger than the value I_0 given by Equation 8. This current spike along with a 50% voltage reversal will severely reduce the capacitor life.

To understand the dependance of the APPT thrust efficiency on the circuit parameters, consider a situation where the dynamical impedance is small with

respect to the total circuit impedance (i.e., "un-coupled"). In this case, one may approximate the current waveform in advance, leaving only Equation 5 to solve independent of Equations 14 and 15.[4] The PPT_1 current is $I_N(t) \approx I_0 \cdot e^{-t/\tau}$, where $\tau \equiv (R_1+R_2)/L_{10}$. The slug speed becomes

$$u_e \approx \frac{I_o^2 L_1 \frac{dL}{dx}}{4m_1 (R_1 + R_2)} \quad . \tag{16}$$

With thrust efficiency defined as $\frac{1}{2}m \cdot u_e^2 \div \frac{1}{2}CV_0^2$ along with Equation 5, then

$$\eta_T \sim \frac{(\frac{1}{2} CV_o^2)}{2 m_1 L_1} \left[\frac{\frac{dL}{dx}}{R_1 + R_2} \right]^2 \quad . \tag{17}$$

Although PPT_2 maximum current is larger than PPT_1, it has the same pulse shape, so Equations 16 and 17 also apply.

APPT Slug Model Results

Three cases of specific fuel consumption (sfc) values based upon the estimated mass loss range at 70J (see Figure 2) were used for the APPT study; sfc ≡ 1.00μg/J, 1.57μg/J, and 1.86μg/J. The lumped PPT parameter values were chosen as $R_1 \equiv 30m\Omega$, $R_2 \equiv 40m\Omega$, $L_{20} \equiv 20nH$. Capacitor parameters were chosen as $L_p \equiv 60nH$, $C \equiv 10\mu F$, and $V_0 \equiv 3.75kV$. In all cases the "simple" inductance model for AR≡3 is used. For each sfc case, four L_{10} inductor values (hence, L/R e-fold decay times) were tested; 0.5μH (7.1μs), 1.0μH (14.3μs), 2.0μH (28.6μs), and 4.0μH (57.1μs). It should be noted that at the time of PPT_2 ignition, up to 25% of the original capacitor energy, E_0, is used to accelerate the PPT_1 slug and lost to resistive dissipation. Thus, the PPT_2 slug mass is estimated from the inductively stored energy; $m_2 \equiv sfc \cdot \frac{1}{2}L_1 I_{max}^2$. Table 2 lists the mass values used in the three cases.

Figures 9a-c show APPT results for Case 2 with $L_{1_0} \equiv 1\mu H$ (14μs pulse). Current and voltage profiles in Figure 9a show that the capacitor voltage reversal is 27% with some oscillatory behavior in the L/R pulse. Figure 9b indicates maximum $I_{sp1} \approx 1500s$ and $I_{sp2} \approx 1200s$ attained at about $t \approx 20\mu s$ leading to electrode lengths $x_1 \approx 20cm$ and $x_2 \approx 25cm$, respectively. In accordance with Equation 11, these I_{sp} values give $\eta_1 \approx 15\%$ and $\eta_2 \approx 8\%$ for a total $\eta_T \approx 23\%$. Figure 9c shows the dynamic impedances reach about 50mΩ at 20μs, not quite matching the 70mΩ circuit resistances but doing better than the LRC circuit PPT's. It is important to mention that the trajectory and I_{sp} curves remained basically the same in all cases studied, with only relative shifts in magnitude and time.

Case	PPT_1 Mass, μg	PPT_2 Mass, μg	Total Mass, μg
1	70	55	125
2	110	88	198
3	132	103	235

Table 2. PPT1 and PPT2 Masses

Table 3 shows Case 2 performance for varying storage inductance values. Both I_{sp} and η decrease with increasing L_{1_0} in accordance to Equation 17.

L_{1_0}, μH	I_{sp1}, s	I_{sp2}, s	η_1, %	η_2, %	η_{total}, %
0.5	1,627	1,190	18	7	25
1.0	1,473	1,162	15	7	22
2.0	1,316	1,065	12	6	18
4.0	1,117	917	9	4	13

Table 3. m_1=110μg, m_2=88μg, Values at PPT_2 Length ≡ 15cm

The best I_{sp} and η_{total} values for Cases 1 and 3 were exhibited with $L_{1o}\equiv 0.5\mu H$. Case 1 gave $I_{sp1}\approx 2{,}350s$ and $I_{sp2}\approx 1{,}600s$ with $\eta_{total}\approx 33\%$ while that of Case 3 was $I_{sp1}\approx 1{,}400s$ and $I_{sp2}\approx 1{,}050s$ with $\eta_{total}\approx 22\%$.

Figure 10a shows that PPT_1 maximum current is practically invarient to the slug mass but decreases with increasing L_{1o} inductance. This is expected since the energy within the inductor is almost always the same ($\approx 54J$), $I_{1max}\approx 10.4\div\sqrt{L_{1o}}$. Note that the currents are less than the 25kA limit on the 10μF capacitor. Maximum I_{sp} variations with L_{1o} and total mass is shown in Figure 10b. Note that for any given L_{1o}, I_{sp} tends to decrease with increasing mass m_1 in accordance to Equation 17 because of inertial effects. However, I_{sp1} tends to *increase* slightly (contrary to Equation 17) with L_{1o} while I_{sp2} *decreases*.

Figures 11a,b show the time and position of maximum slug de-acceleration ("critical" position and time) as a function of current pulse e-fold time. Using the case shown in Figure 9b as an example, this time is at approximately 15μs for both PPT_1 and PPT_2. Figure 11a shows that the critical position scales at a rate of ~0.8 cm/μs with e-fold time. Figure 11b shows the critical time is nearly 1:1 with the e-fold time for all three cases. Again, this suggests that low inductor values of ½μH to 1μH (7μs to 14μs e-fold) are favorable from the viewpoint of electrode mass savings (the mass of an electrode of constant width scales linearly with length).

Concluding Remarks

A one-dimensional slug model based upon an L/R decaying current pulse shows that total thrust efficiencies of 33%, 25%, and 22% are achievable at the 70J level for specific fuel consumption rates of 1.0μg/J, 1.57μg/J, and 1.86μg/J. Both time (μs) and position (cm) of maximum slug de-acceleration scales almost

1:1 with the current e-fold time. This, along with the observation that I_{sp} and η_T are most affected by slug mass suggests that small L_{10} values are favorable as long as the maximum current is kept within the capacitor's operating range.

References

[1] Myers, R.M., *Electromagnetic Propulsion for Spacecraft*, AIAA paper 93-1086.

[2] Yuan-Zhu, K., *Effects of Propellant Geometry on PPT Performance*, IEPC paper 84-94, 1984.

[3] Vondra, R.J., Thomassen, K., and Solbes, A., *Analysis of Solid Teflon Pulsed Plasma Thruster*, AIAA paper 70-179, 1970.

[4] Jahn, R.G., *Physics of Electric Propulsion*, McGraw-Hill Series in Missle and Space Technology, McGraw-Hill Book Co., New York, 1968.

[5] Vondra, R.J. and Thomassen, K.I., *Flight Qualified Pulsed Electric Thruster for Satellite Control*, AIAA paper 73-1067, 1973.

[6] Hirata, M. and Murakami, H., *Exhaust Gas Analysis of a Pulsed Plasma Engine*, 17th IEPC paper 84-52, JSASS/AIAA/DGLR, Tokyo, Japan, 1984.

[7] Thomassen, K.I. and Vondra, R.J., *Exhaust Velocity Studies of a Solid Teflon Pulsed Plasma Thruster*, AIAA paper 71-194, 1971.

[8] Thomassen, K.I. and Tong, D., *Interferometric Density Measurements in the Arc of a Pulsed Plasma Thruster*, AIAA paper 72-463, 1972.

[9] Terman, F.E., *Radio Engineer's Handbook*, 1st Ed., McGraw-Hill Book Co., New York, 1943, pg. 53.

[10] Solbes, A. and Vondra, R.J., *Performance Study of a Solid Fuel-Pulsed Electric Microthruster*, AIAA paper 72-458, 1972.

[11] Vondra, R.J., *M.I.T. Lincoln Laboratory Notes*, Book 1, Oct.-Nov., pg. 152.

[12] Turchi, P.J., *private communications*, 1995.

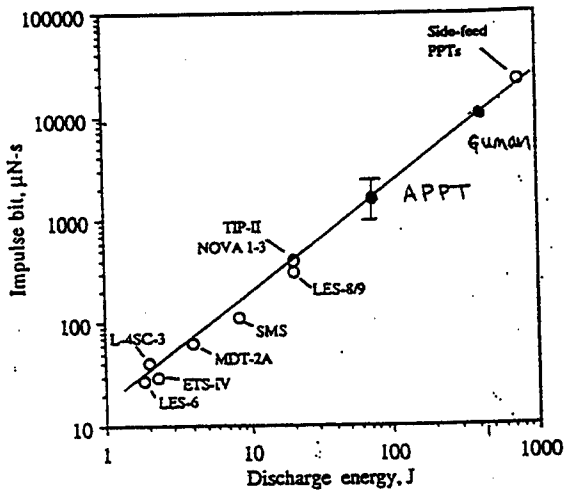


Figure 1

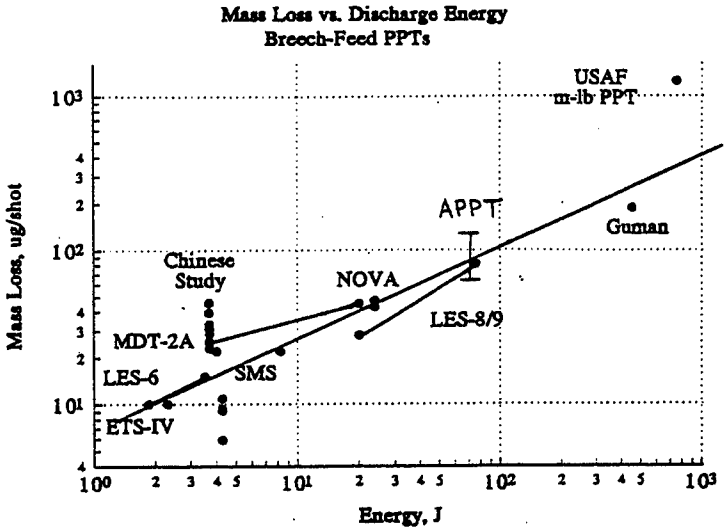


Figure 2

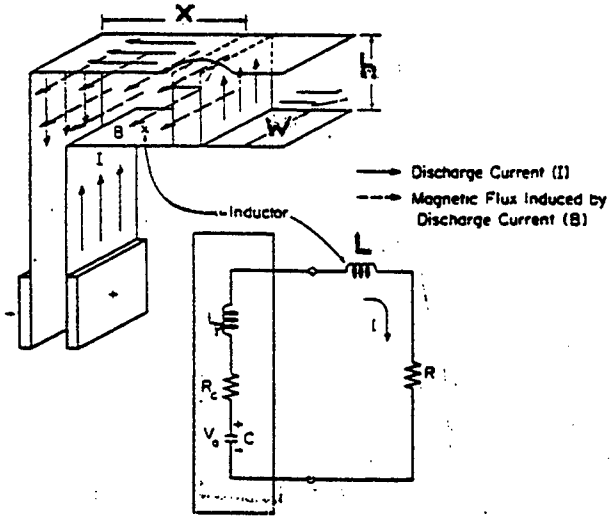


Figure 3

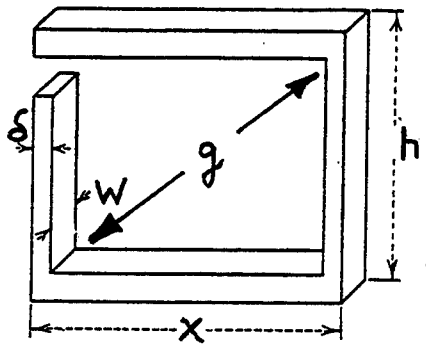


Figure 4

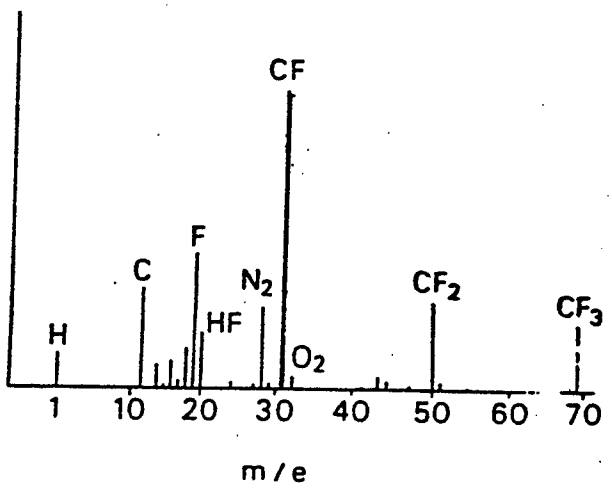


Figure 5a

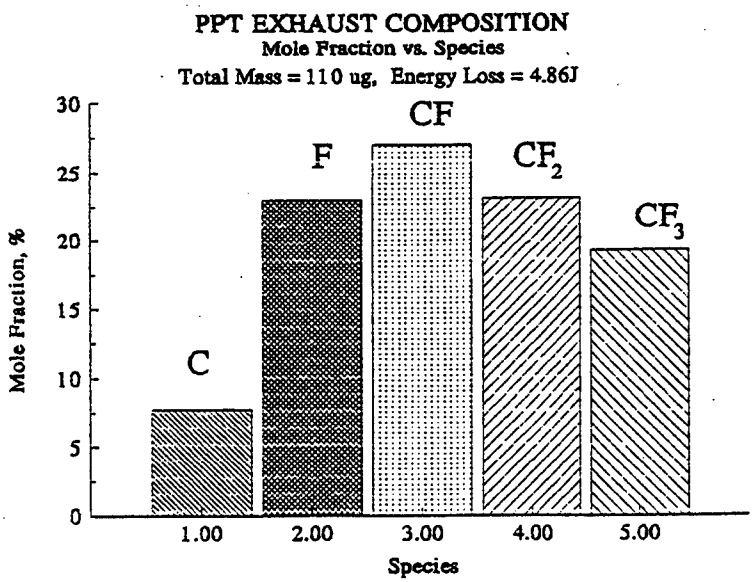


Figure 5b

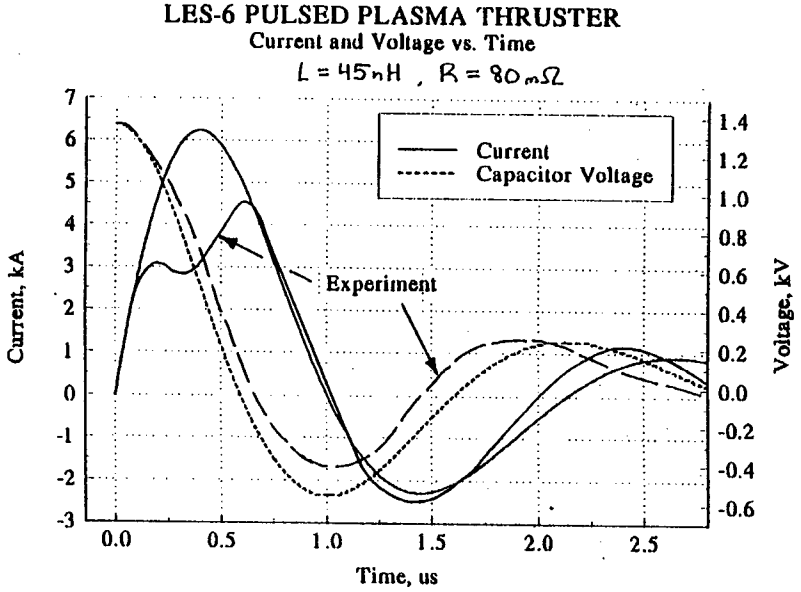


Figure 6a

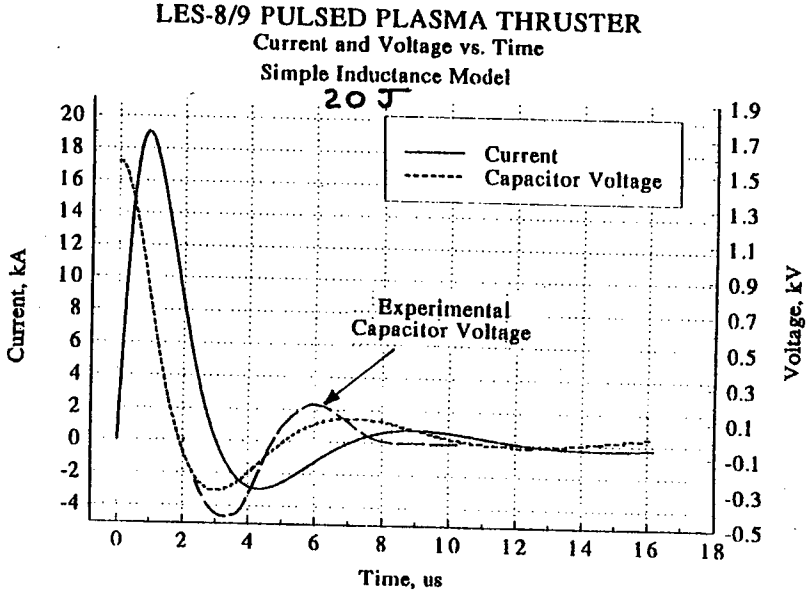


Figure 7a

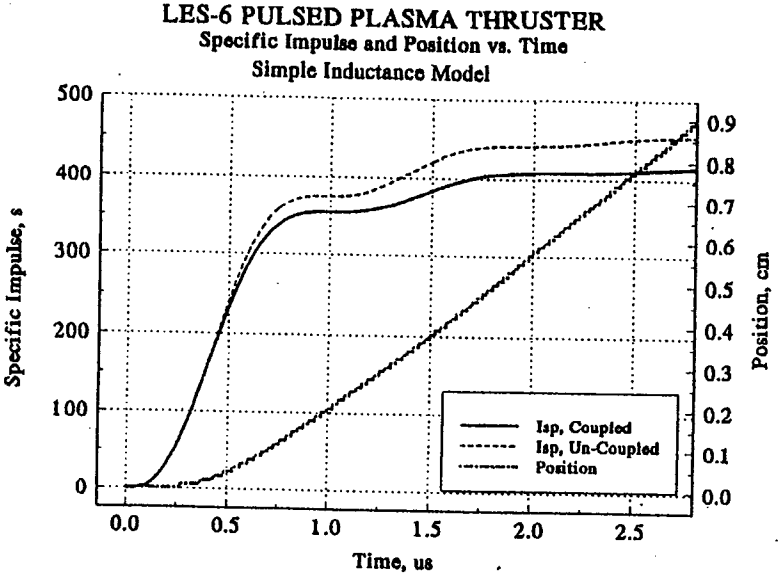


Figure 6b

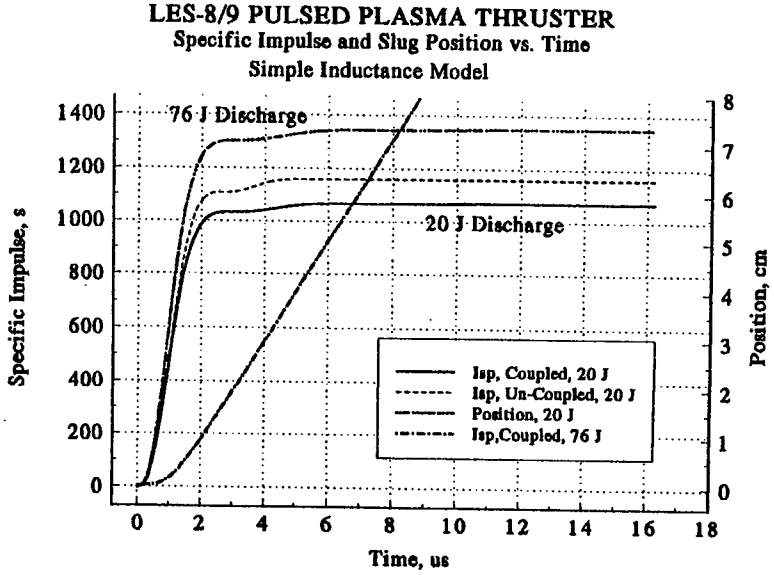


Figure 7b

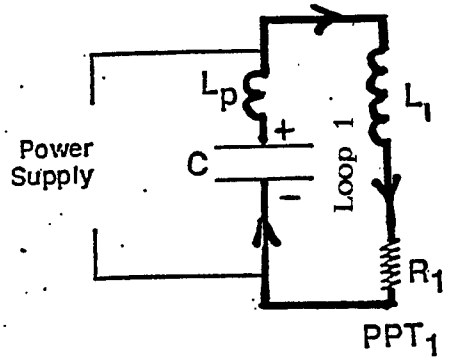


Figure 8a "Charging Phase"

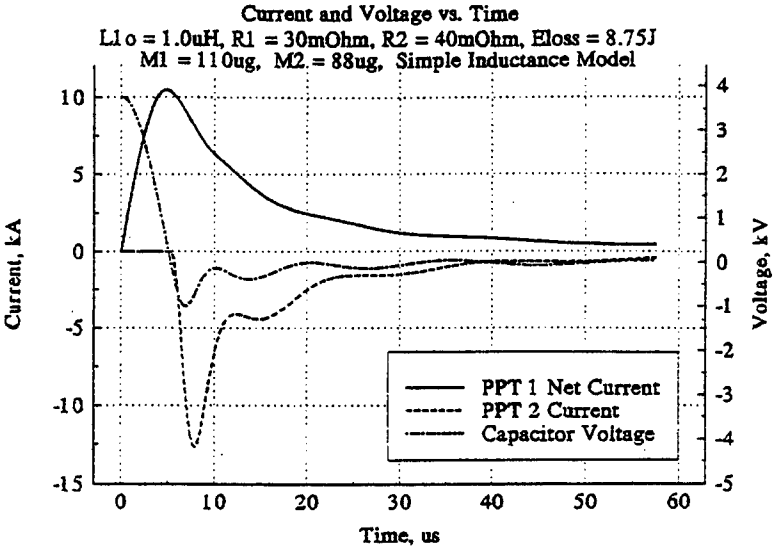


Figure 9a

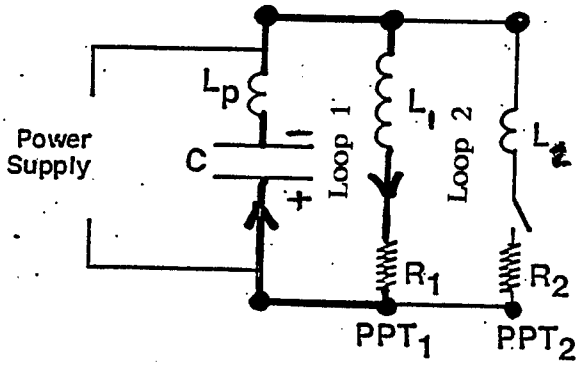


Figure 8b "Switching Phase"

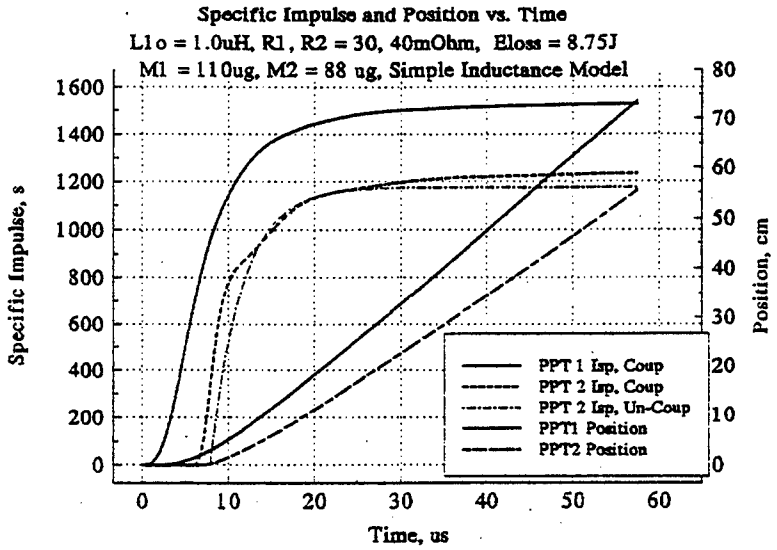


Figure 9b

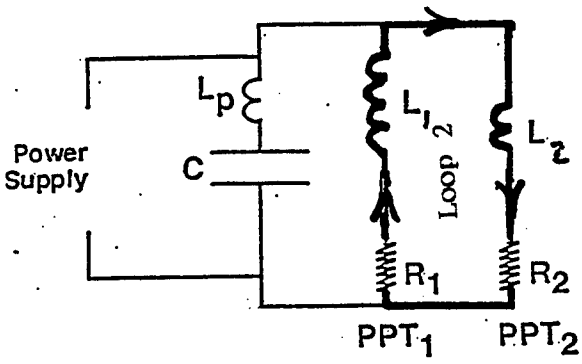


Figure 8c "L/R Decaying Pulse"

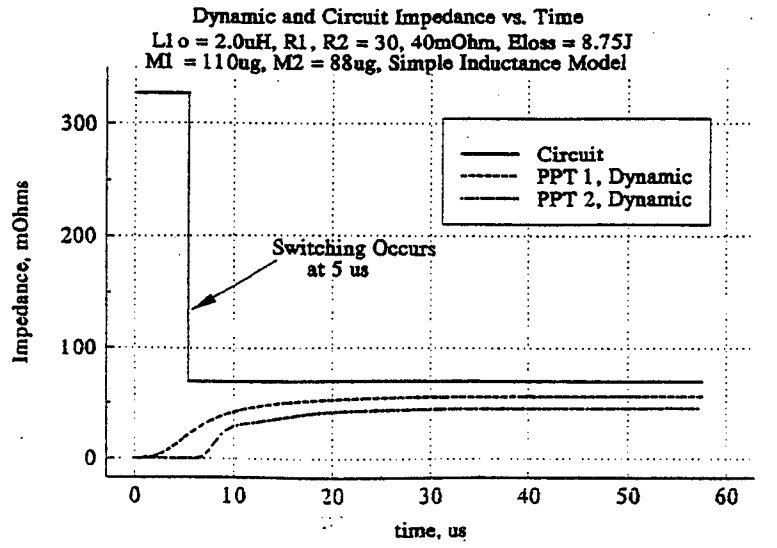


Figure 9c

THE OHIO STATE ADVANCED PULSED PLASMA THRUSTER
Maximum Current vs. Store Inductance

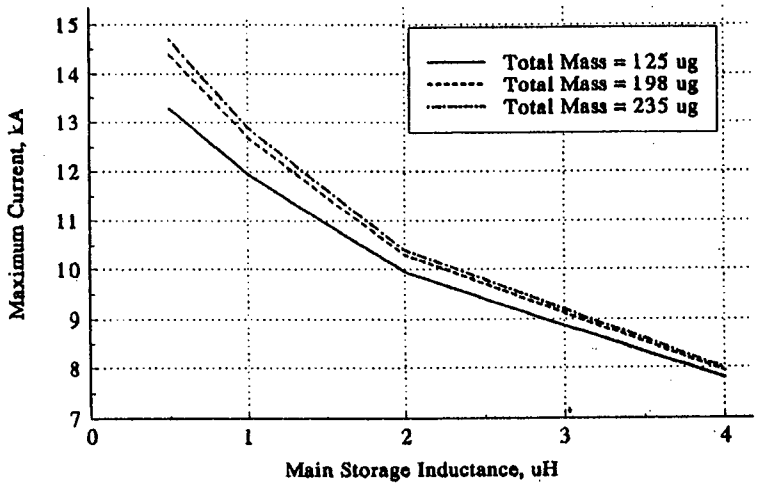


Figure 10a

THE OHIO STATE ADVANCED PULSED PLASMA THRUSTER
Position of Maximum De-acceleration vs. E-Fold Time

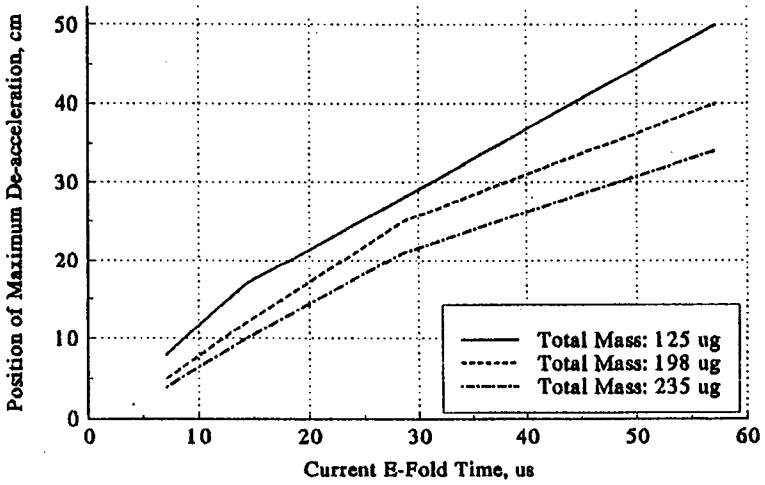


Figure 11a

THE OHIO STATE ADVANCED PULSED PLASMA THRUSTER
Maximum Specific Impulse vs. Store Inductance

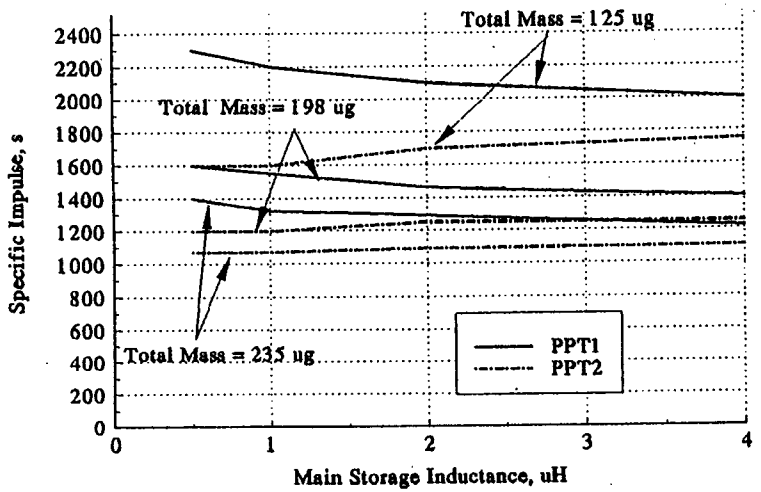


Figure 10b

THE OHIO STATE ADVANCED PULSED PLASMA THRUSTER
Time of Maximum De-acceleration vs. E-fold Time

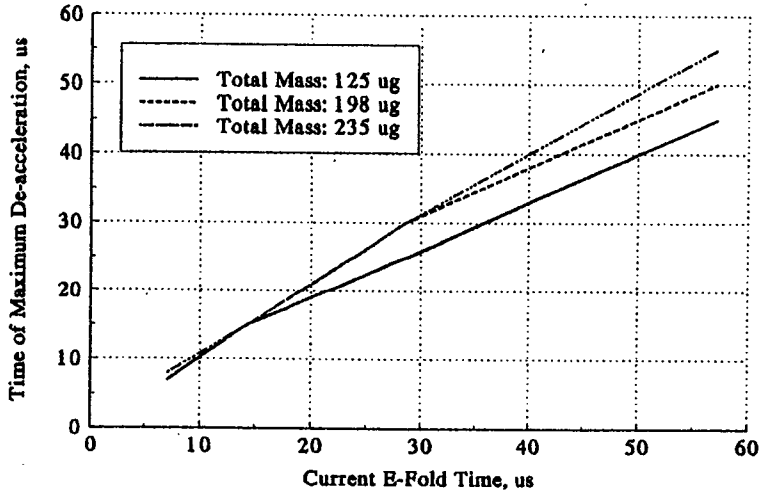


Figure 11b

NEW SIMPLE METHOD OF GETTING SUPERRESOLUTION
NEW METHOD OF OBTAINING POLARIZATION PROPERTIES

David Maker
Graduate Student
Electrical Engineering and Applied Physics

Oregon Graduate Institute
PO Box 91000
Portland Oregon
97291-1000

Final Report for:
Summer Graduate Student Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Laboratory

September 1995

NEW SIMPLE METHOD OF GETTING SUPERRESOLUTION

David Maker
Graduate Student
Electrical Engineering and Applied Physics
Oregon Graduate Institute

Abstract

In this paper we show how to get superresolution by writing down the Kirchhoff integral for a wide aperture and a narrow aperture. We solve for the pupil function in the narrow aperture Kirchhoff integral and put it in the pupil function for the wide aperture. Also we curve fit a polynomial to the actual image at the screen of the narrow aperture. This polynomial must be spatially compressed to take into account the fact that the wide aperture pupil function contains higher spatial frequencies. We have also done simulations of this superresolution using this algorithm and they indicate that this type of superresolution works.

NEW SIMPLE METHOD OF GETTING SUPERRESOLUTION

David Maker

Introduction:

Analyticity of the Fourier Spacial Transform of an Image

Recall that the fourier transform of a dirac delta source is a constant and spread over all spacial frequencies. But measurment devices cannot measure all spacial frequencies, so clearly information about a finite object is lost by finite aperture devices(or is it?). Thus there is a socalled "cutoff" frequency associated with the device. But if the function is piecewise continuous, Riemann integrable, and bounded in amplitude the integral (spectrum) is analytic. But a "function of a complex variable is determined thoughout the entire Z-plane from a knowledge of its properties within an arbitrarily small region of analyticity"[2] . Thus if there is exact knowledge of the spectrum over any region we can determine the spectrum over all regions, so that we can determine the spectrum well beyond the cutoff. We then have a resolution that is **limited only by the Signal to Noise Ratio** (SNR) of the system and so virtually unlimited resolution especially for bright objects because of their large SNR s. **The time is now ripe to take advantage of this technique using the increasily higher SNRs that are available with space telescope technology.** What is exciting here is that the **resolution is not limited by the size of the aperture here**, *only(obviously) by the SNR*[4] , which as we said, is large for bright objects and so is also not a limiting factor here. This would not be true for ground based telescopes because of atmospheric turbulence.

The alternative methods of superresolution such as the poisson map algorithm are dealt with elsewhere[5] . I believe this method is more straightforward than the entropy maximization methods and will yield greater resolution than the standard superresolution iterative algorithms. For example scattered planetary images could easily be misconstrued for lessened entropy and thrown out using these algorithms. Other issues, such as whether the image here is "ambiguous", are dealt with below elsewhere[6] also. It is important to note here though that these researchers, generally, have just rehashed the above observation that the method is limited by the signal to noise ratio (SNR), which is obvious, since the G curve would not have the ideal shape given by the object if noise is included and the resulting calculations of the image N(x,y) would be faulty. But all conceivable methods are limited by the SNR.

Signal to Noise Ratio and The Criteria for the Use of Superresolution

T=Tolerance in terms of spectrum magnitude of the second peak as percent of the 0 frequency level (first peak) in the spectrum. This value the astronomer sets. Might be 5% for double vs star cluster test.

C_1 = the ratio of the peak of the sidelobe to the central peak for the null object which we take to be about 1/14 for a source image that occupies a small field of view.

ω_c = the cutoff spacial frequency for the object.

σ_n=noise power. For good receiver system this is just shot noise.

χ= object extent

ω_e =maximum extropolation frequency. This is the largest spacial frequency that you can meaningfully extrapolate to.

Using methods discussed elsewhere[7] that *take into account* the ambiguous image problem we can show that

(8) $\omega_e - \omega_c = \frac{3}{2\chi}\left[\sin c^{-1} C_1 - \sin c^{-1}\left(\frac{C_1 T}{\sigma_n}\right)\right]$ where the $(\omega_e - \omega_c)$ term on the right side represents the bandwidth extension. The larger it is, the better the superresolution. Note that the bandwidth extension is additive with respect to ω_c .Note also the presence of the tolerance T in the sinc function. Also note though that if the noise σ_n is very, very small (long exposure, high quality optics) then the second sinc function is very nearly zero and **T tolerance plays no role.** For example if $T = 5\sigma_n$ and $C_1 \approx 1/14$ plugging into equation (8) we get:

$\omega_e - \omega_c = \frac{3}{2\chi}\left[\sin c^{-1} C_1 - \sin c^{-1}\left(\frac{C_1 T}{\sigma_n}\right)\right] = \frac{3}{2\chi}(6.79 - 2.22) = \frac{6.85}{\chi}$ .Note (given the properties of the sinc function) that for a given tolerance, if your noise σ_n is very small (so in the above 2.22 is replace by 0) you should be able to get a bandwidth extension of $1.5*6.8/\chi = 10/\chi$. If the cutoff frequency is of the order $/\chi$ which in our applications it probably is (ie., the whole rayleigh criterion region in our field of view) then essentially the bandwidth extension is about $10\omega_c$ and basically you have about 10 times the rayleigh criterion resolution. **A 4 meter telescope would have the resolution usually stated for a 40 meter telescope!**

Note that equation (8) is only valid if $T \le \frac{\sigma_n}{C_1}$ and so:

9) $\sigma_n < T \le \frac{\sigma_n}{C_1}$ so

(10) $1 < \frac{T}{\sigma_n} \le \frac{1}{C_1}$ Note that with C_1 at about 1/14 for a small image and $T \approx .05$ we can **specify the allowable noise** in our system for meaningful superresolution. So 14*.05=.7maximal signal power. In our algorithm we find it convenient to put our own noise in the system using a random number generator to simulate an SNR.In a noisy system it may be possible to use the standard superresolution algorithms (such as poisson map estimation) or wavelet filtering algorithms after the above algorithm is used so our polynomial doesn't violate one of our POCS constraint conditions(eg.,jump to negative infinity values). One limitation we must state here is that the instrument sampling rate ω_s must be at least twice the cutoff frequency ω_c for accurate extrapolation to be possible. So if the cuttoff frequency is 100 then the number of pixels along a line should be 200.

Simple Superresolution

A direct way of getting superresolution is to set up two Kirchoff integrals, one for a wide aperature and one for a narrow aperature. We note the fourier transform nature of the relation between the position on the pupil (x) versus the position of the screen (X) in the Kirchoff integral anologous to ω and t in the paradigm frequency-time fourier transform. Thus we write:

(11) $\int_w f(x) e^{ikxX} dx = U_{LR}(X)$

(12) $K\int_W f(x) e^{ikxX} dx = U_{HR}(X)$ where "W" means wide aperature, "w" means narrow aperture, HR stands for High Resolution, LR is short for Low Resolution, f(x) is the pupil function..Because of the

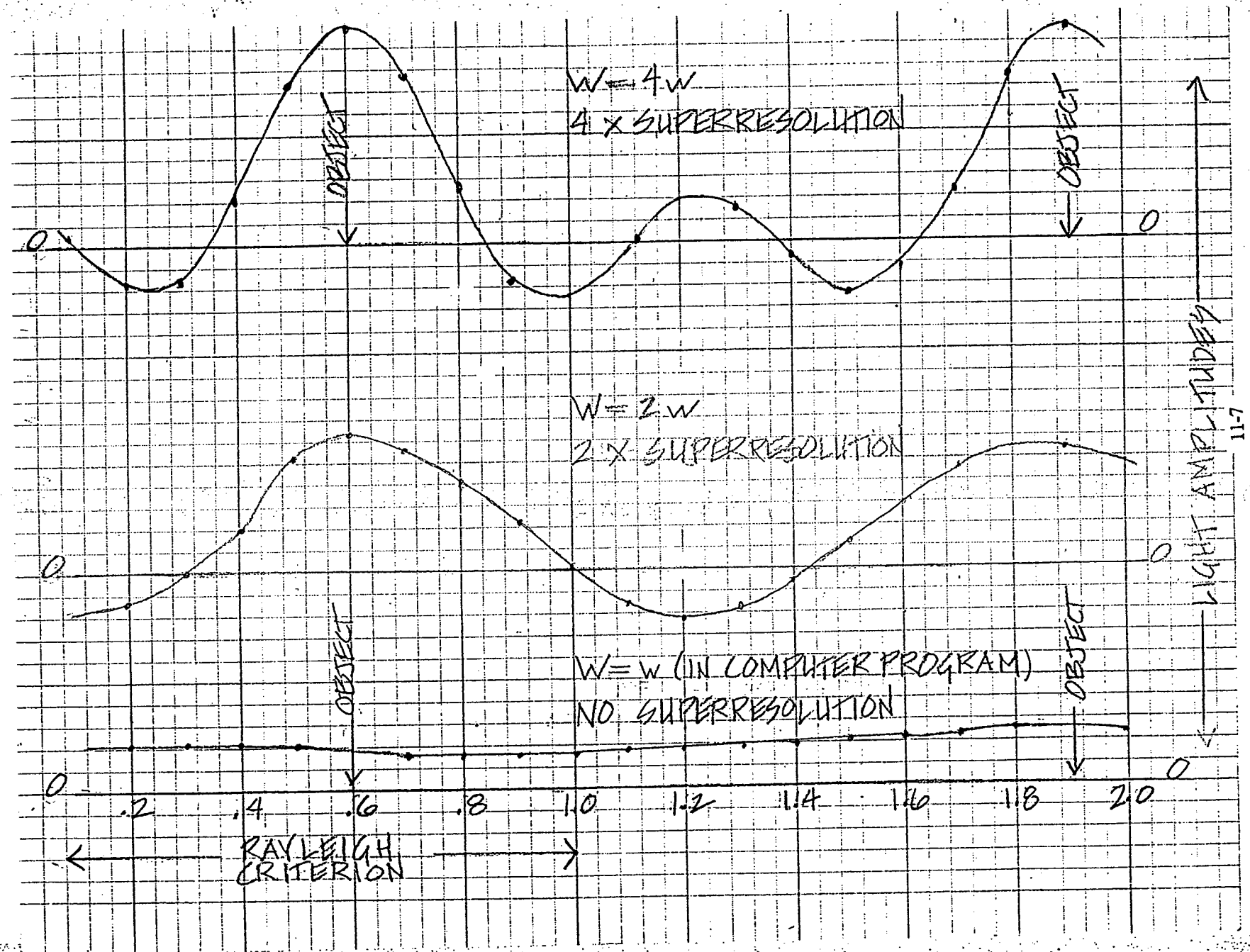
fourier transform nature of the relation between x and X we can solve for the pupil function in equation (11) and plug the result into equation (12) giving us:

(13) $K\int_W f(x)e^{ikxX}dx = U_{HR}(X) = K\int_W\left(\int_w U_{LR}(X)e^{-ikxX}dX\right)e^{ikxX}dx = U_{HR}(X)$. We fix the constants k and K by making sure that $U_{HR} = U_{LR}$ if W=w. Note the interesting possibility exhibited in equation 13 of getting a *High resolution image from the low resolution image* in equation 11. There is however one complication here. The f(x) s in equation 11 and 12 are not the same except in the the width w. The f(x) in equation 12 includes the higher spacial frequencies of $U_{HR}(x)$ created by the higher resolution of detail Thus the key to this type of superresolution is to compensate for the differences in these two f(x) s by spatially compressing the width of U_{LR} by an amount w/W in doing the integral in equation 13 and so putting in by hand these spatial frequencies. Our simulation does this by using sinc(kx) functions and a W/w scaling of the "x" dependence. *For real data we fit a polynomial and do the same "x" scaling on the terms of the polynomial.* Here we include some graphs from the resulting simulations. Thus by changing "W" we change the ammount of superresolution. Thus it is easy to increase the superresolution and note how little distortion there is in the final product. For example with no superresolution the test light sources are barely resolvable as separate sources and they are obviously very sensitive to low SNR s.. With superresolution turned on they become sharper as the artificial aperture size "W" is made larger. SNR s on the order of one still allow for this type of superresolution.

---

[2] Guillemin,"The Mathematics of Circuit Analysis"(John Wiley & Sons,Inc)
[4] Nadar,S, "Analysis of the Limit to Superresolution in Incoherent Imaging"SPIE Vol.2308, pp.1244
[5] Sementilli.P.,Analysis of the Limits of Superresolution...",Opt. Soc.of Am.Vol.10,No.11 Nov.1993,P2265
[6] Toraldo di Francia,"Resolving power and Information"J.Opt.Soc.Am.,45,497(1955)
[7] Sementilli.P,"Analysis of limit of..."Vol.10,No.11/Nov.1993/J. Opt. Soc. .Am.. A.

RAYLEIGH CRITERION
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
OBJECT
OBJECT
W=W
W=2W
W=4W
W=8W
0
0
0
1
LIGHT AMPLITUDES


W = 4w
4 x SUPERRESOLUTION
OBJECT
0
W = 2w
2 x SUPERRESOLUTION
OBJECT
0
W = w (IN COMPUTER PROGRAM)
NO SUPERRESOLUTION
OBJECT
0
0
.2
.4
.6
.8
1.0
1.2
1.4
1.6
1.8
2.0
RAYLEIGH CRITERION
LIGHT AMPLITUDE

```
$ for l.for
      PROGRAM LITE
      INTEGER Q ,X,Y,A,Z,G,F,W
      DOUBLE PRECISION XX1,XX2,XY2,XA2, A2,A3,A4,A5,A6
      DOUBLE PRECISION  ARP,ARP1,ARP2,ARN,ARN1,ARN2
      DOUBLE PRECISION  X1, X2, XXX1, XXX2,XXY2, PP2, PP3
      DOUBLE PRECISION N1,N2,N3,N4,N5,N6,SNR,U,N9
      DOUBLE PRECISION C(200),B(100),T(200,200)
      G=100
      W=200
      WH=W*1.0
      F=100
      WL=F*1.0
      A2=1.0
      A3=0.0
      PP2=1.9
      PP3=0.0
      XX2=0.0
      SNR=100000000000.0
      DO 18 A=1,20
      XA2=.1*A
C     CCD electron counts
C     OPEN(10,file='num.dat',FORM='FORMATTED',STATUS='old')
C     n equal pixel number read in
C     n=50
C     A9=0.0
C     DO 3 I=1,n
C     READ (10,*) C(I)
C     J=201-I
C     C(J)=C(I)
C     P1=C(I)
C     A9=A9+P1
C 3   CONTINUE
C     DO 8,I=1,n
C     B(x)=A9/(P*1.0)
C     C(I)=C(I)/W(x)
C     x1=.01*I
C     DO 7,J=1, 50
C     T(I,J)=x1EJ
C 7   CONTINUE
C     C(J)=C(J)/W(x)
C  8  CONTINUE
C     CALL GAUSS(D,T,C)
C     CALL BSOLVE(D,T,Z,C)
```

```
      DO 15 X=1,W
      X1=.1*X-(WH/20.0)
      DO 12 Y=1,F
      X2=.1*Y-(5.0)
      ARP=3.14159*((WH/WL)*(X2+0.6)+.00000000001)
      ARN=3.14159*((WH/WL)*(X2-0.6)+.00000000001)
      N1=1.0*sin(ARP)/ARP
      N2=1.0*sin(ARN)/ARN
C     With real data, lines down to N6 gone
      ARP1=3.14159*((WH/WL)*(X2+PP2)+.00000000001)
      ARN1=3.14159*((WH/WL)*(X2-PP2)+.00000000001)
      ARP2=3.14159*((WH/WL)*(X2+PP3)+.00000000001)
      ARN2=3.14159*((WH/WL)*(X2-PP3)+.00000000001)
      N3=(sin(ARP1)/ARP1)*A2
      N4=(sin(ARN1)/ARN1)*A2
      N5=(sin(ARP2)/ARP2)*A3
      N6=(sin(ARN2)/ARN2)*A3
C     Data polynomial XY2, Scaled by WL/WH.
C     XY2=0.0
C     DO 11 N=1,100
C     XY2=(Z(N)*(WH/WL)*X1E+N)+XY2
C     CONTINUE
C     RANDOM NUMBER GENERATOR
      U=0.0
C     N9=Q*1.0
C     N9=(N9*.843314861+.453816693)
C     U=N9*.4656612
      XY2=((1.0)/(2.0))*(N1+N2+N3+N4+N5+N6+(U*.5/SNR))
      XX2=(XY2*cos(((2*3.14159*X1*1.0)+.0000001)*X2))+XX2
 12   CONTINUE
      XX1=(XX2*1.0*cos(((-2*3.14159*X1*1.0)+.0000001)*XA2))+XX1
      XX2=0.0
 15   CONTINUE
      PRINT * ,XA2,XX1
      XX1=0.0
18    CONTINUE
20    END
```

```
C       SUBROUTINE BSOLVE(G,V,X,D)
C       INTEGER I,J,Q
C       REAL V(100,100), X(100), D(100),SUM
C       DO 200 I=G, 1, -1
C       SUM=D(I)
C       DO 100 J=I+1,G
C       SUM=SUM-V(I,J)*X(J)
C 100   CONTINUE
C       X(I)=SUM/V(I,I)
C 200   CONTINUE
C       RETURN

C       SUBROUTINE GAUSS(G,L,M)
C       INTEGER J,K,PIVOT, I,INDEX
C       DO 100 I=1,G
C       INDEX=I
C       PIVOT=ABS(L(I,I))
C       DO 200 J=I+1,G
C       IF (ABS(L(J,I)) .GT. PIVOT) THEN
C       PIVOT=ABS(L(J,I))
C       INDEX =J
C       ENDIF
C 200   CONTINUE
C       IF (INDEX .GT. I) THEN
C       DO 400 K=I,G
C       TEMP=L(I,K)
C       L(I,K)=L(INDEX,K)
C       L(INDEX,K)=TEMP
C 400   CONTINUE
C       TEMP=M(I)
C       M(I)=M(INDEX)
C       M(INDEX)=TEMP
C       ENDIF
C       DO 300 J=I+1,G
C       RATIO= L(J,I)/L(I,I)
C       DO 500 K=I+1,G
C       L(J,K)=L(J,K)-L(I,K)*RATIO
C 500   CONTINUE
C       M(J)=M(J)-M(I)*RATIO
C 300   CONTINUE
C 100   CONTINUE
C       RETURN
C       END

[EOB]
```

NEW METHOD OF OBTAINING POLARIZATION PROPERTIES USING MUELLER MATRICES

David Maker
Graduate Student
Electrical Engineering and Applied Physics
Oregon Graduate Institute

Abstract

When imaging an object we would like to be able to not only use phase and intensity information but also polarization information. For this purpose we obtain polarization transfer functions we call **BRDFs** We will be able to add 7 Mueller matrix two dimensional maps on top of the usual intensity map of the object we are imaging. Thus we will have about 7X the information concerning the nature of the surface of the object than with just a simple intensity map

NEW METHOD OF OBTAINING POLARIZATION PROPERTIES
USING MUELLER MATRICES

POLARIZATION TRANSFER FUNCTIONS USING MUELLER MATRICES

David Maker

Introduction

Here we list some examples of useful things polarization might tell you about an image. One simple example is that one can determine whether a reflecting surface is dihedral or not. Dihedral surfaces(ie., two intersecting planes) effect LCP and RCP [5] light in different ways. A dihedral surface reflects incident RCP light as RCP light; the return from a dihedral target will then be strongly rejected by the LCP filter. **Thus one might map the macroscopic surface curvature by using RCP and LCP light.**

Another more spectacular example is the effect of almost total depolarization. It can be shown[3] that

$$p(I)=\frac{1}{2}(1+\mathrm{P})\exp\left[-\frac{2I}{<I>(1+\mathrm{P})}\right]-\frac{1}{2}(1-\mathrm{P})\exp\left[-\frac{2I}{<I>(1-\mathrm{P})}\right]$$

Where P is the polarization, I an intensity , <I> the average intensity, and p is the probability that the intensity will be above the value I. On the next page are values of p plotted at various values of I using a basic program. The above equation provides the algorithm for this program. One thing to note from this out put is that the intensity drops rapidly as $\mathrm{P}\rightarrow 0$. Thus a **depolarizing surface destroys the speckle pattern** or at least decreases the intensity by so much as to decrease the SNR by a large ammount and make the speckle pattern useless.

If we wanted to have a computer algorithm to automatically include all of these effects we need a unifying principle that incorporates all of these cases and perhaps others. That principle is the Bidirectional Reflectance Distribution Function (BRDF). We would like to be able to incorporate an algorithm(to be coded) that can be *turned on whenever we would like to measure not just intensity but also the other 3 stokes parameters.*

Review of Stokes parameters

Let E_x be the real component of the electric field in the X direction and E_y be the corresponding Y component. When then define:

$\frac{E_x}{E_y}=\frac{a_2}{a_1}e^{-i\delta}$ in general . For a linear[2] polarized wave $e^{-i\delta}=(-1)^m$ so $\delta=m\pi$ where m=0,±1,±2,..

For right handed circularly polarized wave $\delta=\frac{\pi}{2}$. For lefthanded circularly polarized wave $\delta=-\frac{\pi}{2}$.

Thus a_1 and a_2 are like amplitudes of E_y and E_x and δ is like the phase difference between E_x and E_y. .To characaterize the polarization ellipse three independent quantities are needed. Here we will use the Stokes parameters to characterize the field. Thus we can write:

$$S_0 = a_1^2 + a_2^2$$

$$S_1 = a_1^2 - a_2^2$$

$$S_2 = 2a_1a_2\cos\delta \; S_3 = 2a_1a_2\sin\delta$$

Also $s_0^2 = S_1^2 + s_2^2 + s_3^2$ so that essentially we have three free[1] parameters, s_0 represents the intensity of the wave and s_3 represents ellipticity of the wave. The stokes paramters can be measured using a polarizer and a quarter wave plate..

Let $I(\theta, \varepsilon)$ represent intensity of the light E field vibrations in the direction making an angle θ and ε making a phase retardation of the y component relative the x component. Thus to obtain the stokes parameters we need two polarizers and a quarter wave plate set up in such a way that:

$$S_0 = I(0°,0) + I(90°,0)$$

$$S_1 = I(0°,0) - I(90°,0)$$

$$S_2 = I(45°,0) - I(135°,0)$$

$$S_3 = I\left(45°,\frac{\pi}{2}\right) - I\left(135°,\frac{\pi}{2}\right)$$

But what can we use the stokes parameters for? For light that has been beamed to a surface and reflected back to a receiver we use them to calculate the the relative amplitudes of the Mueller Matrice terms and use the Mueller matrix amplitudes to understand the nature of the reflecting surface.

Algorithm For Calculating Mueller Matrices and So Determining the Polarization State

Let $\tilde{E}$ be the stokes matrix for the transmitted beam. Let $\tilde{L}$ be the stokes matrix for the reflected beam. Let $\tilde{M}$ be the Bidirectional Reflectance Distribution Function (BRDF). It is a matrix here[6] and is the sum of the Mueller matrices. Then we can write:

(1) $\tilde{L} = \tilde{M} \bullet \tilde{E}$ In symbolic form we can solve for the BRDF:

(2) $\tilde{M} = \frac{\tilde{L}}{\tilde{E}} = \tilde{L} \bullet \tilde{E}^{-1}$ where the $\bullet$ stands for matrix multiplication. Here we list some Mueller matrices.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \tilde{I}$$ Region devoid of matter; or ideal isotropic nonabsorbing plate, transmittance=1

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U \quad \text{Ideal depolarizer}$$

Ideal linear Depolarizer, Azimuth of transmission axis:

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \frac{1}{2} = \tilde{H} \quad \text{Horizontal, } 0^\circ$$

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \frac{1}{2} = V \quad \text{Vertical, } 90^\circ$$

Ideal Homogenous Circular Polarizer

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \frac{1}{2} = \tilde{R} \quad \text{Right Circular}$$

$$\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \frac{1}{2} = \tilde{L} \quad \text{Left Circular}$$

Ideal Homogenous Linear Retarder
retardance = 90° ($\lambda/4$ wave plate)
Azimuth of Fast Axis

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = \left[\frac{\lambda}{4}\right]_{+45^\circ} = \tilde{P} \quad +45^\circ$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} = \left[\frac{\lambda}{4}\right]_{-45^\circ} = \tilde{Q} \quad -45^\circ$$

Retardance = 180° ($\lambda/2$ wave plate)

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} = \left[\frac{\lambda}{2}\right]_{90^\circ} = \left[\frac{\lambda}{2}\right]_{0^\circ}$$

There are other such Mueller matrices. For our measurements though we need only the top eight matrices. So we will not use the $\lambda/2$ plate set for example. If we look carefully at these matrices and notice the symmetries they have in common we note that we can write the sum of the matrices in this compact form:

$$(3)\quad \begin{bmatrix} f_1 & f_2 & 0 & f_4 \\ f_2 & f_3 & 0 & f_7 \\ 0 & 0 & f_5 & 0 \\ f_4 & f_6 & 0 & f_8 \end{bmatrix} = \tilde{M} = \frac{\tilde{L}}{\tilde{E}} = \text{the matrix BRFD .}$$

Note that we can now rewrite equation 2 in matix form as

$$(2)\quad \begin{bmatrix} L_0 \\ L_1 \\ L_2 \\ L_3 \end{bmatrix} = \begin{bmatrix} f_1 & f_2 & 0 & f_4 \\ f_2 & f_3 & 0 & f_7 \\ 0 & 0 & f_5 & 0 \\ f_4 & f_6 & 0 & f_8 \end{bmatrix} \begin{bmatrix} E_0 \\ E_1 \\ E_2 \\ E_3 \end{bmatrix}$$

Note that there 8 unknowns in equation 2 but only 4 equations . The four equations are:

$$\begin{aligned} L_0 &= f_1 E_0 + f_2 E_1 + f_4 E_3 \\ L_1 &= f_2 E_0 + f_3 E_1 + f_7 E_3 \\ (2)\text{A} \quad L_2 &= f_5 E_2 \\ L_3 &= f_4 E_0 + f_6 E_1 + f_8 E_3 \end{aligned}$$

Thus we need to make two completely different sets of measurements with two completely different sets of stokes parameters of the same object(at each point X,Y). We can do this by just changing our transmitter polarization from vertical to horizontal and adjusting the degree of elliptical polarization (to assure that the second set of equations are linearly independent of the first set) thus creating a new linearly independent set of E_i' and associated $L_i's$.

$$L_0' = f_1 E_0' + f_2 E_1' + f_4 E_3'$$
$$L_1' = f_2 E_0' + f_3 E_1' + f_7 E_3'$$
(2)B $$L_2' = f_5 E_2'$$
$$L_3' = f_4 E_0' + f_6 E_1' + f_8 E_3'$$

Equations 2 and 2A are unsuitable for use in solving for the f_i s using gaussian elimination algorithms so we convert them to standard form:

$$\begin{pmatrix} L_0 \\ L_1 \\ L_2 \\ L_3 \\ L_0' \\ L_1' \\ L_2' \\ L_3' \end{pmatrix} = \begin{pmatrix} E_0 & E_1 & 0 & E_3 & 0 & 0 & 0 & 0 \\ 0 & E_0 & E_1 & 0 & 0 & 0 & E_3 & 0 \\ 0 & 0 & 0 & 0 & E_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & E_0 & 0 & E_1 & 0 & E_3 0 \\ E_0' & E_1' & 0 & E_3' & 0 & 0 & 0 & 0 \\ 0 & E_0' & E_1' & 0 & 0 & 0 & E_3' & 0 \\ 0 & 0 & 0 & 0 & E_2' & 0 & 0 & 0 \\ 0 & 0 & 0 & E_0' & 0 & E_1' & 0 & E_3' \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \end{pmatrix}$$

Equations (2)A and (2)B constitute a set of 8 linearly independent equations and 8 unknowns.

Let us write out $\overline{M}$ and equation (3) in terms of the the Mueller matrices we defined on the previous pages:

(4) $$\tilde{M} = a_I \tilde{I} + a_U \tilde{U} + a_H \tilde{H} + a_V \tilde{V} + a_R \tilde{R} + a_L \tilde{L} + a_+ \tilde{P} + a_- \tilde{Q} = \frac{\tilde{L}}{\tilde{E}}$$

Once we have $\tilde{M}$ then we set up the 8 simultaneous equations in equation 4 and solve for the a_i. s.

$f_1 = a_I + a_U + (1/2)a_H + (1/2)a_V + (1/2)a_R + (1/2)a_L + a_P + a_Q$

$f_2 = (1/2)a_H - (1/2)a_V$

$f_3 = a_I + (1/2)a_H + (1/2)a_V$

$f_4 = (1/2)a_R - (1/2)a_L$

$f_5 = a_I + a_P + a_Q$

$f_6 = a_P - a_Q$

$f_7 = -a_P + a_Q$

$f_8 = a_I + (1/2)a_R + (1/2)a_L$

If we put these equations in matrix form for coding purposes the result is:

$$\begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \end{pmatrix} = \begin{pmatrix} 1 & 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 & 1 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 \\ 1 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{pmatrix} \begin{pmatrix} a_I \\ a_U \\ a_H \\ a_V \\ a_R \\ a_L \\ a_+ \\ a_- \end{pmatrix}$$

These a_i s become functions of X and Y. We can then make an image of just these values of a_i (as a function of X and Y) given by these solutions. ***Note that we will have 8 images of the same object.*** That means we will have obtained 8 times the information that we obtained from just intensity measurements alone.

Examples:

As an example, if a particular area on the image was an ideal depolarizer then the a_U constant would dominate at that X,Y region. Thus we have a map of points that are strong depolarizers.

The a_+ and a_- could be used to map macroscopic (dihedral) grooves in the object.

The a_V and a_H maps could be used to map microscopic linear grooves in the object. For example in a region of the X,Y map in which the a_V coefficients were large we can assume that the surface has microscopic vertical grooves in it.

If a_+ and a_- alternate in intensity across the object then there are alternating vertical and horizontal grooves on the object. Note that with this method we are able to, in a *systematic* manner, discover all the polarization effects that we discussed in the introduction. Simulation s have been run to see if this algorithms is actually mapping polarization and it appears to be.

---

[5] Shapiro,"Target Reflectivity...",Applied Optics,Vol.21,No.18,15 Sept 1982.

[3] Dainty J,"Laser Speckle and Related Phenomena", page 29.

[1] Born and Wolf, "Principles of Optics", pagre554.

[6] Kavaya, Michael ,"Polarization Effects..."Applied Optics,/Vol.26, No.5/1 March 1987

```
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Username: MAKER
Password:

==========================================================

                Last System Backup : 09-SEP-95

      PROGRAM POL
      INTEGER I,J,L,P,A,Y,S,M
      DOUBLE PRECISION T(8,8),C(8),Z(8),Ei(2,4),Eo(2,4),R
C     Input and echo data
      P=8
      R=1.0
      X1=2
      Y1=1
C     Function F() puts  in artificial Stokes parameters.
C     These data would come from the receiver in an experiment. In
C     the anolog-digital conversion the digital signal is coming from the
C     polarimeter stokes parameter measurements.
C     Eo(1,i) (i goes from 1 to 4) is the first set of stokes parameters at a
C     given X,Y point.  Eo(2,i) is the second set after the beam polarization
C     has been changed. This set is also taken at the same point X1,Y1.
C     Experimentally you will have 8 different sets of data for each given
C     point on the object(the Eo s). Plus you will have the eight different
C     beam stokes parameters (Ei s) for a total of 16 pieces of data.
      Eo(1,1)=1.0
      Eo(1,2)=0.9999
      Eo(1,3)=0.01
      Eo(1,4)=0.01
      Eo(2,1)=1.0
      Eo(2,2)=.99
      Eo(2,3)=.01
      Eo(2,4)=.01
      Ei(1,1)=1.0
      Ei(1,2)=.6
      Ei(1,3)=.2
      Ei(1,4)=0.02
      Ei(2,1)=1.0
      Ei(2,2)=0.4
      Ei(2,3)=0.1
      Ei(2,4)=0.1
C Input of artificial stokes parameters ends here.
  15  DO 19 I=1,8
      DO 18 J=1,8
```

```
      T(I,J)=.00001
   18 CONTINUE
   19 CONTINUE
      T(1,1)=Ei(1,1)
      T(1,2)=Ei(1,2)
      T(1,4)=Ei(1,3)
      T(2,2)=Ei(1,1)
      T(2,3)=Ei(1,2)
      T(2,7)=Ei(1,3)
      T(3,5)=Ei(1,2)
      T(4,4)=Ei(1,1)
      T(4,6)=Ei(1,1)
      T(4,8)=Ei(1,3)

      T(5,1)=Ei(2,1)
      T(5,2)=Ei(2,2)
      T(5,4)=Ei(2,3)
      T(6,2)=Ei(2,1)
      T(6,3)=Ei(2,2)
      T(6,7)=Ei(2,3)
      T(7,5)=Ei(2,2)
      T(8,4)=Ei(2,1)
      T(8,6)=Ei(2,1)
      T(8,8)=Ei(2,3)

      C(1)=Eo(1,1)
      C(2)=Eo(1,2)
      C(3)=Eo(1,3)
      C(4)=Eo(1,4)
      C(5)=Eo(2,1)
      C(6)=Eo(2,2)
      C(7)=Eo(2,3)
      C(8)=Eo(2,4)

   20 CALL GAUSS(P,T,C)
C     If no solution then add delta to elements
C     Find solution by back substitution
      CALL BSOLVE(P,T,Z,C)

      IF(R .LT. 1.2)THEN
      GOTO 27
      ENDIF
   25 PRINT * , ' '
      PRINT * , 'a(I)=',Z(1),'a(U)=',Z(2)
      PRINT * , 'a(H)=',Z(3),'a(V)=',Z(4)
      PRINT * , 'a(R)=',Z(5),'a(L)=',Z(6)
      PRINT * , 'a(+)=',Z(7),'a(-)=',Z(8)
C     Plot statement belongs here instead of print in real application
C     Plot Z(1), at all points X1,Y1 with the magnitude of Z(1) being
C     grey scales of ordinary intensity of the reflected light. Plot Z(2)
C     at all points in the X,Y plane with the gray scales representing
C     representing degree to which reflected light is Vertically
C     polarized. Finally at Z(8) you will have the eighth plot.
C     Thus you will have eight different photos of the same object
C     with each photo telling you something different about the state
C     of polarization of the light reflected from the object and therefore
C     telling you something different about the surface of the object.

      IF (R .GT. 1.5) THEN
      GOTO 700
      ENDIF
```

```
 27   DO 30 I=1,8
      C(I)=Z(I)
 30   CONTINUE
      DO 605 I=1,8
      DO 600 J=1,8
      T(I,J)=.00001
 600  CONTINUE
 605  CONTINUE
      T(1,1)=1
      T(1,2)=1
      T(1,3)=.5
      T(1,4)=.5
      T(1,5)=.5
      T(1,6)=.5
      T(1,7)=1
      T(1,8)=1
      T(2,3)=.5
      T(2,4)=-.5
      T(3,1)=1
      T(3,3)=.5
      T(3,4)=.5
      T(4,5)=.5
      T(4,6)=-.5
      T(5,1)=1
      T(5,7)=1
      T(5,8)=1
      T(6,7)=1
      T(6,9)=-1
      T(7,7)=-1
      T(7,8)=1
      T(8,1)=1
      T(8,5)=.5
      T(8,6)=.5
      R=2.0
      GOTO 20
 700  STOP
      END

      SUBROUTINE BSOLVE(G,V,X,D)
      INTEGER I,J,Q,G
      DOUBLE PRECISION V(G,G), X(G), D(G), SUM
C     Starting with Z(L-1)=C(L)/T(L,L),
C     solve for Z(L-1), Z(L-2), ....Z(1)
      DO  200 I=G, 1, -1
      SUM = D(I)
      DO 100 J = I+1, G
      SUM = SUM - V(I,J)*X(J)
 100  CONTINUE
      IF(V(I,I) .LT. ABS(.0000000005)) THEN
      V(I,I)=.00001
      ENDIF
      X(I)=SUM/V(I,I)
 200  CONTINUE
      RETURN
      END
```

The Behavior of a Free Beam with Piezoelectric Vibration Control System

Roy E. Martin III
Masters Graduate Student
Department of Mechanical Engineering

Worcester Polytechnic Institute
100 Institute Road
Worcester, MA 01609

Final Report for:
Graduate Student Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Laboratory
Hanscom Air Force Base, MA

September 1995

The Behavior of a Free Beam with Piezoelectric Vibration Control System

Roy E. Martin III
Masters Graduate Student
Department of Mechanical Engineering

Abstract

The focus of the summer research sponsored by the Air Force Office of Scientific Research is the application of piezoelectric ceramic thin sheets as actuators and sensors as part of an active vibration control (AVC) system. The aim of the AVC system is to increase the decay rate of a vibrating free beam with concentrated masses. Continuous beam theory and a model developed by Crawley-de Luis predict the frequencies and deflected shapes of the beam and the decay rate of the beam under feedback controls, respectively. The predicted feedback controls increase the decay rate of the freely vibrating beam by one to two orders of magnitude. An experiment was designed and built to test the theory, but the development of an analog control circuit was delayed due to difficulties. The future work based on this summer research, besides expanding theoretical models, is to develop a control system, perhaps digital, to confirm the predicted decay rate.

The Behavior of a Free Beam with Piezoelectric Vibration Control System

Roy E. Martin III

1. Introduction

The relatively recent interest in Active Vibration Control (AVC) has come about because of the emergence of piezoelectric ceramics and polymers as the sensors and actuators for an AVC system. Piezoelectricity is an old phenomenon that has repeatedly found new applications. There are two different effects that describe piezoelectricity. The direct piezoelectric effect entails the development of an electric charge in a material when stressed in the appropriate direction (figure 1a). The converse piezoelectric effect describes the development of deformations in a material when an electric field is applied across the material (figure 1b). The piezoelectric ceramics and polymers operate as both sensors and actuators by using the direct and converse effects respectively. Though piezoelectric polymers are generally used, more as sensors than actuators. In the AVC systems, the piezoelectric sensors generate an electrical signal which can be used as feedback control for the piezoelectric actuators. Their small size and mass allow for an increase in the number of sensor/actuator pairs and therefore an increase in their effectiveness as an active vibration control system.

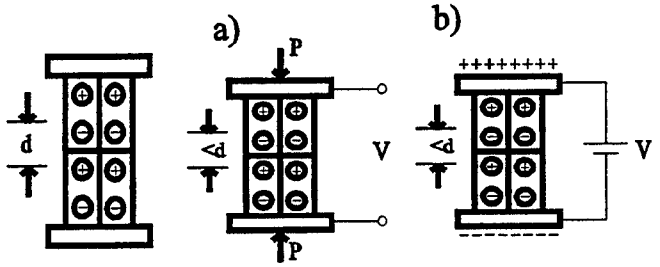


Figure 1: a) direct piezoelectric effect, b) converse piezoelectric effect.

The active vibration control system to be studied in this paper and in follow up research is for a beam with concentrated masses under micro gravity and other boundary conditions. Micro gravity will be achieved by using a drop capsule dropped either in a tower or from a balloon platform. The development of a balloon drop capsule capability at the Engineering Division of Phillips Laboratory is of high interest. However, the focus of the summer work and this paper is on the experiment to be flown in the drop capsule.

The problem to be modeled in the experiment is based on a real problem encountered by a scientist at Philips Laboratory. The scientist's experiment was on a satellite to measure the upper atmosphere by using very sensitive accelerometers. The difficulty occurred at perigee which was the most critical time for the experiment. This was also the time when the attitude control jets fired to maintain the satellite's orientation. The firing of the jets caused vibrations in the solar panels which flooded the accelerometers for a significant interval of experimental measurements. The aim for any vibration control methods would be increasing the damping in the solar panels to decrease the vibrating time and increase the data gathering time. A piezoelectric AVC is feasible for this problem because the PE ceramic sensors and actuators can meet satellite restrictions on mass and power consumption[1].

2. Theory

2.1. Basic Model

The basic beam model, based on a symmetric satellite with two solar panels, analyzed is a central mass with two equal length beams extending symmetrically from the mass. Due to initial concerns of unwanted high fundamental frequencies, end masses were added to the end of the beams to lower the frequencies of the experiment. The concern with high frequencies developed from initial design constraints and possible difficulties in gathering data at high frequencies. Therefore the various free beams to be analyzed were the continuous beam, a beam with central mass only, a beam with end masses only and a beam with central and end masses (figure 2 a,b,c and d respectively).

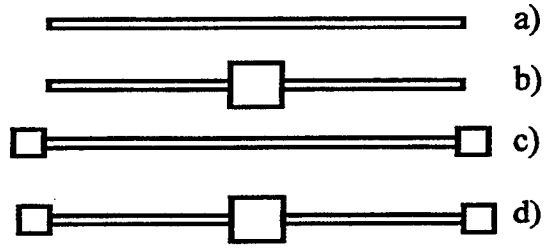


Figure 2: a) continuous beam, b) beam with central mass, c) beam with end masses, and d) beam with central and end masses

2.2. Mass-Spring Model

A mass-spring model was developed as a first approximation of the various beams. This allowed for a rough sizing of the beam lengths and masses. It also allowed for an estimate of the fundamental frequencies of the various beam configurations. The mass-spring model used consists of five concentrated masses connected by four massless springs (figure 3). The matrix equation derived from this model is equation 1. The masses in equation 1 can be correlated to the end masses (m_1 and m_5), central mass (m_3), and an equivalent beam mass representing the mass of the beam in the massless spring elements (m_2 and m_4). However, because the Crawley-de Luis model is used to model the piezoelectric forces which requires the

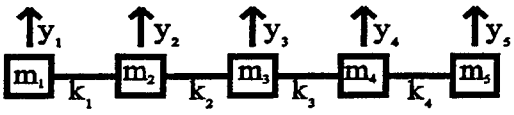


Figure 3: Mass-Spring Model

$$\begin{bmatrix} m_1 & 0 & 0 & 0 & 0 \\ 0 & m_2 & 0 & 0 & 0 \\ 0 & 0 & m_3 & 0 & 0 \\ 0 & 0 & 0 & m_4 & 0 \\ 0 & 0 & 0 & 0 & m_5 \end{bmatrix} \cdot \begin{bmatrix} \ddot{y}_1 \\ \ddot{y}_2 \\ \ddot{y}_3 \\ \ddot{y}_4 \\ \ddot{y}_5 \end{bmatrix} + \begin{bmatrix} k_1 & -k_1 & 0 & 0 & 0 \\ -k_1 & k_1+k_2 & -k_2 & 0 & 0 \\ 0 & -k_2 & k_2+k_3 & -k_3 & 0 \\ 0 & 0 & -k_3 & k_3+k_4 & -k_4 \\ 0 & 0 & 0 & -k_4 & k_4 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = 0 \qquad (1)$$

analytical solutions for the mode shapes of the beam, emphasis was placed on developing the analytical models for the beams.

2.3 Continuous Beam Model

The analytical solution for continuous beams solves the differential equation (equation 2). The initial step is to apply a separation of variables to the displacement u, namely $u(x,t) = \phi(x)\cdot q(t)$. The separation of variables yields the equations 3 and 4. $\phi(x)$ describes the deflection of any point on the beam from the non-deflected shape, referenced from the left end of the beam. q(t) describes the time dependant motion of any point on the beam. In equation 3 and 4, $\beta^4 = m\cdot\omega^2/E\cdot I$, where m is the mass per unit length of the beam, ω is the frequency, E is the Youngs Modulus, and I is the moment of inertia which is equal to $w_b\cdot t_b/12$ for beams with rectangular cross-sections. The terms w_b and t_b represent the width and thickness, respectively, of the beam. Equation 3 determines the mode shape of the beam or the shape of the deflected beam at different frequencies. Equation 4 determines the time dependent motion of each point on the beam. Applying the appropriate boundary conditions yields the solutions for different beam models.

$$EI\frac{\partial^4 u(x,t)}{\partial x^4}+m\frac{\partial^2 u(x,t)}{\partial t^2}=0 \quad (2)$$

$$\frac{d^4\phi}{dx^4}-\beta^4\cdot\phi=0 \quad (3)$$

$$\frac{d^2 q}{dt^2}+\omega^2\cdot q=0 \quad (4)$$

The boundary conditions for a free-free beam are that the shear force is zero at the beam ends and the bending moment is zero at the beam ends, as shown in equations 5 and 6 respectively. In the equations l_b is the length of the beam. Applying the boundary conditions in equations 5 and 6 to equation 3 yields the implicit frequency equation, equation 7, and the mode shape equation, equation 8. There are many values of βl_b for which equation 7 is valid. These values of βl_b correspond to the natural frequencies of a set geometry beam. Notice that βl_b can be zero, this corresponds to a natural frequency of zero, which indicates that the free beam can translate. The values of βl_b corresponding to the fundamental and first harmonic frequency are 4.730041 and 7.853205. A plot of the mode shape, equation 8, versus the non-dimensional length, x/l_b for these two values of βl_b can be seen in figure 4. Notice that zero deflection occurs on the beam at x/l_b =0.224 and 0.776 for the fundamental frequency. Suspending the beam at these two points approximates the free beam assumption allowing the tests to be conducted in a lab. Thus, for different geometries of the continuous beam the fundamental frequency and harmonics as well as the mode shapes for these frequencies can be predicted, which aids in the design process for the beam.

$$\frac{\partial}{\partial x}\left[EI\frac{\partial^2 u(x,t)}{\partial t^2}\right]\Big|_{x=0\ \&\ l_b}=0, \quad or \quad u'''(0,t)=u'''(l_b,t)=0 \quad (5)$$

$$EI\frac{\partial^2 u}{\partial x}\Big|_{x=0\ \&\ l_b}=0, \quad or \quad u''(0,t)=u''(l_b,t)=0 \quad (6)$$

$$\cos(\beta l_b)\cosh(\beta l_b)=1 \quad (7)$$

$$\phi(x)=A[\cos(\beta x)+\cosh(\beta x)]+\sin(\beta x)+\sinh(\beta x)$$

$$A=-\frac{\sin(\beta l_b)-\sinh(\beta l_b)}{\cos(\beta l_b)-\cosh(\beta l_b)}=\frac{\cos(\beta l_b)-\cosh(\beta l_b)}{\sin(\beta l_b)+\sinh(\beta l_b)} \quad (8)$$

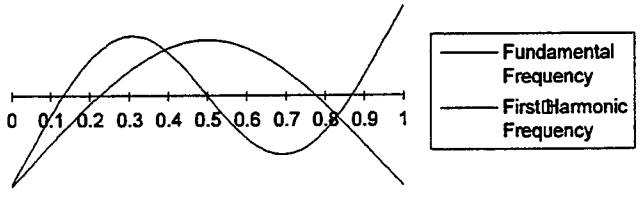


Figure 4: Shape of the deflected beam beam.

The solution for the continuous beam has been worked out in a variety of texts.[2,3] The application of end masses therefore only requires the changing of the beams boundary conditions. The two boundary conditions are that the bending moment is zero, assuming that the rotational inertial effects of the concentrated mass are negligible, at the ends of the beam and an inertial load exists at the ends of the beam, equations 6 and 9 respectively. The variable m_e is the concentrated mass at the end of the beam causing the inertial load. Applying the boundary conditions in equations 6 and 9 to equation 3 yields the frequency equation, equation 10, and the mode shape equation, equation 11. Again βl_b equal to zero is a solution of equation 10 indicating that the beam can translate. However, with the α term in equation 10, the βl_b's for the fundamental frequency and harmonics are dependent on the geometric and mass properties of the beam. That is the solutions for the βl_b's and the mode shape can not be found without including the beam dimensions and material properties. I.e., the solution can not be non-dimensionalized similar to the continuous beam problem. Generally, the addition of end masses will decrease the fundamental frequencies for similar length beams. The zero deflection points corresponding to the fundamental frequency on the beam will move closer to the ends of the beam, until for a very large end mass the zero deflection points will be at the ends of the beam. This would correspond roughly to a hinged support.

$$m_e\frac{\partial^2 u(0,t)}{\partial t^2}=-EI\frac{\partial^3 u(0,t)}{\partial x^3}$$
$$m_e\frac{\partial^2 u(l_b,t)}{\partial t^2}=EI\frac{\partial^3 u(l_b,t)}{\partial x^3} \quad (9)$$

$$\cos(\beta l_b)\cosh(\beta l_b)+2\alpha^2\sin(\beta l_b)\sinh(\beta l_b)=1$$
$$\alpha=\frac{\frac{m_e\omega^2}{EI}}{\beta^3}=\frac{m_e}{m^{\frac{3}{4}}}\left(\frac{\omega^2}{EI}\right)^{\frac{1}{4}} \quad (10)$$

$$\phi(x)=A(\cos(\beta x)+\cosh(\beta x)+2\alpha^2\sinh(\beta x))+\sin(\beta x)+\sinh(\beta x)$$
$$A=-\frac{\sin(\beta l_b)-\sinh(\beta l_b)}{\cos(\beta l_b)-\cosh(\beta l_b)-2\alpha\sinh(\beta l_b)} \quad (11)$$

When a central mass is added to the beam, a decrease in the frequency is expected from the continuous beam and the zero deflection points for the fundamental frequency are expected to move inward. Unfortunately the addition of a central mass to the beam proved difficult as it was not a change in a boundary condition of the differential equation. What was changed was the functions describing the moment of inertia and the mass per unit length of the beam. They became functions with a discontinuous pulse at the center of the beam. The exact nature of this pulse function and the need for differentiation and integration of a non-continuous function to solve the differential equation indicate that this problem would be difficult to solve with analytical methods. The nature of this problem actually calls for finite element methods, but skills in vibrational finite elements are lacking. So this potion of the theoretical development was delayed until the return to WPI where support in finite element methods for vibrations could be obtained.

2.4. Crawley-de Luis Model

The Crawley-de Luis model[4] predicts the behavior of a continuous beam with piezoelectric ceramic thin sheets symmetrically joined to the top and bottom of the beam. This model develops expressions for the applied bending force to the beam for actuators with an applied voltage. Equation 12 represents the bending force of a perfectly bonded actuator pair on the beam. The subscripts a, b and s indicate values associated the actuators, beam and sensors

$$F=-\frac{6E_bI_b}{6+\psi}[\phi''(a+\frac{l_a}{2})\phi'(a+\frac{l_a}{2})-\phi''(a-\frac{l_a}{2})\phi'(a-\frac{l_a}{2})]q-\frac{E_bt_b^2w_b}{6+\psi}(\phi'(a+\frac{l_a}{2})-\phi'(a-\frac{l_a}{2}))\frac{d_{31}V_a}{t_a} \quad (12)$$

respectively. The term ψ is the stiffness ratio of the beam with respect to the piezoelectric elements. It is defined as E_bt_b/E_at_a or E_bt_b/E_st_s since E_at_a equals E_st_s, as the same material is used for sensors and actuators. The term a is the distance from the left end of the beam to the center of the actuators. d_{31} is the piezoelectric constant which describes the relationship between the voltage and deflection in the actuators. V_a is the voltage applied to the actuators.

$$K_a=\frac{6E_bI_b}{6+\psi}[\phi''(a+\frac{l_a}{2})\phi'(a+\frac{l_a}{2})-\phi''(a-\frac{l_a}{2})\phi'(a-\frac{l_a}{2}) \quad (13)$$

$$F_V=\frac{E_bt_b^2w_b}{6+\psi}(\phi'(a+\frac{l_a}{2})-\phi'(a-\frac{l_a}{2}))\frac{d_{31}}{t_a} \quad (14)$$

Equation 12 can be separated into two parts, a stiffness component K_a associated with the displacement q, equation 13, and a forcing term F_V associated with the voltage V_a, equation 14. Thus, the single mode equation of motion can be expressed as equation 15. When the voltage is proportional to the velocity $\dot{q}$, which can be done through a control circuit, equation 15 becomes equation 16, where C_α is a constant of proportionality. Solving the differential equation 16 for q yields equation 17. The first term in equation 17, in square brackets, represents the harmonic response of q. The other term represents the decaying exponential response of q. Substituting the expressions for B and M into the second portion of equation 17 yields equation 18. The effective damping coefficient for the beam is represented by equation 19. The effective damping of the beam will be increased for positive values of F_V and C_α.

$$M\ddot{q}(t)+B\dot{q}(t)+(K+K_a)q=-F_VV$$
$$M \text{ - modal mass } =\frac{K+K_a}{\omega_n^2} \quad (15)$$
$$B \text{ - modal damping } =2\zeta_b\omega_nM$$
$$K \text{ - modal stiffness } =E_bI_b\int_0^{l_b}(\phi''(x))^2dx$$

$$M\ddot{q}+(B+F_VC_\alpha)\dot{q}+(K+K_a)q=0 \quad (16)$$

$$q(t)=\left[A_1e^{i\sqrt{\frac{K+K_a}{M}-\frac{(B+F_VC_\alpha)^2}{4M^2}}t}+A_2e^{-i\sqrt{\frac{K+K_a}{M}-\frac{(B+F_VC_\alpha)^2}{4M^2}}t}\right]e^{-\frac{B+F_VC_\alpha}{2M}t} \quad (17)$$

$$q_{decay}(t)=e^{-(\zeta_b+\frac{F_VC_\alpha\omega_n}{2(K+K_a)})\omega_nt} \quad (18)$$

$$\zeta_{eff}=\zeta_b-\frac{F_vC_\alpha\omega_n}{2(K+K_a)} \quad (19)$$

3. Design and Manufacture

3.1. Beam and Weight Design

The design of the beam and weights has to meet a set of criteria. The first criteria is a restriction on the over all length so that the beam fits inside a drop capsule for either a balloon or tower micro gravity drop. The overall length of the beam was initially restricted to one foot. The second criteria is that the beam is a thin beam so that the piezoelectric actuators can effect the motion of the beam. The expression thin beam means that the length of the beam is much greater

than its width and its width is much greater than its thickness. A rule of thumb[4] for the thickness of the beam is ten times the thickness of the piezoelectric actuators. Piezoelectric thin sheets are approximately 10 thousands of an inch thick, so the beam's thickness should be roughly a tenth of an inch or, to simplify dimensions, an eight of an inch. The final criteria placed on the beam is that it has a low fundamental frequency. This restriction is primarily for data gathering purposes and the top frequency limit is 100 Hertz. A low frequency means less data points are needed per sample which requires less expensive instrumentation for the experiment. With these criteria and the models from the previous section, a design developed.

The theoretical models are used to vary and compare the effects of various parameters. The parameters in order of importance are frequency, length, concentrated masses, and width. The thickness of the beam is set primarily by the thickness of the actuators. Using the continuous beam model, the relationship between the frequency and length is, for the fundamental frequency of the continuous beam, $\beta l=4.730041$ with $\beta^4=m\omega^2/EI$ and $m = 0.098$ lb/in^3, $E = 10^7$ psi, $t = 0.125$ in, and $w = 1.75$ in. This relationship is plotted in figure 5. From figure 5, for the 100 Hz restriction on frequency the 12 inch limit on length is too strict. Therefore the length of the beam was extended to 18 inch, which meets the frequency limit with a fundamental frequency of 78.78 Hz. Use of the end mass model provides understanding as to how the end mass will effect the frequency. Using the set dimensions of the beam (l=18 in, w=1.75 in, and t=0.125 in), the end mass is varied to determine the relationship between m_e and ω. From equation 10, a range of masses exist in which small changes in the mass does not effect the frequency. Therefore if the mass is chosen from this area, small changes for ease of manufacture can be made without greatly effecting the vibration problem. The end mass weight used in this experiment is 0.593 lb. Since a model for the central mass was not found, it is only expected that the addition of the mass will further reduce the frequency. For this experiment the central mass was set at roughly twice an end mass or 1.102 lb. Finally, width does not effect the frequency. The width of the beam is dependant on the size of the piezoelectric actuators and sensor. The final design and dimensions of the beam and weights are in the drawing in figure 6.

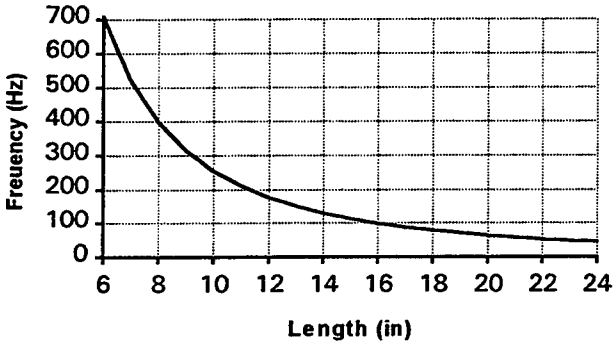


Figure 5: Length versus frequency

3.2. Piezoelectric Actuator and Sensor Design.

The piezoelectric actuator and sensor design for the 12 week time period of the on-site summer research is largely dependent on the availability of materials. The desired parameters are low cost for a relatively small amount of material and attractive piezoelectric properties, i.e. a high piezoelectric coefficient d_{31} and a Youngs Modulus close to that of aluminum. The material that meets these criteria is Piezo System, Inc.'s PSI-5A-S4-ENH a piezoelectric single sheet. This material has a Young's Modulus of 6.6×10^{10} N/m^2, piezoelectric strain coefficient $d_{31} = -190\times10^{-12}$ m/V,

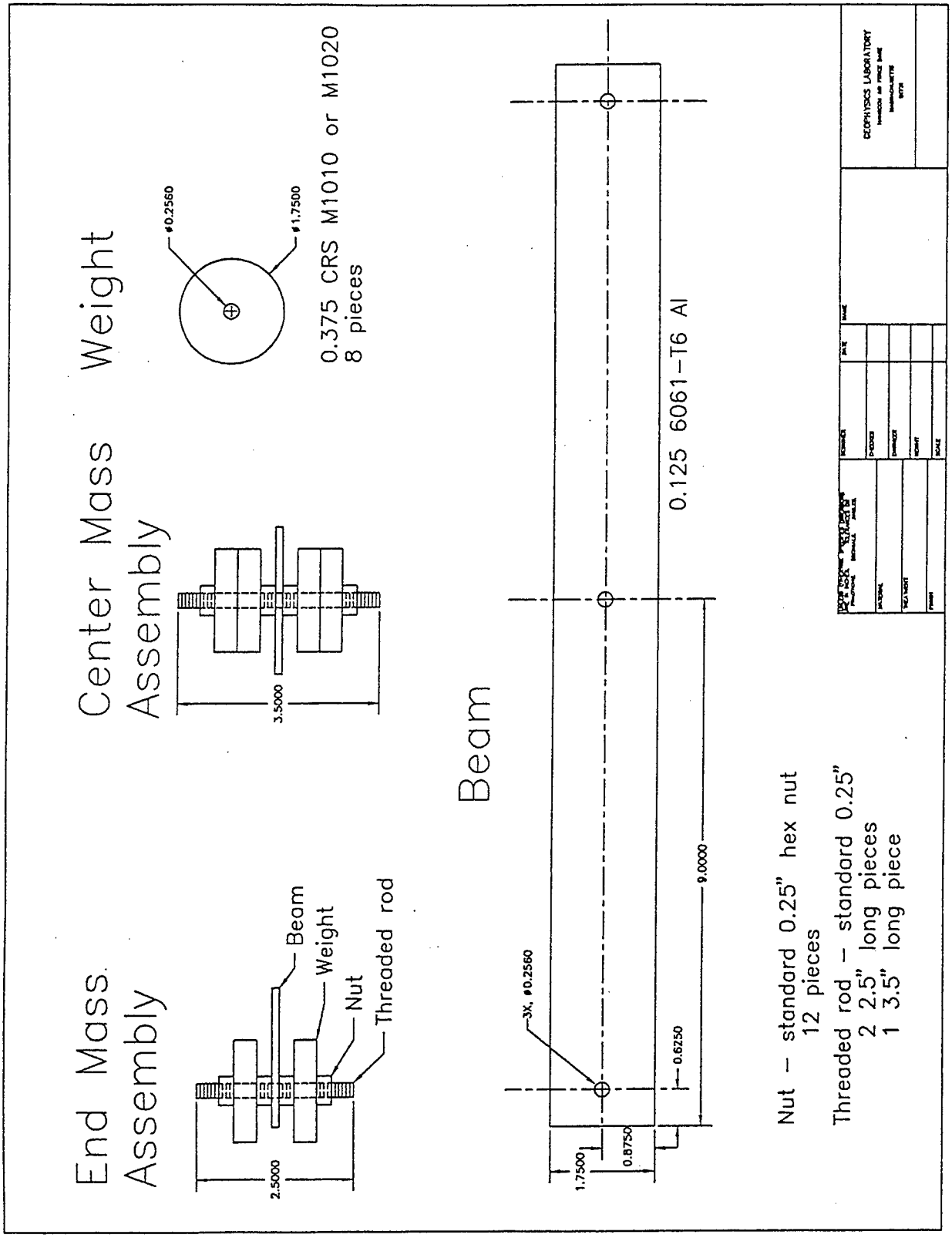


Figure 6: Beam and Weight Design

piezoelectric voltage coefficient $g_{31} = -11.6 \times 10^{-3}$ N/m^2 The dimensions for the single sheet are 2.85 inch by 2.85 inch by 0.0075 inch ± 0.0005 inch from which 6 actuator 1.425 inch by 0.950 inch, 16 sensors 1.425 inch by 0.356 inch, or 4 actuators and 4 to 6 sensors can be produced.

The placement of the actuators and sensors is important for the stability of the feedback controls, i.e. the response of the beam is decreased, and the maximum effectiveness of the control system. The actuators and sensors should be collocated on the beam, in this design they are side by side, as this leads to inherent stability[5]. They should also be at or near the maximum strain in the beam. The maximum strain in the continuous beam can be found by taking the second derivative of the mode shape, equation 8. The resulting curve, shown in figure 7, has its maximum point at the center of the beam. As the central mass is located there, the actuator/sensors are located symmetrically as close as possible to the center.

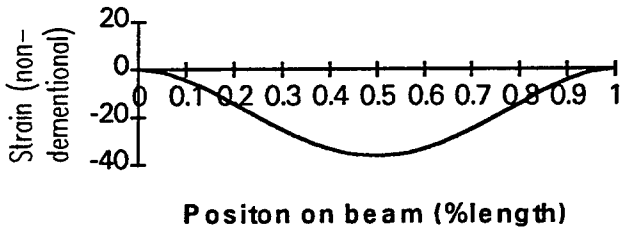


Figure 7: Strain curve for continuous beam at fundamental frequency

3.3 Manufacture Methods

The beam and weights were made to specifications at Wentworth Institute of Technology. The actuators and sensors are made and placed on the beam by following the procedures set forth by Lieutenant Hixenbaugh in his thesis[6]. The main difference is in making electrical contact to the piezoelectric. Hixenbaugh solders copper leads to the thin sheets, but there is some concern as to whether or not the temperature of soldering adversely affects the properties of the ceramic. Therefore conductive epoxy is used to make the electrical contacts. The adhesive bond between the actuator or sensor and the beam consists of a small dot of conductive epoxy surrounded by a rectangle of a strong non-conductive epoxy, see figure 8. This layout effectively makes the beam the electrical ground with a separate wire epoxied to the beam connecting it to the control instrument and circuits. The leads from the free surfaces of the piezoelectric are attached using the conductive epoxy with care taken to keep the epoxy from the edges of the actuators and sensor to avoid shorting the circuit. All the leads and the ground wire are then routed off the beam at the zero deflection points to minimize their affect on the motion of the beam.

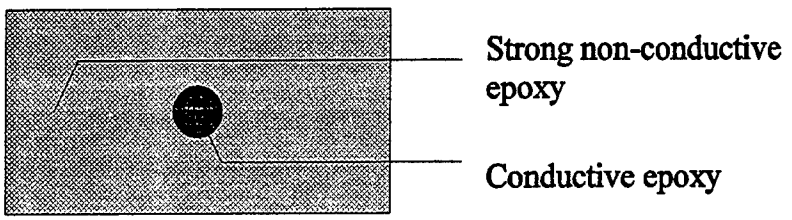


Figure 8: Adhesive layer between beam and actuators/sensors

3.4. Electrical Control Circuit

The electrical control circuit needs to be a PD, proportional derivative, controller so that the actuator forcing terms affect the damping of the beam. The reason is that the sensors produce a voltage signal that is proportional to the strain in the sensor as described in equation 20. Assuming a perfect bond between the piezoelectric and the beam, the strain in the sensor, ϵ_s, is related to the strain on the surface of the beam as in equation 21[4]. The strain on the surface of the beam can be expressed as in equation 22, where the s+ and s- indicate the right and left ends of the sensors. The voltage developed by the sensors can now be expressed as equation 23, where a indicates the midpoint of a sensor pair as located from the left end of the beam. Sending the voltage developed by the sensors through the PD control should generate the voltage needed by the actuators to effect the damping of the beam. A simple expression of this can be seen in equation 24, where G is the gain provided by the proportional part of the controller. The expression, equation 19, for the effective damping when applying equation 24 will become equation 25, where the only difference between 19 and 25 is the addition of the gain G. This gain should be between zero and one based on previous stability analysis.[5]

$$V_s = t_s E_s g_{31} \epsilon_s \quad (20)$$

$$\epsilon_s = \frac{\psi}{\psi+6}\left[\frac{\epsilon_b^{s+}+\epsilon_b^{s-}}{2} + \frac{\epsilon_b^{s+}-\epsilon_b^{s-}}{2}\cdot\frac{x}{l_b}\right] \quad (21)$$

$$\epsilon_b^{\,s} = -\frac{t_b}{2}\frac{d^2\phi(x)}{dx^2}q(t) \quad (22)$$

$$V_s = \frac{-t_s t_b E_s g_{31}}{4}\frac{\psi}{\psi+6}\left[\phi''(a+\tfrac{l_s}{2})+\phi''(a-\tfrac{l_s}{2})+(\phi''(a+\tfrac{l_s}{2})-\phi''(a-\tfrac{l_s}{2}))\frac{a}{l_b}\right]q(t) \quad (23)$$

$$or$$

$$V_s = C_a q(t)$$

$$V_a = G\frac{dV_s}{dt} = GC_a\frac{dq(t)}{dt} \quad (24)$$

$$\zeta_{eff} = \zeta_b - \frac{F_V G C_a \omega_n}{2(K+K_a)} \quad (25)$$

The design and manufacture of an analog control circuit was delayed because of unfamiliarity with circuit design and that a unknown phase shift in the sensor circuit could not be determined theoretically or during brief experimentation. This delay extended into the last weeks of the summer until it was unfeasible to design, build and test a circuit on site.

4. Testing and Results

4.1. Experimental Setup

Due to the difficulty with the analog control circuit, the experiments were used to verify the model prediction for the fundamental frequencies and to determine the nature damping coefficient, ζ_b, of the aluminum beam though the use of the logarithmic decrement. To accomplish these tests the beam was suspended at the zero deflection points with two pieces of fishing line two feet long. All four actuators were wired to a signal generator which could output a stable sinusoidal signal over a large range of frequencies. One pair of sensors was wired to an oscilloscope, a Data Precision Data 6000 Digital Oscilloscope, from which the frequency, amplitude, and phase shift of the drive signal and sensor signal could be read.

The second step in setting up was to calibrate the sensors so that the voltage read from the sensors could be correlated to the deflection of the beam. Initially, the calibration was to be done by statically deflecting the beam with a know deflection and reading the voltage generated by the sensors. This is not possible because of the nature of piezoelectrics. The electric model of a piezoelectric thin sheets, figure 9, is a perfect voltage source (proportional to the strain) connected in parallel to a capacitor. The capacitance for the actuators and sensors used was 51 nanofarads and 19.5 nanofarads, respectively. Such small capacitors will discharge any charge built up by the direct piezoelectric effect through the instrument circuit, the piezoelectric ceramic, and even the air. Therefore, the sensors needed to be dynamically calibrated. However as the need for this type of testing was not planned and arranged for earlier, dynamic calibration would not be possible during the remaining weeks at Phillips Labs. Without the sensors calibrated, they could still be used to determine the resonance frequencies and any phase shift between the beams motion and the signal output by the sensors. The natural damping of the beam can still be determined because the logarithmic decrement uses the ratio of deflection amplitudes at the same point in different cycles. As the sensor voltage is proportional to the deflection by a constant amount, the ratio of the voltage amplitudes is the same as that of the deflection amplitudes. So the planned experiments for frequency determination, natural damping, and phase lag could still be conducted without calibration.

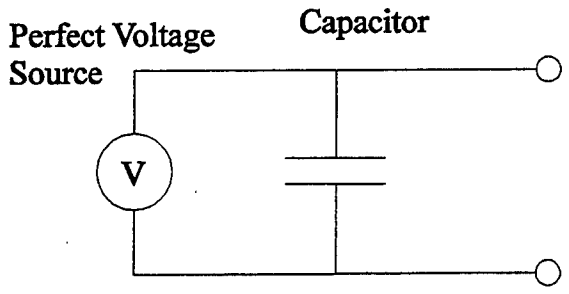


Figure 9: Electric Model of Piezoelectric thin sheet

4.2. Determining the Fundamental Frequencies

The fundamental resonance frequencies of the beam with different symmetric mass configurations were found by using the piezoelectric actuators to create vibration in the beam. The signal generator attached to the actuators could be finely adjusted to find the driving frequency at which the peak amplitude response of the sensors occurred. This driving frequency is close to the natural frequency of the beam as the beam has only a small amount of damping exists within the beam. The results for the different configuration are shown in table 1.

Table 1. Results for Fundamental Frequencies for Different Beam Configurations

	Driving Frequency (oscilloscope)	Peak Amplitudes (oscilloscope)		Natural Frequency (predicted)
		Input Signal (actuator)	Output Signal (sensor)	
beam only	79.5 Hz	± 4.66 V	± 7.69 V	78.78 Hz
beam with central mass only	57.5 Hz	± 4.63 V	± 8.68 V	-
beam with end masses only	41.6 Hz	± 4.65 V	± 4.86 V	39.69 Hz
beam with end and central masses	22.5 Hz	± 4.96 V	± 8.49 V	-

The model predictions for the beam alone and the beam with end masses were within 2% and 5% of the experimental values, respectively. The predicted values varied from the natural frequency because they did not consider the effects the actuators and sensors would have on the frequencies. The greater error for the end mass model is because of the idealization of a concentrated mass at the end of the beam. The end mass model was adjust to more appropriately reflect the actual beam by shortening the length of the beam in the model. The length in the model went from the center of the left end mass assembly to the center of right end mass assembly, or l_b = 16.75 in. The addition of the central mass reduced the frequency as expected. An unexpected result was that the addition of the central mass caused a significancy larger response in the beam and the end masses reduced the response amplitudes. This occurred because the addition of the masses changed the mode shape of the beam and this changed the strain in the beam at the sensor locations.

4.3. Determining the Damping Coefficient

To determine the damping coefficient the actuators were used to excite the beam. Then the voltage signal was be disconnected from the actuators and the signal decay output by the sensors was recorded in the digital oscilloscope. By use of the logarithmic decrement, equation 26, the damping coefficient ζ can be calculated from the captured signal. Four separate vibrational decays were recorded with 3 samples of 11 data points separated by 1 cycle taken from each decay. The decay rate calculated from this data was ζ=0.002±0.0002. The natural frequency was recorded as 79.8 Hz, so that the ratio of driving frequency to natural frequency is 79.5/79.8 or 0.996.

$$\zeta = \frac{\delta}{2\pi} \quad \textit{for small } \delta$$

$$\delta = \frac{1}{j}\ln\frac{x_1}{x_{j+1}} \qquad (26)$$

j = number of complete cycles between data points

x_1 = *first data point*

x_{j+1} = *second data point*

With the beam's damping coefficient, the effect of the control system can be predicted with equation 25. The term in equation 25 due to the actuators has to be duplicated to account for two sets of actuator/sensor pairs. The values for F_{V1} and F_{V2} are 0.008 charge/meter, C_{a1} and C_{a2} are -1.645x10^4 and -1.417x10^4 kg m/s^2 charge, and $K+K_a$=4.7845x10^4 kg/s^2. The gain G is set at 0.5. The resulting effective damping coefficient is 0.6367, an increase in two orders of magnitude over the natural damping coefficient of the beam.

4.4. Determine the Phase Shift

The determination of the phase shift requires a known harmonic excitation source. The suspended beam can not be used because a phase shift could exist between the signal generator and the response of beam to the actuation. A proposed solution is to use a shaker table with accelerometers attached to the table. The beam is fixed to the table through its center hole. The beam is modeled as two symmetric cantilever beams. In this way a comparison can be made between the accelerometer readings and the piezoelectric sensors.

The electrical model for the sensor-oscilloscope circuit is an resistor-capacitor (RC) circuit, figure 10. The phase shift in an RC circuit is found by plotting the impedance of the capacitor against the impedance of the resistor, as shown in figure 11. When the capacitor's impedance is large in comparison to the resistance, the phase shift for the circuit is 90°. When the resistance is large in comparison to the impedance of the capacitor, the phase shift is zero. The impedance of a capacitor, in this case the sensor, is X_c = $1/(2\pi\omega C)$ or $1/(2\pi(40\ Hz)(19.5\ nanofarads))$, which equals 2.04×10^5. The resistance in the oscilloscope is 1 MΩ. The phase shift according to figure 11 should be approximately 11.5°. If the resistance in the oscilloscope is increased to 10MΩ, the phase shift reduces to 1.2°.

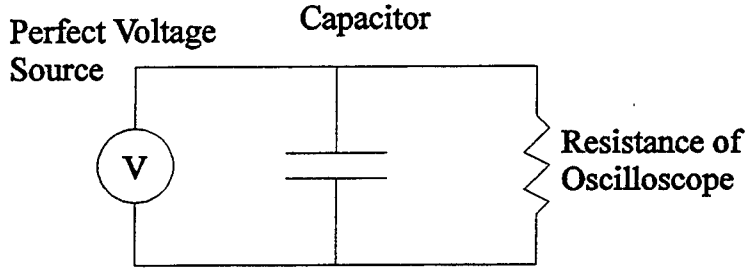


Figure 10: Electrical model of sensor-oscilloscope circuit

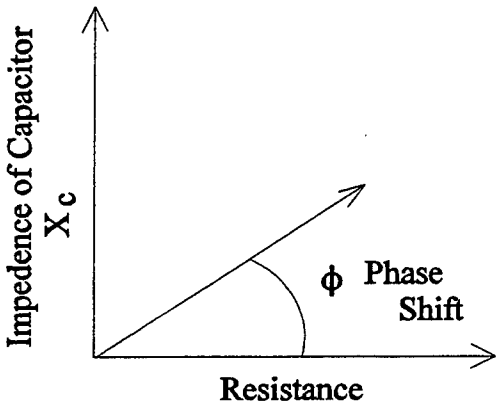


Figure 11: Graphical method for determining phase shift

The initial testing of the beam was done with the resistance of the oscilloscope set at 1 MΩ. The resulting sensor and accelerometer signals were roughly 180° out of phase, because the sensors are proportional to the position of the beam, the accelerometers are proportional to the acceleration of the beam, and the response is harmonic. If the one of the two signals is inverted and compared with the other, the phase shift was found to be approximately 10° which compares well with the above analysis of the circuit. However when the resistance was increased to 10 MΩ, the phase shift did not decrease as predicated but increased. The reason for this result was not determined and as this occurred during the last week on base, more detail analysis will have to wait.

5. Future Work

This work is proposed to be continued at Worcester Polytechnic Institute during the academic year starting in 1995, with the possibility of being incorporated into a masters thesis.

5.1. Theory

1. A model for the beam with a central concentrated mass needs to be developed. If an analytical model can not be found, the development of a finite element model is required. A concern with a finite element model is whether or not it will generate a mode shape function that can be used with the Crawley-de Luis model. If not, then the Crawley-de Luis model has to be adjusted or changed to achieve similar results.

2. A theoretical model for the controls needs to be developed and analyzed for stability.

5.2. Design and Manufacture

1. The control circuit needs to be built for the lab suspension tests and for the eventual micro gravity tests either in a drop tower or from a balloon.

2. A micro gravity drop experiment package for a drop tower or balloon needs to be designed and built for this experiment.

5.3. Testing

1. More precise measurements of the phase shift in the sensors are needed along with an analysis of what is occurring.

2. With understanding of the phase shift, the control circuit can be built and tests made to compare the achieved damping with the predicted damping.

3. The micro gravity drop experiment package has to be fully tested before any flight(s) can be made.

Acknowledgments

I would like to acknowledge the Air Force Office of Scientific Research and the Research and Development Laboratories for the opportunity of this summer research program. I would like to thank George Jumper, who made my time at Phillips Laboratory possible, enjoyable, and valuable. I would also like to thank Will Thorn, George Mcphetres, Dave Curtis, John Otis, and Jim Anderson for their aid and guidance during the twelve weeks of the summer research program at Phillips Laboratory.

References

1. Inman, Daniel J., "Modeling and Control of Intelligent Flexible Structures," Air Force Office of Scientific Research, Bolling Air Force Base, DC, Tech. Rept. AFOSR-94 0353, March 1994, pp. 2-4.

2. Tse, Francis S., Ivan E. Morse, and Rolland T. Hinkle, *Mechanical Vibrations: Theory and Applications*, Allyn and Bacon, Boston, 1978, pp. 269-276.

3. Meirovitch, Leonard, *Elements of Vibration Analysis*, McGraw Hill, New York, 1986, pp. 220-227.

4. Crawley, Edward F., and Javier de Luis, "Use of Piezoelectric Actuators as Elements of Intelligent Structures," *AIAA Journal*, Vol. 25, No. 10, October, 1987, pp. 1378.

5. Fanson, J. L., and T. K. Caughey, "Positive Position Feedback Control for Large Space Structures," *AIAA Journal*, Vol. 28, No. 4, April, 1990, pp 717.

6. Hixenbaugh, Franklin D., *A Study on Piezoelectric Actuators and Sensors for Vibration Control of Flexible Space Structures*, Master's Thesis, Naval Postgraduate School, Monterey, CA, September, 1993, pp. 3-14.

INPUT FILE TOOL FOR MACH(MAGNETOHYDRODYNAMICS SIMULATION PROGRAM)

Michael McGuinness
Department of Computer Science

State University of New York at Albany
1200 Wasgington Avenue
Albany, NY 12222

Final Report for
Graduate Student Research program

Sponsored by:
Air Force Office of Scientific Research
Bolling AFB, DC

and

Phillips Laboratory

August 1995

1 Problem Overview

The problem is to create a properly formatted input file for the MACH simulation program. MACH is a magnetohydrodynamics simulation program for problems with complex shapes. Any particular problem for MACH is set up using an input file (input deck). The problem is to let a user interactively create, or modify an existing, MACH input file through a graphical interface.

A MACH input file is comprised of lines of plain text. On each line there is either a comment, a namelist name, a namelist end symbol, or a MACH variable assignment statement. All MACH variables are logically grouped into fields called namelists.

There are eleven namelists. Some of the namelists have special meaning for the control of the simulation to be run. Others simply specify the values for that problem's particular parameters.

The most common use of the software will be modifying an existing MACH input file. For example, a user has an existing MACH input file. The user has successfully run MACH with this file several times. Now, the user wants to make some changes to their problem set-up. The user then makes changes to their MACH input file, using this software, then runs MACH with the new file. This software is designed to make it easier for users of MACH to make changes to their input files.

2 Software Requirements for the MACH Input File Tool

- To let a user interactively create a properly formatted MACH input file.
- The program will run in an X-Windows Unix environment.
- There are up to 548 MACH variables that the user can choose from. The program must let the user choose from among the variables in a logically ordered fashion. There are some variables that must be be specified in a MACH input file and some that can be left out. Other variables are 1 or 2 dimensional arrays. All these types of variables should be easily and logically grouped so a user can specify (or modify) them.
- MACH specifies problem geometries in a 2 dimensional fashion using blocks. These blocks each have four corners. The corners are specified as points which are x-y coordinate pairs of some user-defined scale. The program must let the user create (or modify) a specific problem geometry using an interactive drawing tool. The drawing tool will be constructed as follows:

  1. The program will let the user add and remove points, corners, blocks, and arcs. Arcs are circular sections defined by the user in one of two ways: By two endpoints and a center point or by two endpoints and an angular direction.
  2. The program will let the user specify the values of each point, corner, and block either by typing in values or by using a pointing device. Arcs may be added by typing in values for the endpoints and a value for the midpoint or angular direction.
  3. The program will let the user change modes while in the ezgeom drawing screen. There will be three drawing modes:
     (a) Add Points Mode.
     (b) Add Block Mode.

(c) Move Points Mode.

4. The current drawing mode will be displayed in the drawing screen.
5. The program will let the user move points around the drawing screen. The program will let the user simultaneously view the changes to the values of the points as they are being moved.
6. The program will let the user choose their own scale.
7. The program will display the current units (meters or inches) in the points window.
8. The program will let the user scroll through the enter-point-data window. The number of entry boxes for the point arrays is determined by the user in the define indices window.
9. The program will let the user redefine any of the block, point, and arc array indices at any time during the $ezgeom session.

- The program will be designed to be maintainable. As the MACH program changes, the input file specifications may also change. The program must be written such that any changes can be made by a programmer. The maintenance programmer will be familiar with C programming and Tcl/Tk GUI programming.
- The program will force the user to enter a filename for the file to be generated by the program.
- The program will check the problem name entered by the user. The problem name must be less then 17 characters in length and have no spaces.
- The program will check the length of the problem description entered by the user. The problem description must be less than 81 characters in length.
- The program will help create a syntax error-free MACH input file.
- The program will assume that any user is familiar with the MACH program. The user is expected to know what type of values can be entered for most of the MACH variables. In many cases the program will force the user to enter data if that data is required by MACH. This will help the user create (or modify) a syntactically correct MACH input file file.
- As the user is entering data in the main loop, the program will let the user go back, review, and edit what they have done so far.
- The current modtim will be displayed in the main menu screen.
- At anytime, the user should be able to view the MACH input file that they are creating. All the variables that have been entered up to that point should be included in that file.
- The program will create windows where the user will enter data for their MACH input file. The windows will be grouped according to namelist and sub-grouped by MACH variable characteristics (i.e. variables with similar functionality are grouped together within a namelist.).
- The program will not let the user violate any of the restrictions listed in section 3.3.

- The program will distinguish between the two variables, both named t, in the $contrl namelist and the $restrt namelist.
- All MACH true-false variables will be represented by radiobutton choices in the GUI.

3 Outline of Mach Input File and Namelists

Problem generation for a MACH simulation is controlled by information read from a local file. An input file contains, first, a one-line descriptive label of fewer than 80 characters that begins in column one, a one-line problem name of fewer than 17 characters that begins in column 1, followed by some of the FORTRAN namelists: $contrl, $inmesh, $curnt, $ezgeom, $ezphys, $output, $modtim, $laser, $matmdl, $remap, and $restrt. All but the first two lines of this input file and comment lines must have its first column blank. Below, a general description of the function of the eleven namelists is given.

1. $contrl contains only variables that control the duration of the simulation and the activity of the various physics routines of MACH . There are no block or boundary-dependent variables in the namelist $contrl.
2. $inmesh contains a few scalar quantities that describe the behavior of the mesh generator, as well as, the various arrays that describe block-dependent initial data, and the block- and boundary-dependent physical boundary conditions to be used for a particular simulation.
3. $curnt contains the variables that describe applied currents as well as the elements of external electric circuits if such a circuit is required for the simulation of interest.
4. $ezgeom contains the parameters that describe the shape of the computational mesh. The mesh can also be specified by a number of input parameters in namelist $inmesh, but $ezgeom is much easier to use.
5. $ezphys contains parameters that are used to set initial physical data throughout the computational region. Using $ezphys allows simpler input files for many problems.
6. $output contains all parameters that control MACH output.
7. $modtim serves a special purpose. Each occurrence of this namelist in an input file enables the user to re-specify most input parameters at the predetermined time, *tmod* or $t +$ *tmodinc*, or cycle, *ncycmod*, or *ncyc* + *ncmodinc*.
8. $laser
9. $matmdl Materials model.
10. $remap contains variables that let the user of MACH alter the problem geometry after some amount of execution time. $remap must come after the $modtim namelist.
11. $restrt should occur only in input files that are designed to read restart files, and should be the first namelist in the file.

3.1 More on $modtim

The $modtim namelist is special. It is used to create a new environment in a MACH simulation after some user-specified time interval. If the user of MACH wishes to use $modtim, they simply add the $modtim namelist and the $modtim variable assignments to their MACH input file. Once a modtime is read by MACH, all the variables in all the other namelists (except $restart and $ezphys) can be re-specified with new values. $ezgeom may be used after $modtim only if $remap is used.

The construction of the namelists and the special use of $modtim lead to a logical data-structure implementation for the MACH namelists and variables. See Section 5 for further discussion on the data structures.

3.2 More on $ezgeom and $remap

The $ezgeom and $remap namelists are also special. The MACH problem geometry is set up in $ezgeom. This geometry can be altered, but not like the other namelists; it must be changed in conjunction with the $remap and $modtim namelists.

Once the geometry has been specified, the user of MACH can alter the problem geometry at a certain time by adding a $modtim namelist then a $remap namelist. The remap namelist does exactly as its name suggests: it remaps the problem geometry by letting the user change the number of blocks. The number of blocks is originally specified in $ezgeom.

The geometry that the user has originally specified in $ezgeom does not go away when the user wants to use $modtim and $remap. The original geometry is just altered. This is the reason that the data the user has entered for the $ezgeom namelist can't be cleared for each loop through the main menu (see section 5). The program must maintain this $ezgeom data throughout, but if the user wants to do a remap, it must let the user alter the existing variables in $ezgeom.

If the user chooses to do a $modtim but does not want to alter the geometry with a $remap, the program will not let the user alter the geometry.

3.3 Some Restrictions

MACH will stop reading input under the following conditions (not a complete list):

- If MACH reads a non-valid namelist.
- If $restrt is read but it is not the first namelist in the file.
- Namelist $ezgeom comes somewhere after a $modtim declaration without a $remap declaration.
- Namelist $ezphys comes somewhere after a $modtim declaration

4 Data Structures

The data structures empolyed by the program are driven by the characteristics of a syntactically correct MACH input file. The program stores data in two ways:

1. In C program arrays and linked lists of structures.

2. As global variables in the Tcl interpreter.

4.1 Data Structures in C

The program stores MACH data read from a MACH input file into an array of 100 elements. This value is defined as in the header file imt.h. This value was defined at 100 by the customer as the maximum possible number of $modtim namelists that could normally occur in a MACH input file. The name of the array is therefore called the modtime array. The modtime array is the main data structure used by the program to store MACH variable data.

Each modtime array element is a structure. Each one of these structures contain head pointers to lists of another type of structure. This list structure is defined in the imt header file (imt.h) as NameElement. These list structures contain three things:

1. Line of data read from, or to be written to, a MACH input file.
2. A boolean switch. This switch is true or false depending upon whether the line contained in the list structure is a comment line.
3. Pointer to the next element in the list

The philosophy of the modtime array is very simple: it only contains lines that will be read from, or written to, the MACH input file. I have left the job of reading and writing syntactically correct lines (i.e. variable assignment statements and comment lines) to other helper functions.

4.2 Data Structures in Tcl

Tcl does not handle complex data structures such as linked lists and pointers to structures. It works well with global variables and associative arrays. All of the MACH variables will have the same name in Tcl as they do in a mach input file. The arrays in a mach input file are also stored as arrays in Tcl. All the mach variables are stored as global variables in Tcl. Tcl keeps all of these variables in, what it calls, the interpreter. Once the interpreter is destroyed, all the variables and their values will also be destroyed.

A global variable in Tcl can be created in a least two ways:

1. Any textvariable from any entry box will be global.
2. Setting a variable in the root widget will make the variable visible to any Tcl procedures in the interpreter. Any Tcl procedure that needs to use a global Tcl variable must use the "global" command.

5 High Level Design Specification (HLDS)

The high level design of the MACH input file tool is driven by the characteristics of the MACH namelists.

The problem is broken into 2 parts at the highest level:

1. Create a MACH input file from scratch.
2. Modify an existing MACH input file.

The global algorithm is fairly straight forward, however; there are many details to deal with as the namelists are being edited by the user. See figure HLDS In the Appendix.

5.1 High Level Design Preliminaries

The user is first asked for a filename, a problem description and a problem name. The problem name must be specified. MACH uses this name as a subdirectory name to be created to store output. The problem description is another special line that must occur at the beginning of a MACH input file. See Section 3 for a description of these two special lines.

If the user wants a restart file the first FORTRAN namelist to occur must be $restrt. This is the next step in the HLD. If the user is asked if they want to do a $restrt. If so, the user is shown the $restrt window, else the user goes on to the main menu.

5.2 High Level Design Main Loop

The main loop consists of a menu-bar system wherein the user enters data for eight namelists: $contrl, $ezgeom, $ezphys, $matmdl, $inmesh, $laser, and $curnt. Each namelist is made up of separate modules. See section 6 for individual detailed module descriptions.

The user pulls down menu's and enters data for whatever namelist they wish. Also displayed on the main menu screen is the current modtim. Since the user may re-specify all or some of these main menu variables for each modtime, (See section 3.1) it is necessary to inform the user for which modtime they are entering data.

When the user is finished with these eight namelists, they have 3 choices:

1. Continue. (i.e. add a $modtim and/or $remap)
2. Done. (Save and Exit)
3. Exit. (No save)

When the user has chosen Continue, this leads the user down a path in the HLD. At this point three crucial things happen in the background (transparent to the user).

1. All the data the user has entered into the main menu windows gets saved to the appropriate C data structures. This fills the next array element in the modtime array (See section 4.1 for C data structure descriptions).
2. All the variables for the eight namelists in the main menu must be cleared (except $ezgeom and $ezphys. See section 3.2). This will let the user enter new data for as many variables as they want.
3. The array counter for the modtim array and the display counter for the current modtim (in the main menu screen) are incremented.

After these background tasks have been completed, the user is prompted to enter data for the $modtim namelist, and if they choose, the $remap namelist. When this is complete, the user is back at the main menu again. The user then enters their data for the new modtim.

If the user chooses Done (Save and Exit) from the main menu, the program saves the data the user has just entered in the main menu, saves everything to the file, and exits. This is the only path in the HLD that lets the user exit and save the data to a syntactically correct input file.

6 Module Descriptions.

The module descriptions contain some basic information about each module in the HLD. Input and output are discussed as well as some other major properties of each. For detailed descriptions of module algorithms or to find out what certain parts of code are doing, please see the comments found in the source code.

6.1 Main.

The purpose of this module is to not input/output, rather; it purpose is twofold:

1. To initialize the program and some global variables.
2. To act as the root of the Tk/Tcl GUI hierarchy tree.

6.2 Filename.

The input to this module is from the user. The user chooses the values for the following items:

1. Filename.
2. Problem name.
3. Problem Description.

The filename is used as the name of the file which the user is creating. This item must be specified. The problem name is a string of characters with no spaces and less than 17 characters long. MACH uses this string as the name of a subdirectory that it creates for its output. MACH uses the problem description to let the user type an identifying line of text into the top of an input file.

The output of this module is the three items described above. The program open a file for writing with the user-supplied filename. The other two items are written to this file.

6.3 Restart.

Input into this module is user-supplied. The user enters values for variables in the $restrt namelist. The output is these values saved into the Tcl interpreter.

6.4 Namelist Menu System.

This module does not explicitly perform input/output. Its function is somewhat similar to that of the main module. Its purpose to to implement a menu system wherein the user enters data into MACH namelists. Its other purpose is to act as a root to the Tk/Tcl GUI hierarchy tree for those namelists. See comments in the source code for more on this subject.

6.5 Contrl.

Input to this module is user-supplied values to the variables in the $contrl namelist. There are 100 variables in the $contrl namelist. To let the user easily enter data, the variables have been grouped by functionality (i.e. algorithm controls, multigrid controls etc.). The user pulls down the $contrl menu and chooses any of the eleven types of variable groups. When the user is entering data into these windows, they have the choice of going to the next $contrl window by clicking on the "Next" button or by pulling down the menu again and choosing another group.

The output is these values saved into the Tcl interpreter.

6.6 Ezgeom.

Input to this module is user-supplied values to the variables in the $ezgeom namelist. This is a complex module that lets the user view the data in a canvas as it is being entered. See the comments in the source code for more details about how this module is coded.

The output is these values saved into the Tcl interpreter.

6.7 Save Variables.

This module has an important task: it saves all the MACH variable data currently stored as Tcl variables into the main C data structure. See section 4 for more on the data structures.

Input to this module is all the MACH variables and their values. The output is putting all these variables and their values into the proper place in the data structure.

6.8 Unset All.

This procedure unsets all of the MACH Tcl variables that could have been set by the user. Each call to UnsetVar clears that variable from the interpreter.

There is really no input/output to this module. The module does its work on global Tcl variables.

6.9 Modtime.

This module is different from the other namelist modules under the namelist menu system. This module is on a direct path in the HLD. Also, the user can go down the path to the remap module or continue without a remap. This module is not under the hierarchy of the namelist menu; it is directly under the main module root for the whole Tk/Tcl heirarchy tree.

Input into this module is user-supplied. The user enters values for variables in the $modtim namelist. The output is these values saved into the Tcl interpreter.

6.10 Remap.

Input into this module is user-supplied. The user enters values for variables in the $remap namelist. The output is these values saved into the Tcl interpreter.

6.11 Array To File.

Input into this module is the main C data structure: the modtime array. Output is writing this array to file.

7 User Manual.

This user manual is a guideline on how to create (or modify) a MACH input file using this program. It is not intended as an instruction manual on the MACH program variables. It assumes that you are familiar with the X-Windows system.

The program leads you through the process of creating a file. The things that must occur at the beginning of a file are prompted for first (e.g. problem name, problem description and the restart namelist). Once you get to the main menu, you can enter data for any of the namelists on the menubar. If you want to use a $modtim namelist you must first finish entering data for the namelists on the menubar. This is the same way you would do it if you were to write your own file from the top down. The rest of the program runs in a similar fashion: it prompts you for information in the order that a file would be written from scratch.

To start the program you must be at an X-windows terminal. The program runs in an X-Windows graphic environment. Go to the imt directory. There are two ways to start the program:

1. Type imt at the command prompt.

2. Type imt with one command-line argument. This argument must be a syntactically correct MACH input file.

The program is very self explainatory. Use the buttons on each window to go to the next window(s).

The first window you see is an introductory window. Click on continue.

The next window will prompt you for a filename, a problem description, and a problem name. You must enter a filename. This name will be the filename of the file you are creating. If you are modifying an existing file, I recommend you use a different name. The program will overwright any file with the same name.

If everytinng is OK, you will see a message telling you that your file has been opened. For the rest of the program you will be writing the varaible assignments to this file.

The next window asks you if this will be a restart file. If you want this to be a restrt file click yes, otherwise click no.

In the restrt window you have entry boxes for the $restrt variables. The only special thing about this window are the reinitbc and reinitcc buttons.

7.1 reinitbc/reinitcc variables for $restrt and $modtim

When you click on either of these two buttons another window will come up. Use this window to specify any number of reinitbc/reinitcc variable assignments. You will not be able to interact with any other part of the program until you have closed this window.

The area just above the "Another" and "Done" buttons shows what will be written to the file.

Click one of the 1-4 radiobuttons at the top to specify the lbdry (l boundary). Type a number for lblk (l Block). (There is no need to move the mouse into the lblk entry box. Just type.) Click on true or false to specify the value. As you move the mouse, you will see what will be written to file.

If you click on "Another", the assignment statement currently in the viewing area will be written to file. Also, the previous values will be cleared and you can enter the value for another reinitbc/reinitcc variable.

If you click on "Done", the assignment statement that is currently visible will be written to file. If you click on "Done" and there is an incomplete assignment statement visible, it will not be written to file. "Done" takes you back to the $restrt window.

7.2 Main Namelist Menu Sytem

This window is the main loop for the program. It lets you enter data into the eight namelists listed. When you are finished with these namelists, go to the File menu and make your choice:

1. Continue (Modtime). This button will save all the data you have entered up to this point. Also, it will open the $modtim namelist window.

2. Done (Save and Exit). This button will save all the data you have entered up to this point then exit the program.

3. Exit leaves the program without a save.

7.3 Contrl

Enter data into this namelist by pulling down the menu and opening the windows that you need. Just fill in the data you want. Any items you leave blank will not be written to file.

In any if the $contrl windows, you can click on "Next" to get to the next window in the namelist. Or, you can go back to the pull-down window and choose what you need.

7.4 Ezgeom.

The first window you see will prompt you for several values: nblk, npts, numarcs, units, and max and min for the scale. Max X, Max Y, Min X and Min Y are the respective max amd min cooridnates of the points on your MACH problem geometry. You must enter all this data before you will be allowed to continue. After the data has been entered correctly click OK.

You now have opened three or four windows (depending of you want any arcs). Use the points window (Just under and to the left of the draw window) to type in data for each of the points in your problem. Use the Return key to move between the entryboxes. You should see the points appear on the draw window at the cooridnates you have just specified. If the points do not appear, you may haved typed bad data for that point or maybe you have incorrectly specified the max and min points. You can use the Define Indices button to check the max and min values.

Use the scroll bar in the points window to get to the rest of the point entry boxes.

You can use the mouse to add points to your problem. Click on the "Add Points" button. Notice the Draw Mode at the bottom-left of the draw screen. It should be in Add points mode.

Click on the draw window to add points. Notice that the values of the points you have just specified appear in the points window.

You can also use the mouse to drag points to new positions on the screen. Click on the "Move Points" button. Check to see that the draw mode has changed. Now, you can drag a selected point around the screen. Notice that the value of the point you are dragging will update as you are moving the new point.

Minimize the points window.

Put the corners window on top. Enter data into the corners window. As you finish each block (four corners) you should see the block written to the draw window. If the block does not appear, you probably typed in a bad point or a point that has no value.

You can add blocks using the mouse. Click on the Add Block button. Check to see that the draw mode has changed. Click on a selected point for the top-left corner. The next 3 points clicked will be the other three corners. As in the points window, the values of the corners will change (or be added) as you create blocks with the mouse.

If you specified any arcs, enter the arc data. You should see a line segment that represents the arc.

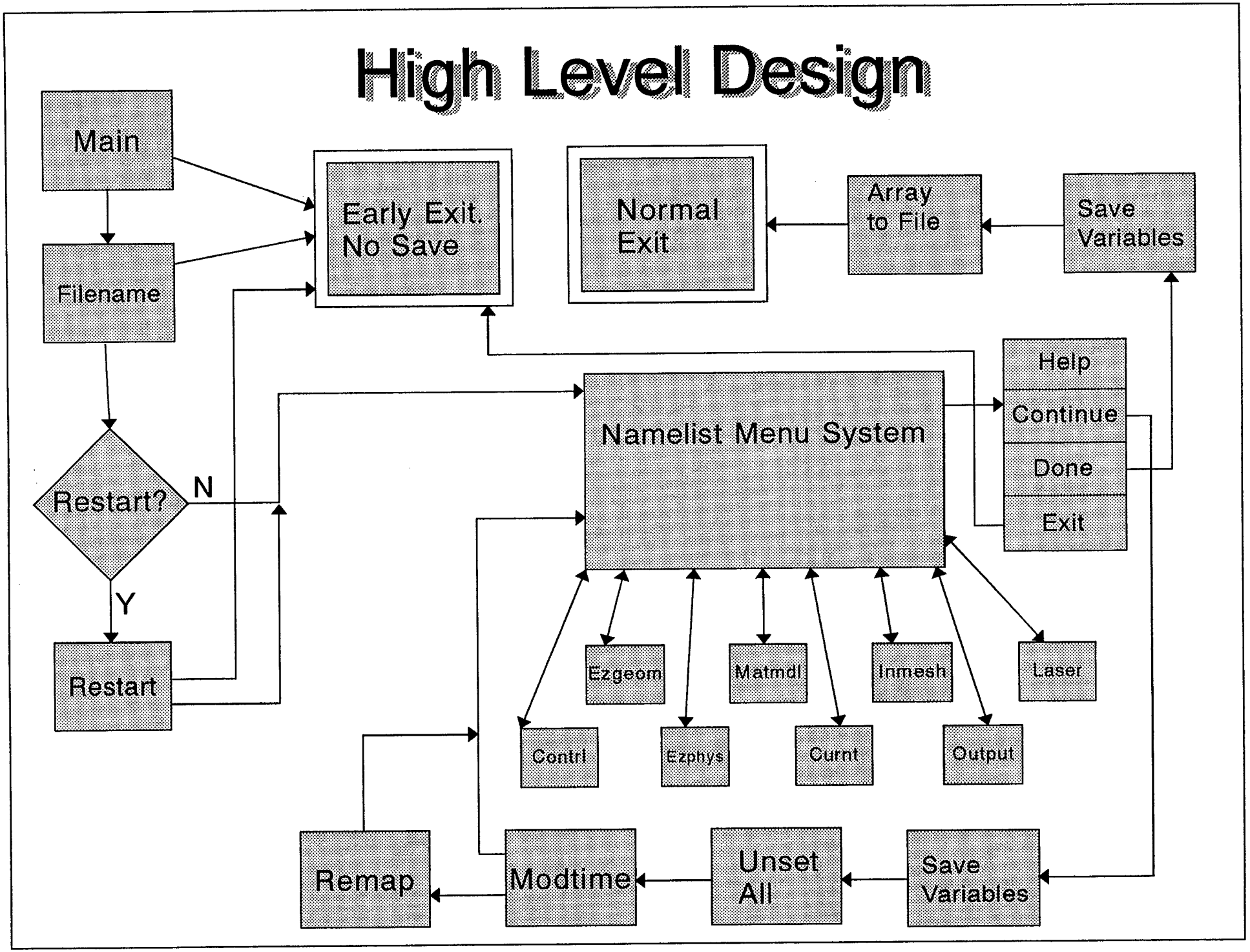
Click the OK button to leave $ezgeom.

7.5 Modtime/Remap

The modtime window lets you add data to the modtime namelist. See section 7.1 for information on reinitbc/reinitcc.

The $remap namelist must come after the $modtim namelist, so choose remap if you want. If not, choose (no remap).

High Level Design



ON THE FLUID DYNAMICS OF SUPERCRITICAL DROPLET BREAKUP: AN ANALOG WITH MISCIBLE FLUIDS

Dimos Poulikakos
Professor

and

Chad J. Mitts
Graduate Student

Department of Mechanical Engineering
The University of Illinois at Chicago
842 West Taylor Street
Chicago, Illinois 60607-7022

Final Report For:
Summer Faculty and Graduate Student
Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington DC

and

Phillips Laboratory

August 1995

ON THE FLUID DYNAMICS OF SUPERCRITICAL DROPLET BREAKUP: AN ANALOG WITH MISCIBLE FLUIDS

Dimos Poulikakos and Chad Mitts
Department of Mechanical Engineering
The University of Illinois at Chicago
842 W. Taylor Street, Chicago IL 60607

Summary

Atomization of liquid propellants and oxidants in rocket engines occurs in a high temperature and pressure environment, at conditions often exceeding the critical temperatures and pressures of the working liquids. It is therefore important that an extensive knowledge base exists on a host of relevant phenomena, exemplified by atomization in sprays undergoing transcritical transition. Such a knowledge base is currently very limited. The work reported herein aims at the above deficiency and reports preliminary results on droplet deformation and breakup in a miscible fluid analog of actual transcritical and supercritical droplet breakup. More specifically, this analog consists of a liquid droplet falling freely into a reservoir of another liquid in which the droplet is miscible. Two combinations of such liquids were examined in the context of the present study: sucrose solution droplets falling in a water reservoir, and petroleum oil droplets falling in a pentane reservoir. The study reveals a host of distinct breakup regimes for the falling droplet depending on the values of the droplet Reynolds number and the viscosity ratio of the miscible liquids. In addition, the importance of solubility becomes apparent and it is recommended that its effect is investigated in a future study.

1. Introduction

As the operating pressures and temperatures in contemporary liquid rocket engines and gas turbines increase towards supercritical conditions, an understanding of the basic mechanisms and regimes of droplet atomization under these conditions becomes a crucial component of cost effective, efficient design. Currently, the majority of the design codes utilize models for fuel atomization that were developed for subcritical conditions. These models do not include the major features of supercritical atomization phenomena such as the vanishing of surface tension, similar properties between the gas and liquid phases, the high solubility of the gas phase in the liquid phase which can lead to a mixture with an increased critical pressure, and property singularities near the critical point. To begin addressing this problem, the behavior of a single liquid droplet under transcritical and supercritical conditions is examined. The knowledge accumulated is a building block toward the understanding of the more complex sprays and jets present in actual injectors in rocket propulsion.

The main focus of the research was the development of a zero surface tension analog of the real droplet deformation and breakup process, through the use of miscible fluids. An additional scope of the work was the

characterization of a novel piezoelectric LOX droplet generator, built at Phillips laboratory. The goal of the miscible fluid supercritical analog work is to characterize the breakup mechanisms of a zero surface tension droplet over a wide range of viscosity ratios and Reynolds numbers. By establishing similitude with the real supercritical case through the proper selection of miscible fluids, it is hoped that the mechanisms of supercritical droplet deformation and breakup can be captured. The investigation of miscible fluids as a supercritical droplet simulation was initiated at Phillips laboratory [1]. The goal of studying the LOX droplet generator is to gain a better understanding of the stable operation of the device so that its use (at a later time) in the study of supercritical phenomena is reliable and devoid of unexpected complexities. To this end, two experiments were established, one which allowed the evolution of a zero surface tension droplet to be observed over a wide range of viscosity ratios and Reynolds numbers, and another which was capable of producing and recording the images of a continuous stream of LOX droplets at atmospheric conditions.

2. Background

Supercritical Droplet Breakup Analog

Traditionally, the deformation and breakup of liquid particles (droplets) has been characterized by the density and the viscosity of the particle and the fluid surrounding, the particle effective diameter, the velocity of the particle relative to the surrounding fluid, and the particle surface tension. In addition, the acceleration of the particle has been included as a parameter when it is important. These properties are grouped into the following nondimensional numbers [1]: the density ratio (DR), the viscosity ratio (VR), the Weber number (We), the Eotvos number (Eo), and the Reynolds number (Re). In many situations, the Ohnesorage number (Oh) is commonly used in place of Re in describing the breakup regimes of subcritical droplets. All of these numbers are defined below.

$$DR = \rho_p / \rho \tag{1}$$

$$VR = \mu_p / \mu \tag{2}$$

$$We = \rho D V^2 / \sigma \tag{3}$$

$$Eo = a \rho_p D^2 / \sigma \tag{4}$$

$$Re = \rho V D / \mu_p \tag{5}$$

$$Oh = \mu / (\rho_p D \sigma)^{1/2} \tag{6}$$

The symbols ρ, σ, μ, D, V and a stand for density, surface tension, viscosity, particle diameter, particle velocity and particle acceleration respectively. The subscript p denotes the particle (droplet). Under transcritical and supercritical conditions, the surface tension approaches zero, thus We, Eo, and Oh all approach infinity and are no longer a useful means of characterizing the droplet behavior. In addition, densities of the liquid phase and the gas

phase under supercritical conditions approach the same order of magnitude, thus the density ratio also becomes uninformative. This suggests that the evolution of a liquid particle in supercritical conditions can be solely described based on the viscosity ratio and the Reynolds number [1]. This appears reasonable since these two numbers describe the influence of the two primary forces that govern particle deformation and breakup in the absence of surface tension, the inertial forces which tend to induce breakup, and the viscous forces which tend to resist particle breakup.

To determine the accuracy with which the deformation and breakup of a zero surface tension particle can be described through the combination of the viscosity ratio and the Reynolds number, a simple experimental setup was developed that utilizes two miscible fluids to simulate the zero surface tension condition of a supercritical environment. The premise of this model is that the surface tension, which is caused by an asymmetry of molecular forces at the interface of two fluids [2], will vanish as the distinct fluid interface disappears due to the effect of the fluids dissolving into each other. While this method does not replicate the thermodynamic singularities or thermal transport properties relevant in supercritical droplet study, it should still capture the purely kinematic nature of particle deformation and breakup. Furthermore, the miscibility of the two fluids is similar to the high solubility of the gas phase into the liquid phase.

Piezoelectric LOX Droplet Generator

One of the primary obstacles in the study of supercritical oxidants and fuels is the stable generation of a monodispersed stream of droplets. To meet this challenge, a piezoelectric droplet generator was developed at Phillips laboratory that is capable of producing LOX droplets by condensing gaseous oxygen (GOX). It utilizes a concentric tube heat exchanger arrangement, with a continuous flow of LN2 serving as the cooling fluid for the GOX. The advantage of this device is two fold. First, LOX is an extremely dangerous substance. Concentrated amounts of oxygen can easily oxidize many items. By condensing GOX instead of feeding the generator directly from a LOX reservoir, the amount of LOX present in the laboratory is greatly reduced yielding a much safer practice. Second, gaseous oxygen and LN2 are readily available in most laboratory environments, whereas LOX is more difficult to obtain and maintain.

Once the stream of oxygen is condensed in the heat exchanger, it passes through a cylindrical piezoceramic tube that is pulsed by a function generator. With each pulse, the tube contracts and relaxes, thus inducing perturbations in the LOX stream. In principle, the optimum frequency of perturbation needed to induce Rayleigh breakup was determined by Weber [3] as a function of the jet diameter and velocity and it is given below in equation (7).

$$\frac{1}{f} = \frac{\sqrt{2}\pi d}{v}\left(1 + \frac{3\mu_L}{\sqrt{\rho_L \sigma d}}\right)^{0.5} \qquad (7)$$

The symbols f, v, ρ, μ, and d refer to the optimum frequency of oscillation, the mean jet velocity, the fluid density, viscosity, and the diameter of the jet respectively. In this experiment, only one orifice diameter was available, therefore, the jet diameter remains relatively constant. Thus, a change in the piezo oscillation frequency must be balanced by the appropriate change in the mass flow rate of the oxygen in order to excite the optimum instability leading to a Rayleigh breakup mode. The effect on the droplet stream due to changes in the frequency (for $Re_{jet} < 10^3$) should be reflected mainly in a change in the velocity of the droplets. The diameter of the droplets will largely be determined by the orifice diameter.

3. Experimental Setup and Procedures

Supercritical Droplet Breakup Analog

A significant advantage to the study of zero surface tension particles using the miscible fluid analog is the simple and inexpensive experimental setup that is required (Fig. 1). The experimental setup consisted of a transparent, Plexiglas test section of four inches square in cross section (normal to the direction of motion of the particle) and twelve inches in height. Above the test section, a two mm diameter flat tip tube and syringe were mounted on a calibrated rail such that the height of the tube above the surface of the fluid in the test section, denoted as h in Figure 1b, could be easily adjusted and measured. The particle fluid would collect at the end (tip) of the tube until the force of gravity overcame the surface tension and caused the droplet to release and fall towards the test section. A strobe light was mounted behind the test section, and it was used in combination with a Stanford Computer Optics 4 Quik 05 CCD camera with an image intensifier to record the evolution of the particle as it traveled through the test section. The image acquisition system consisted of a Data Translation QuickCapture Frame Grabber board installed in a Macintosh 7200/80 Power PC (Fig. 1). The image acquisition / analysis software used was NIH Image V1.57.

Two different sets of miscible fluids were used in the experiments: the first was a sucrose droplet falling into water, and the second was a petroleum oil droplet in pentane. In each of the cases, the substance with the greater density was used as the particle fluid so that the it would fall freely through the test section. Measured quantities of the ambient fluid were added to the particle fluid to change its viscosity. The viscosity and density were then measured as a function of the volume percentage of particle fluid to ambient fluid. The viscosity measurements were taken using a Cannon-Feaske viscometer according to the standard guidelines established in ASTM D446-74. In addition, for both the oil/pentane mixture and the sucrose water/mixture, the initial diameter of the particle formed from the tube was measured using the CCD camera and the NIH image analysis software.

For both the sucrose/water and the oil/pentane combinations, particle deformation and breakup images were digitally recorded for particle fluid volume percentage mixtures ranging from 100% to 20%, and for values of h ranging between one cm and six cm. The varying mixtures and heights provided a viscosity ratio range from 2.2 to

843 and a Reynolds number range from 2 to 1760 for the sucrose and water combination. For the oil and pentane experiments, the viscosity ratio ranged between 2.9 and 835 and the Reynolds number between 6.2 and 2,100 . The image data were used to characterize the breakup regimes for the different conditions, and to quantify various dimensional aspects of the particle.

Piezoelectric LOX Droplet Generator

A droplet generator was constructed and mounted to a stand such that the stream of LOX droplets was directed downward. A flow control valve on the gaseous oxygen feed line allowed the flow rate to be easily adjusted. A function generator and an amplifier were used to produce a square wave of 10 μs duration with a variable frequency and amplitude to drive the piezoelectric ceramic device. Reference 4 contains a detailed description of the droplet generator. A schematic of the experimental setup and droplet generator is given in Figure 2. Quantitative measurements of the droplet diameters were made using a Greenfield Instruments Model 700 Spray Analyzer. This device consisted of a CCD camera and a strobe light combination that sent shadowgraph images of the droplets to a Gateway 2000 PC for analysis using the Greenfield Instruments software. The frequency of the piezo oscillations was varied between 200 Hz and 2000 Hz. At each different frequency, the flow rate of the oxygen was adjusted to attain a Rayleigh breakup mode, and the diameter data was collected.

4. Results and Discussion

Supercritical Droplet Breakup Analog

The breakup regimes for the sucrose/water combination are shown in Figure 3 as a function of the viscosity ratio and the Reynolds number. The four distinct regions identified were viscous globules, viscous bag, vortex transition, and vortex ring behavior. The oil/pentane data are presented in Figure 4 again as a function of the viscosity ratio and the Reynolds number. The same basic breakup modes, viscous dominated and vortex ring, were identified for both combinations of fluids, however, the transition from the viscous dominated regime to the purely vortex behavior is much more distinct in the case of the oil/pentane combination. The transition region also occurs over a considerably smaller range of Reynolds numbers for the oil/pentane case, spanning values of 60 to 90, as compared a Reynolds number range of 85 to 500 for the sucrose/water experiments. The noted differences between the two combinations of miscible fluids under similar viscosity ratios and Reynolds numbers indicate that an additional mechanism, the solubility of the two fluids, is affecting the breakup of the particles beyond eliminating the surface tension. Therefore, the complete characterization of the deformation and breakup of a zero surface tension particle must include the solubility in addition to the viscosity ratio and the Reynolds number.

In the absence of surface tension, the primary force that resists breakup of the particle is the internal viscous force. The effect of solubility is that it tends to alter the viscosity of the particle fluid as the ambient fluid is

dissolved into it. Thus, the ability of the particle to resist deformation from the inertial forces is also altered. To demonstrate the significant effect that solubility can have, the distance from the free surface of the test section to the location that primary breakup of the particle occurs is nondimensionalized with respect to the particle initial diameter and plotted as a function of the viscosity ratio in Figure 5. The disparity of the distance to breakup between the oil/pentane and sucrose/water data is roughly an order of magnitude different for viscosity ratios of 400 and greater. At smaller viscosity ratios, the solubility rates are lowered as the concentration gradient between the particle and the ambient fluid is reduced. As a result of the smaller concentration gradients and the relatively weaker internal viscous forces at the lower viscosity ratios, the particle deformation is dominated by the inertial forces and the development of the particle occurs at a fast enough rate, such that the solubility is not a significant factor, thus, the data for the two different fluid combinations begin to converge.

Figure 6 illustrates the viscous globule breakup mode for the sucrose/water combination at a VR of 843 and a Re of 4.8, and it is a typical example for VR of 400 and greater. As the initially spherical particle enters the test section fluid, the pressure distribution around the particle will exhibit a maximum at the stagnation point near the center of the particle leading edge. The mass at the trailing edge of the particle begins to redistribute itself around the periphery of the particle in response to the leading edge mass that has been slowed down at the stagnation point. This redistribution creates the flat disc of particle fluid shown in Figure 6a. Corresponding to the initial stagnation point analysis, a minimum pressure would exist along a circumferential region slightly aft of the leading edge surface. However, due to the relatively high viscosity ratio of this case, the internal viscous force is great enough such that this suction has little effect on the particle development within the viscous globule regime. Due to the internal flow of mass towards the particle periphery, the center of the particle thins until it is no longer capable of resisting the shear stresses created by the pressure difference between the forward and aft sides of the particle, thus it consequently breaks. As a result of the new azimuthal distribution of mass, the maximum pressure shifts from the original stagnation point to the leading edge of the of the circumferential ring that has developed. The pressure distribution around the ring inherently exhibits variations, thus the ring becomes unstable and it breaks into separate, smaller viscous globules as demonstrated in Figure 6c. The smaller globules behave as the initial particle and generally, they would continue to break into smaller components at least one more time.

Figure 7 is the breakup of an oil droplet as it travels through pentane. The initial VR and Re are 835 and 2 respectively, which is similar to the sucrose particle conditions of Figure 6. The initial deformation is very similar between the two cases, with the initially spherical particle flattening out into a disc shape in response to the stagnation point at the leading edge. However, the pentane appears to dissolve into the oil at a faster rate than the water into the sucrose. This results in a diminished internal viscous force and as indicated in Figure 5, the particle breaks into the smaller globules at a much shorter distance than the sucrose particle. In addition, the shape of the disc is more bag like than in the sucrose trials due to the greater effect of the low pressure region present at the aft of the particle combined with the reduction in the particle viscosity. The smaller globules that result after primary

breakup of the initial particle are consequently less viscous, and therefore, the secondary breakup is rapid and is dominated by the dissolving action of the ambient fluid.

As the viscosity ratio is reduced, the shape of the particle becomes even more non-spherical as indicated in Figure 8 showing a sucrose particle at a VR of 144 and a Re of 15, and in Figure 9, showing an oil particle with a VR of 157 and a Re of 42. Due to the decreased initial internal viscous force, both particles are influenced by the low pressure region aft of the particle and the resulting shape is that of a bag. However, the sucrose particles developed a much more distinct bag shape, and it was an identifiable breakup mode for a VR of 20 to 200. As the sucrose bag traveled through the test section, the periphery would thin out until the viscous forces were no longer able to maintain a coherent shape, thus causing the periphery to collapse inwards under the pressure differential between the leading and trailing edges. The mass that had constituted the thin periphery would then tear away from the main particle and form a cloud of slow moving mass that dissolved quickly. The center of the bag, which contained a concentrated amount of mass, would then exhibit the viscous globule behavior described earlier.

For the oil particles, the bag behavior existed only in the vicinity of a VR of 150, and as indicated in Figure 5, the oil case tends to breakup at a distance less than that of the sucrose data. Furthermore, due to differences in the effect of solubility between the two different particle fluids (oil and sucrose), the thinning action of the bag development caused the oil particle to be dissolved faster than in the sucrose particle. Therefore, upon breakup of the primary oil particle, the entire amount of mass exhibited a puff-like behavior, and it quickly dissolved. The bag mode observed in the oil/pentane experiments did not constitute a primary breakup regime as it did in the sucrose experiments due to the bag's shorter life, reduced incident of occurrence, and greater ambiguity. Rather, all of the globule and bag modes were grouped into the general viscous regime denoted in Figure 4.

At viscosity ratios of approximately 30, the bag behavior is still the primary breakup mode for the sucrose particles, as seen in Figure 10. However, at the same viscosity ratio, the oil particles exhibit a vortex ring behavior as illustrated in Figure 11. Full vortex ring behavior was observed for the oil/pentane trials at a VR of 120 and a Re of 95. The vortex ring structure is similar to that observed Peck and Sigurdson [5] for a water droplet falling into water. The vortex ring is initiated by the impact with the free surface, which creates a discontinuous velocity gradient within the particle as a result of the shear forces acting on the particle fluid that is in contact with the ambient fluid. The particle fluid which is not in contact with the ambient fluid continues downward at a greater velocity until it comes in contact with the ambient fluid. The result is a recirculation vortex that is stabilized by the pressure differential between the leading and trailing edges of the azimuthal ring. At the higher viscosity ratios discussed previously, the internal viscous forces of the particle were sufficient to dampen out the differences in inertial forces between the mass in contact with the ambient fluid and the internal mass of the particle. Therefore, the vorticity that would result was diminished.

The oil particles exhibit vortex behavior at higher viscosity ratios and lower Re than the sucrose particles due to solubility. As stated earlier, the solubility effect reduces the viscosity of the particle fluid, and thus

diminishes its ability to dampen out the vorticity created by the inertial and pressure differences within the particle. Even for a viscosity ratio as low as VR=14, as shown in Figure 12, the sucrose particles have still not attained the vortex structure. At that point, the primary mode is neither vortex nor bag dominated. This was therefore termed as the transition regime. In Figure 12c gross vortex motion is observable, however, the particle mass quickly dissolves and the motion never becomes sustained. Figure 13 illustrates the stability of the vortex ring for the oil particles at a VR of 14. It should be noted that the Reynolds number is 320 for Figure 13, whereas for the sucrose particle of Figure 12, Re was only 170. At similar viscosity ratios, the Reynolds number is much greater for the oil/pentane experiments than the sucrose/water case due to fact that while the initial velocities and diameters of the two particle fluids are similar, the oil viscosity is 80% less than that of the sucrose.

At a VR of 7.5 and a Re of nearly 300, the sucrose particles develop a stable vortex ring. The VR at which the vortex behavior is dominant for the sucrose/water combination is over 90% smaller than the value for the oil/pentane combination. Similarly, the Re number that begins the vortex regime for the sucrose particle is 60% greater than that of the oil case. This is a clear demonstration of the major role that solubility plays in characterizing the deformation and breakup modes of a zero surface tension particle. Furthermore, the nature of the vortex mode is also different between the two fluids. As shown in Figure 14, the presence of mass nodes within the vortex structure is distinct, and it was observed to be the dominant factor governing the depth at which the vortex convected through the test section before breaking up. The mass nodes were not as distinct in the oil vortex rings, and while the dynamics responsible for the nodes has not yet been sufficiently explained, it is likely that the presence of solubility will determine to what degree they are present.

Piezoelectric LOX Droplet Generator

The droplet generator readily produced a stream of spherical LOX droplets, as typified in Figure 15. However, the uniformity of size within the stream, and the spacing between droplets exhibited instabilities. Figure 16 is a plot of the length and the Sauter mean of the LOX droplet stream for a wide range of frequencies. The average sample size for the different frequencies was 5200 drops. The striking feature of the data is that the length mean is relatively constant over the span of frequencies for large droplet sample sizes. This is a reasonable result and it corresponds to the Rayleigh breakup theory which indicates that for a single value of the jet diameter, there is an optimum value of the wavelength of the perturbations required to induce a stable breakup. The wavelength is a ratio of the jet velocity and the oscillation frequency, therefore, a change in one parameter must be accompanied by the corresponding change in the other to maintain the optimum wavelength. Thus, increasing the frequency of oscillation and the mass flow rate of the oxygen for a constant orifice diameter will result in a longer jet and an increased droplet velocity. However, to alter the diameter of the droplets markedly, the orifice diameter should be altered.

The variation in the Sauter mean reflects the variance in droplet diameters, whereas the length mean damps out influence of the larger droplets over a large sample size. The fluctuation in droplet sizes can be attributed to three sources. First, to achieve the flow rate of oxygen required for the optimum perturbation wavelength, a more precise control mechanism than the current flow control valve is needed. Second, it has been suggested by Dr. Talley's group at Phillips laboratory that the location at which the condensation of the GOX occurs may fluctuate enough to incur instabilities in the droplet stream. This effect can be reduced by increasing the length of the oxygen line within the heat exchanger. Finally, the orifice used contained many blunt surfaces and therefore, the optimum disturbances induced by the piezoelectric device may be damped out or altered. To reduce this effect, an orifice with smoother contours should be implemented. The next step in this study would be to characterize the droplet velocities and diameters as a function of the orifice diameter and the piezo oscillation frequency.

5. Conclusions

As consequences of supercritical conditions, the surface tension vanishes and solubility between the gas and liquid phase increases. In simulation of this phenomenon, two miscible fluids were utilized to create zero-surface tension particles (droplets). In the absence of surface tension, the primary forces present during the deformation and breakup of a particle are the internal viscous forces and the inertial forces. The breakup regimes of sucrose particles falling into water and oil particles falling into pentane are given as a function of the VR and the Re. The primary regions observed are the viscous regime, transition, and vortex behavior. However, the point at which the various regimes occurred varied between the sets of fluids as a result of the solubility, which tends to alter the viscosity of the particle fluid. Therefore, to characterize the breakup regimes of a zero surface tension droplet, and presumably a droplet undergoing transition in a supercritical environment, the solubility rate must be included with the viscosity ratio and the Reynolds number.

For the piezoelectric LOX droplet generator, the length mean diameter of a large sample size of droplets remained constant at approximately 250 μm over a wide range of piezo oscillation frequencies for the six μm orifice used. This corresponds to the Rayleigh breakup mode which indicates that the diameter of the droplets is a function of the jet diameter. Thus, to alter the jet diameter, the orifice of the generator must be changed. However, the stream of droplets was not entirely uniform in size, and these variations are indicated in the fluctuation of the Sauter mean diameter. The instabilities present resulted from the inability to precisely control the mass flow rate of the oxygen, the variation in the location at which GOX condensation occurs, and blunt corners of the orifice. As these issues are addressed, the generator will become an invaluable resource in supercritical droplet breakup study.

6. References

1. Talley, D. G., R. D. Woodward, and J. C. Pan, "Supercritical Droplet Behavior," AFOSR Task 2308AP (1995).
2. Probstein, R. F., Physicochemical Hydrodynamics," Butterworths, 1989.
3. Lefebvre, A. H., *Atomization and Sprays*, Hemisphere Publishing Corporation, N.Y., 1989.
4. Woodward, R. D., D.G. Talley, T.J. Anderson, and M. Winter, "Shadowgraphy of Transcritical Cryogenic Fluids," presented at the 6th Annual Symposium of the Pennsylvania State University Propulsion Engineering Research Center, NASA LeRC, Cleveland OH, (1994).
5. Peck, B., and L. Sigurdson, "The Vortex Ring Velocity Resulting From an Impacting Water Drop," *Experiments in Fluids*, vol. 18, pp. 351-357, 1995.

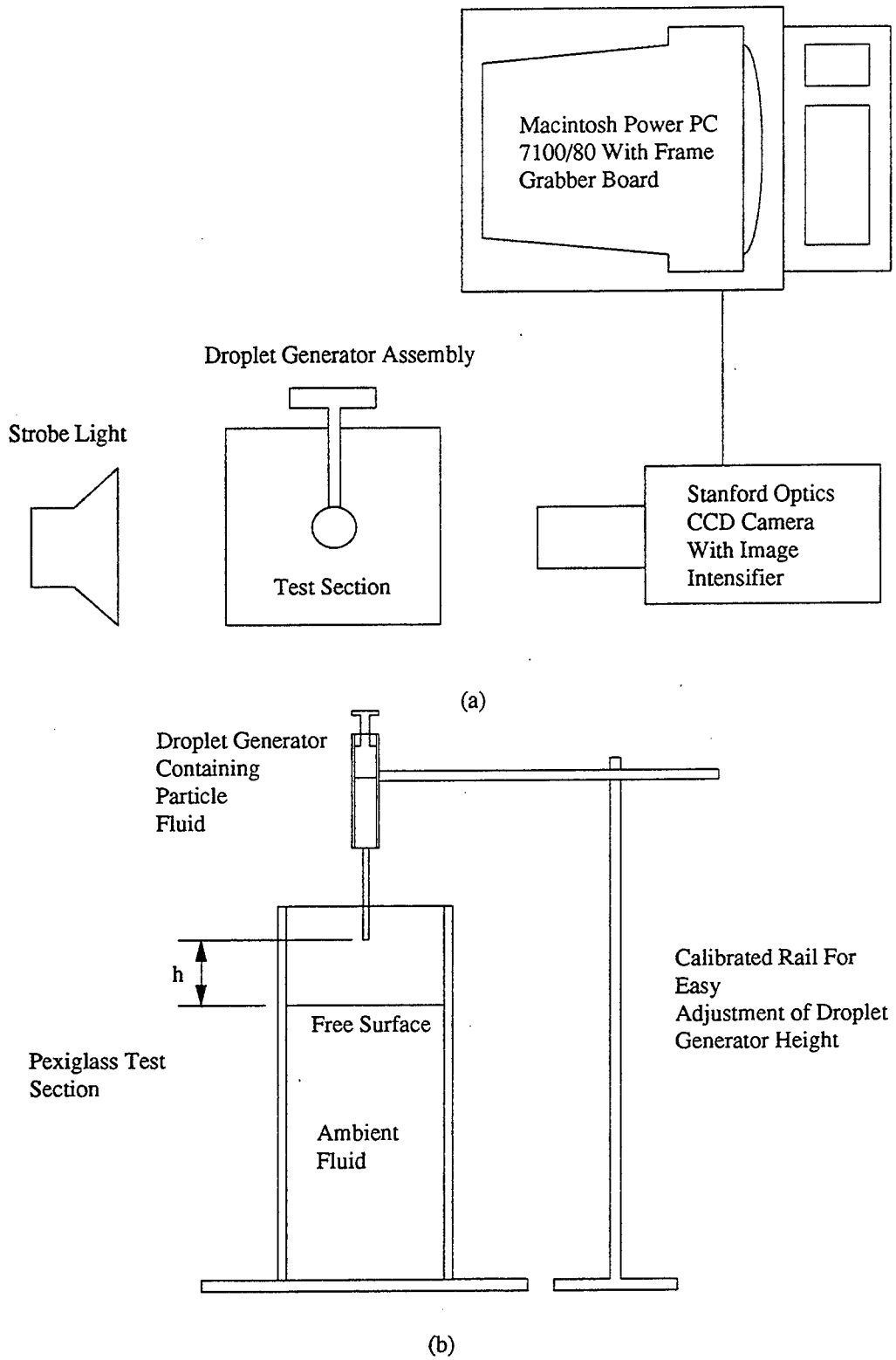
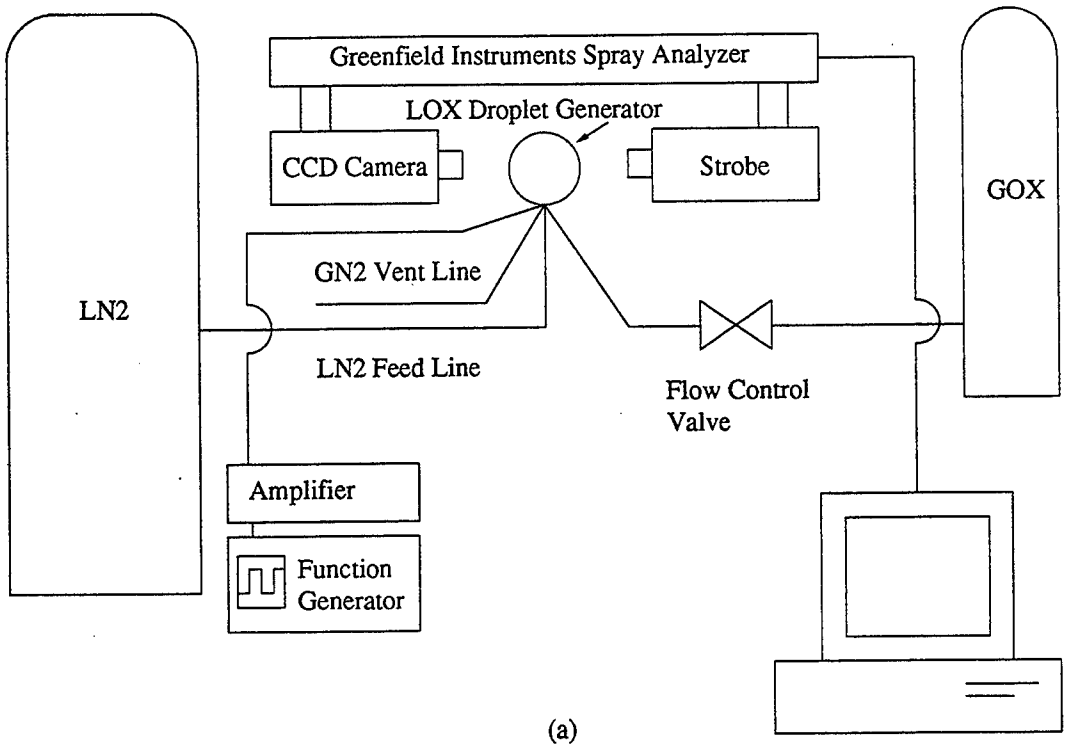
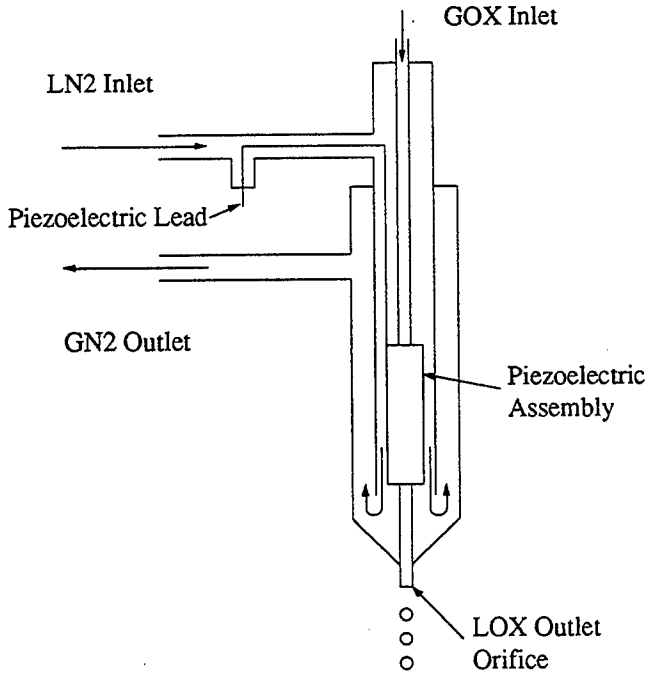


Figure 1. (a) Schematic of experimental setup (b) Schematic of droplet generator and test section



(a)



(b)

Figure 2. (a) Schematic of LOX droplet generator experimental setup (b) Schematic of LOX droplet generator

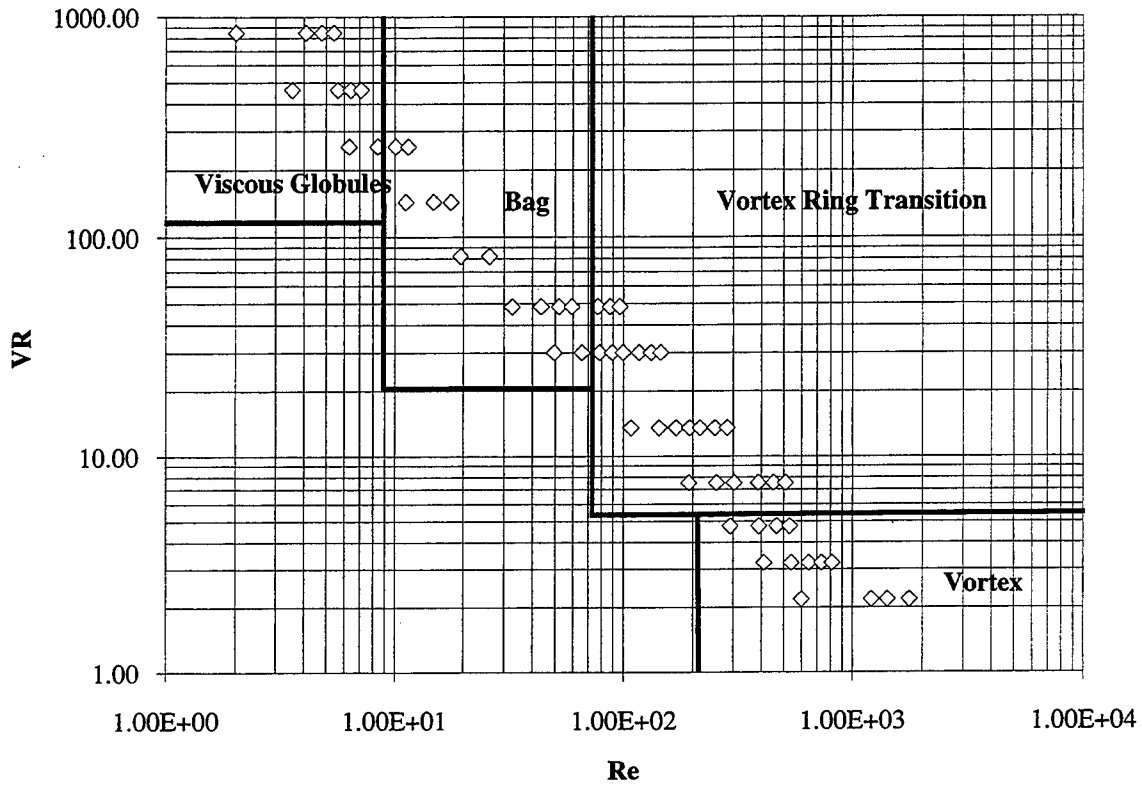


Figure 3. Primary breakup modes of a zero surface tension sucrose particle as a function of VR and Re

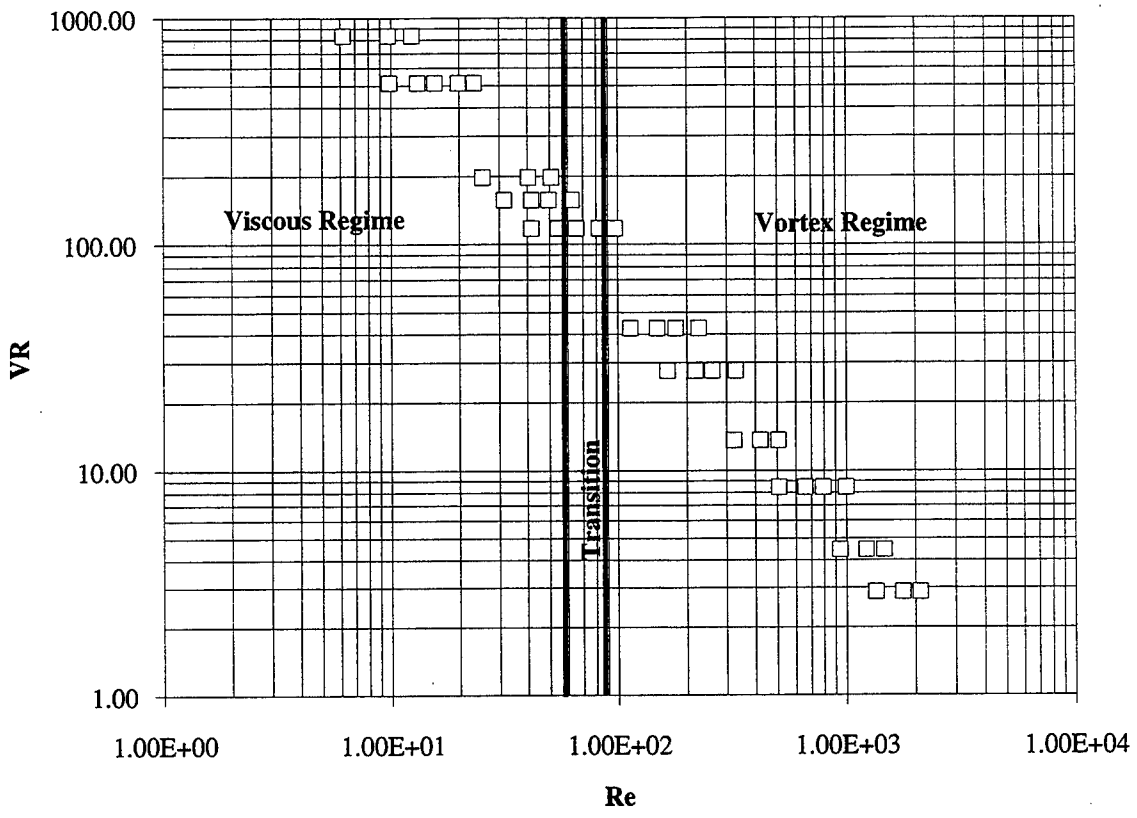


Figure 4. Primary breakup modes for a zero surface tension oil particle as a function of VR and Re

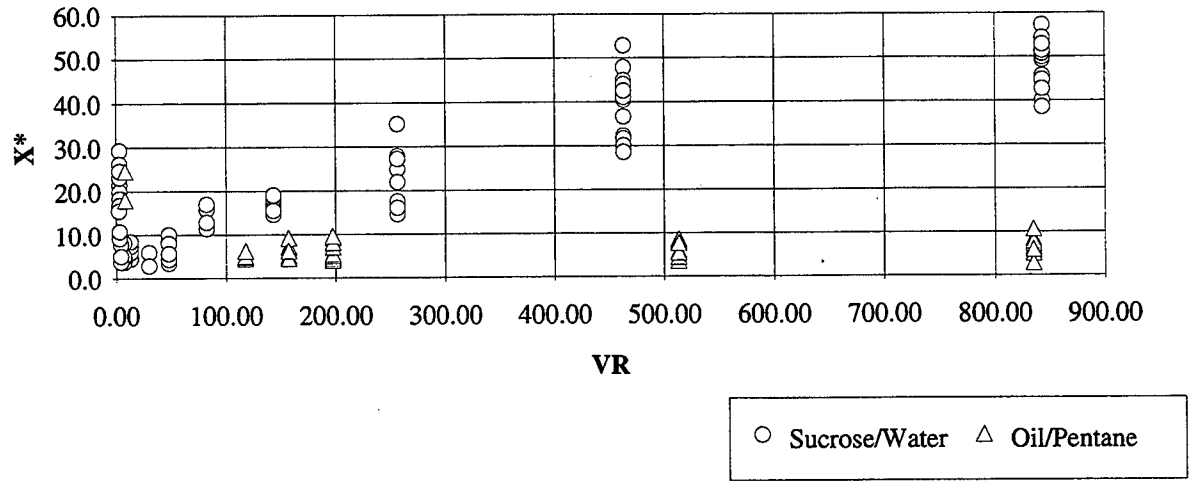
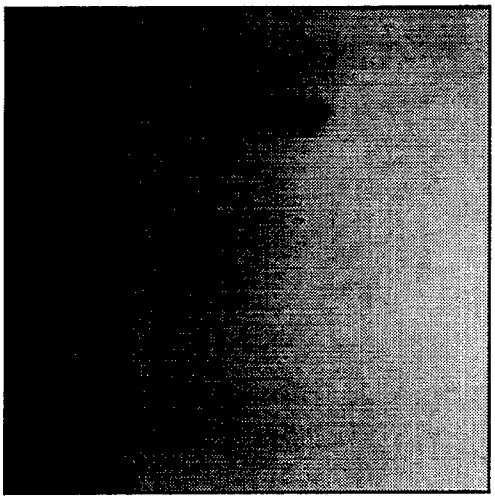
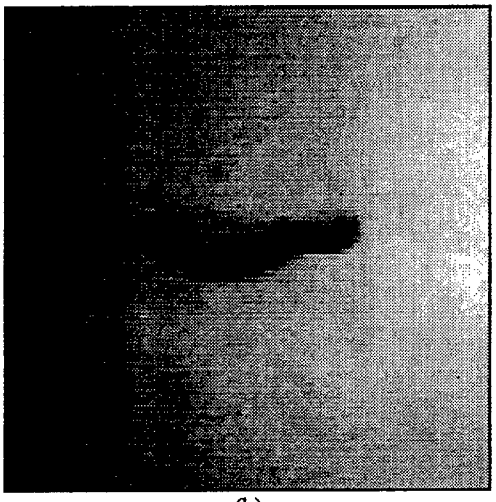


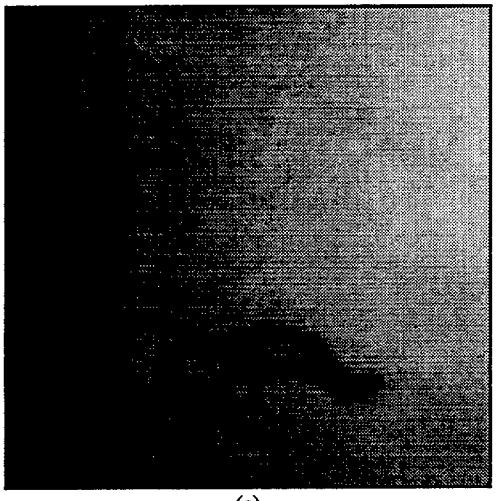
Figure 5. Nondimensional distance to primary breakup, X*, as a function of VR



(a)

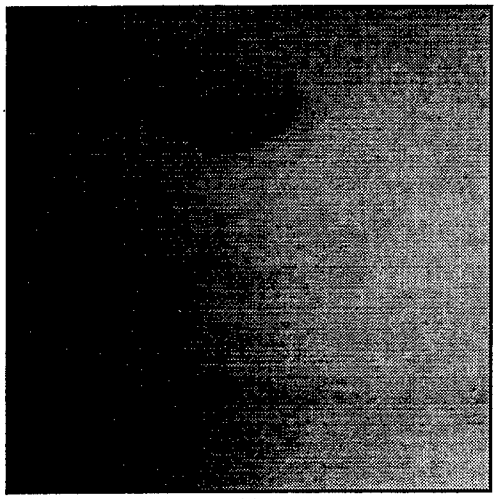


(b)

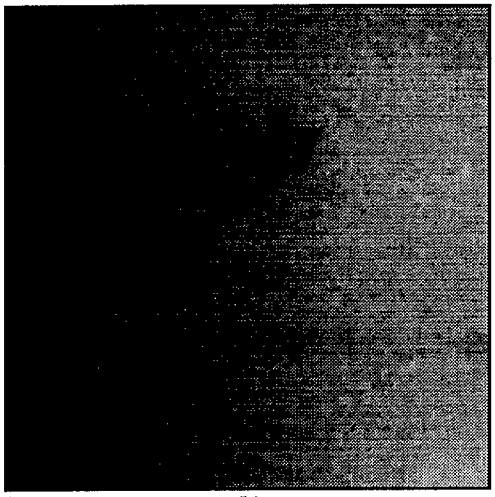


(c)

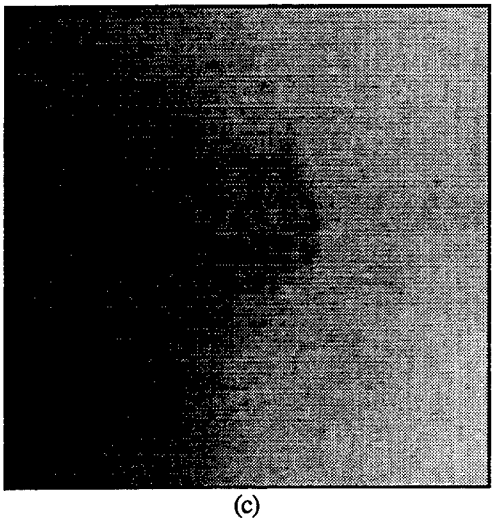
Figure 6. Deformation and breakup of a zero surface tension sucrose particle falling freely in water with initial conditions of VR=843 and Re=5



(a)

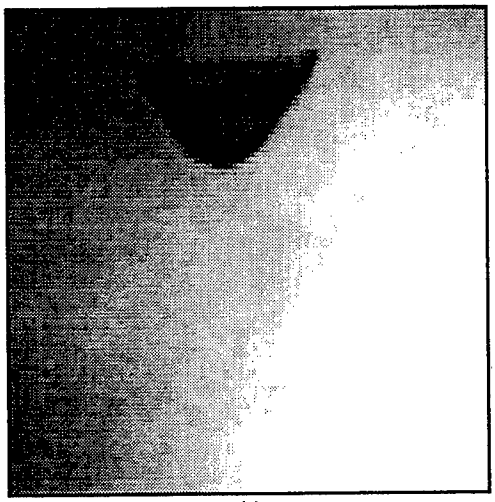


(b)

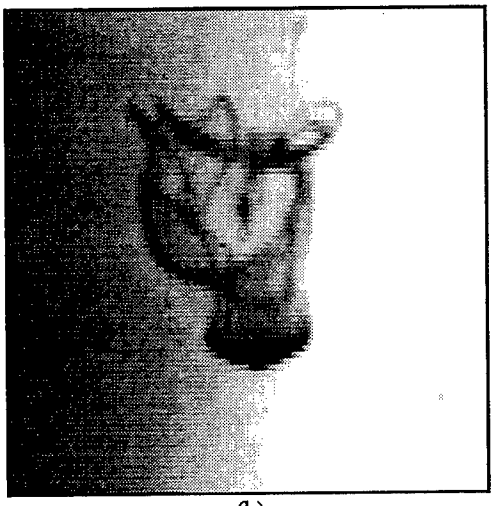


(c)

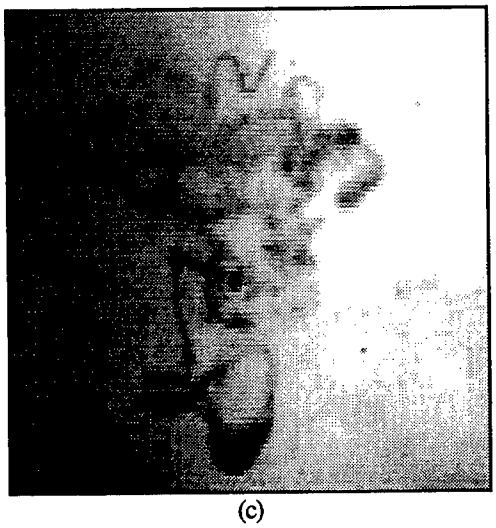
Figure 7. Deformation and breakup of a zero surface tension oil particle falling in pentane with initial conditions VR=835 and Re=6



(a)

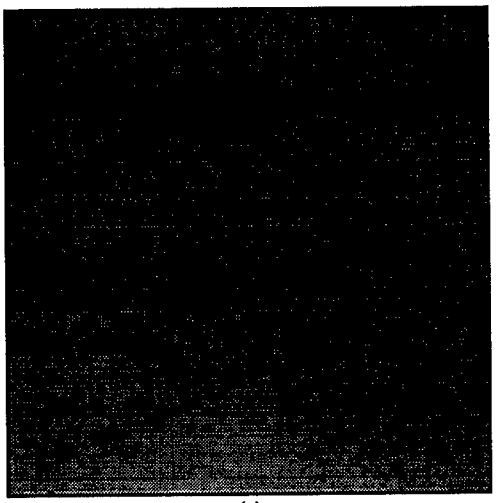


(b)

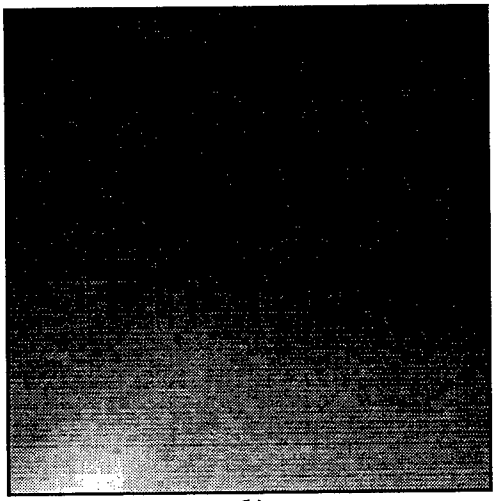


(c)

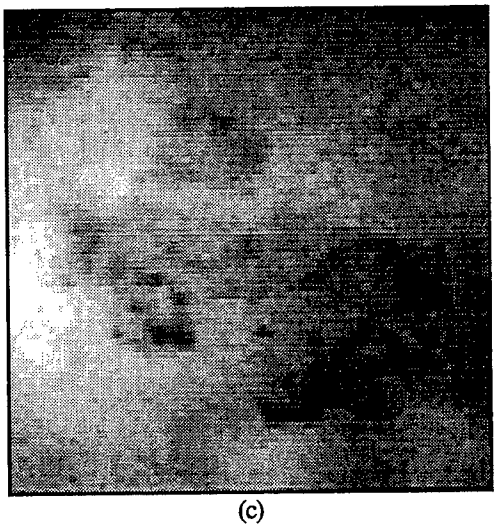
Figure 8. Deformation and breakup of a zero surface tension sucrose particle falling in water with initial conditions of VR=144 and Re=15



(a)



(b)

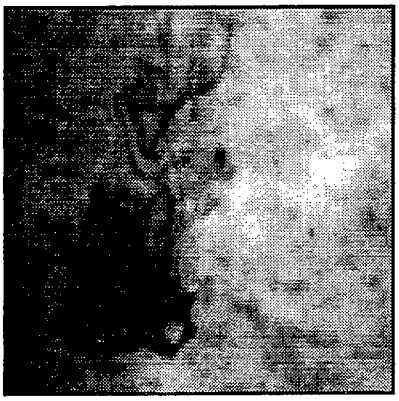


(c)

Figure 9. Deformation and breakup of a zero surface tension oil particle falling in pentane with initial conditions of VR=157 and Re=42

(a)

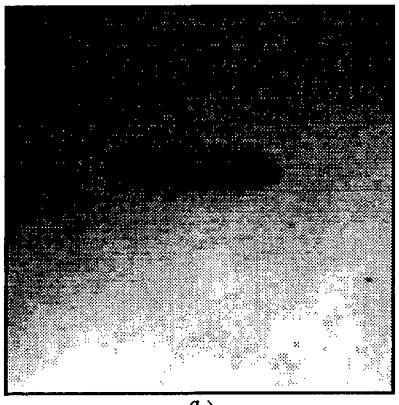
(b)



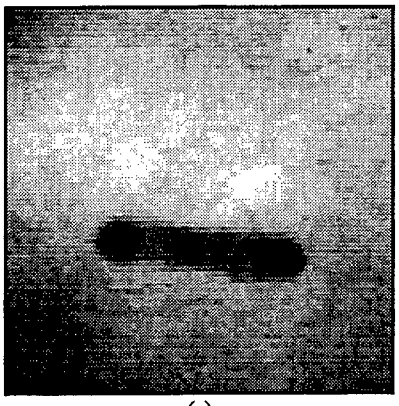
(c)

Figure 10. Deformation and breakup of a zero surface tension sucrose particle falling through water with initial conditions of VR=30 and Re=66

(a)

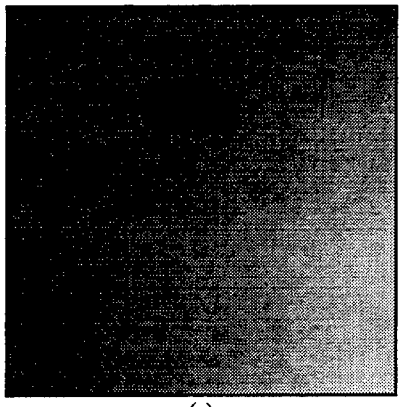


(b)

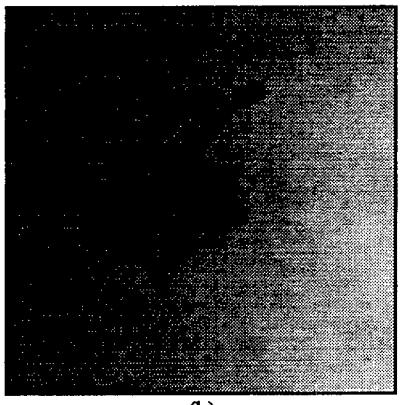


(c)

Figure 11. Deformation and breakup of a zero surface tension oil particle falling through pentane with initial conditions of VR=28 and Re=260



(a)



(b)

(c)

Figure 12. Deformation and breakup of a zero surface tension sucrose particle falling through water with initial conditions of VR=14 and Re=170

Laboratory Studies of Current Driven Plasma Processes with the Versatile Toroidal Facility (VTF)

Daniel T. Moriarty
Graduate Student
Nuclear Engineering Department

Massachusetts Institute of Technology
167 Albany St.
Cambridge, MA 02139

Final Report for:
Graduate Student Research Program
Philips Laboratory, Hanscom AFB, MA

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

August 1995

Laboratory Studies of Current Driven Plasma Processes with the Versatile Toroidal Facility (VTF)

Daniel T. Moriarty
Graduate Student
Nuclear Engineering Department
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Abstract

The Versatile Toroidal Facility (VTF) has been constructed to create plasmas for wave-plasma interaction studies. Three plasma sources may be combined to produce various conditions. The plasma density gradients, field aligned currents, and other geometries and parameter ratios closely match the plasma conditions of the auroral and upper ionospheric regions. Current driven plasma waves have been observed in VTF and instability mechanisms are being studied. Spectral analysis reveals broadband emissions from below the ion cyclotron frequency, f_{ci}, to near the electron cyclotron frequency, f_{ce}. Waves with frequencies corresponding to the whistler and lower-hybrid ranges are consistent with the well known inverse Landau damping of energetic charged particles. Existing theory is modified to investigate current convective modes in VTF. Multi-point cross-correlation measurements have been performed to determine wave numbers of the excited modes and have qualitatively confirmed the modified theory. Comparison of VTF experiments with rocket experiments shows VTF can be used to appropriately study ionospheric plasma processes, and complement active experiments in space.

Laboratory Studies of Current Driven Plasma Processes
with the Versatile Toroidal Facility (VTF)

Daniel T. Moriarty

1 Introduction

The following is a report studying current driven plasma processes with the Versatile Toroidal Facility (VTF). The Ionospheric Plasma Studies group at the MIT Plasma Fusion Center built VTF with the objective to supplement our recent field experiments at Arecibo, Puerto Rico, Millstone Hill, and elsewhere; to explore phenomena observed in shuttle and rocket experiments; and to corroborate theoretical studies of plasma wave processes in the ionosphere.

High power RF ionospheric modification phenomena such as parametric decay, resonant absorption, and anomalous microwave absorption have been studied in both the ionosphere and VTF [*Moriarty*, 1992]. Particle precipitation and other current driven phenomena such as those found naturally in the auroral electrojet or simulated with electron beam experiments performed in space with shuttle- or rocket-deployed payloads contain similarities with hot cathode beam injection plasmas in VTF. Broadband emission from current-containing plasmas and its relevance to theories developed by *Maggs* [1976], *Kindel and Kennel* [1971], *Chaturvedi et al.* [1979,1980], *Ossakow et al.* [1979], and *Keskinen et al.* [1980,1982,1983], is the focus of this report.

Section 2 is an introduction to the laboratory facility, VTF. VTF experiments exploring current driven instabilities and ionospheric turbulence are described in Section 3. Similar phenomena studied in space plasmas will be described in Section 4. The ECHO 7 experiment obtained data most similar to that found in VTF. A review of theories predicting emission in the various frequency ranges is in Section 5. A modified collisional-ion fluid instability is then presented. A review of the experiments as well as a few conclusions is in Section 6.

Table 1: Versatile Toroidal Facility Parameters

Major Radius (R_o)	0.9m
Minor Radius (r)	0.3m
Toroidal Field (B_o)	1800G
Vertical Field (B_v)	20G
2.45 GHz RF Power	3kW
Beam power	50 kW
Electron Density	$10^{10} - 10^{12} cm^{-3}$
Bulk Electron Temperature	$\sim 10eV$

2 Characteristics of VTF

VTF is a medium sized ($R_o = 0.9m$) toroidal plasma device in which ionospheric processes have been studied. Typical operating parameters are given in Table 1. Its vacuum vessel, with a major radius of 92 cm and minor radius of 30 cm, is approximately 6 cubic meters in volume and has 48 ports, giving access to the top, side, and bottom. The vacuum system is capable of creating a base pressure of less than 5×10^{-7} Torr. The toroidal field is produced by eighteen "picture frame-" shaped electromagnets "wound" around the vacuum vessel. A weaker vertical field is produced by a set of Helmholtz coils wound externally. Inductive current drive can be provided via a central solenoid wound with 45 turns capable of 1 Volt-sec with "double swing". An additional set of coils are available for poloidal field shaping.

Plasmas may be produced by any of three sources: RF, hot-cathode beam injection, and Taylor discharge (inductive drive). The RF is injected from a 3kW 2.45 GHz magnetron launched in the O-mode. An electron beam plasma is produced by biasing a hot LaB_6 cathode, which emits into a helical magnetic field. A typical density profile for a beam plasma as a function of radius is shown in Figure 1.

The VTF beam plasma system may provide seconds-long plasmas with densities as high as ten times that of RF plasmas. The principal elements of the electron beam system are several LaB_6 filaments installed at the bottom of the vacuum chamber. Each filament is heated to emit electrons that are subsequently accelerated by a 300 Volt potential applied between each filament and the vacuum chamber. Electrons travel upwards along the helical magnetic field, producing field-aligned currents with intensities exceeding 5 amp/m^2. The geometry of the beam plasma source is shown in Figure 2.

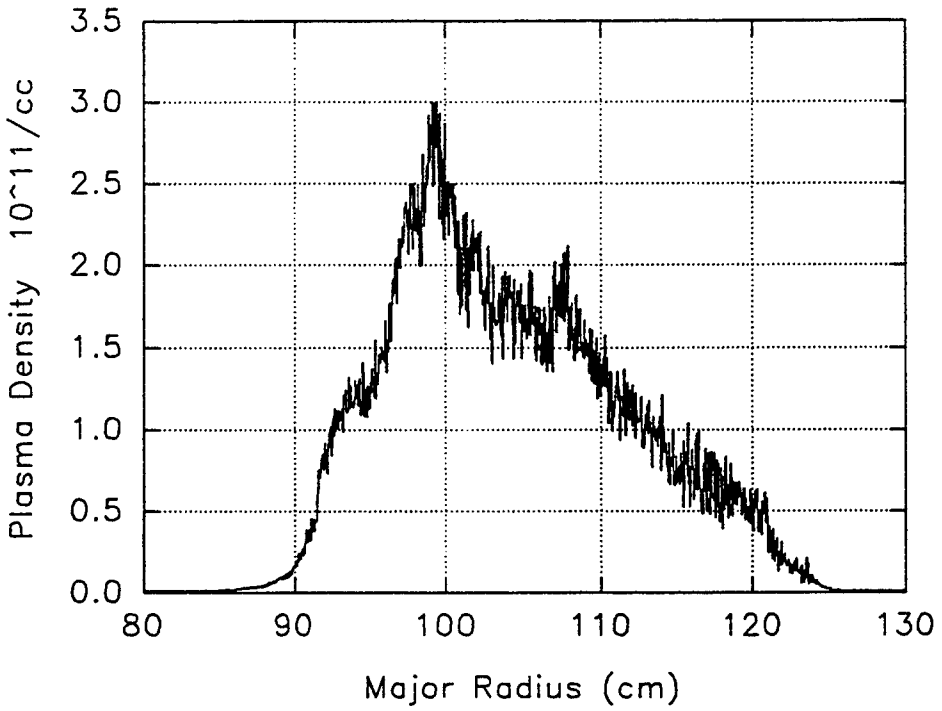


Figure 1: Plasma density profile for a beam produced plasma.

2.1 VTF as an Ionospheric Plasma Laboratory

VTF as an ionospheric plasma laboratory is useful to cross-check wave-injection experiments at Arecibo, Puerto Rico; to explore naturally occurring phenomena; to compare with Shuttle and Rocket experiments; and to support theoretical studies. Arecibo experiments have been set aside since the Observatory has not been available.

The following ionospheric parameters are produced in VTF:

(1) $\nabla n \perp \vec{B}_o$ (2) $f_{pe} \geq f_{ce}$

(3) $\frac{\nu_i}{\Omega_i} \sim 10^{-2}, \frac{\nu_e}{\Omega_e} \sim 10^{-3}$ (4) $2\% < \frac{n_e}{n_o} < 20\%$

(5) H, Ar, O, orNplasma (6) $DC\vec{E}_o$

The first condition says that the density gradient is perpendicular to the imposed DC magnetic field. In actuality plasma density gradients exist both perpendicular and parallel to the Earth's magnetic field. In the auroral region and mid latitudes the variation of density with altitude corresponds to a parallel density gradient; however, there typically exists a much larger gradient perpendicular to the Earth's field (usually North-South). Near the equator the altitude variation in density is perpendicular to the Earth's magnetic field, which is nearly horizontal. Further, the

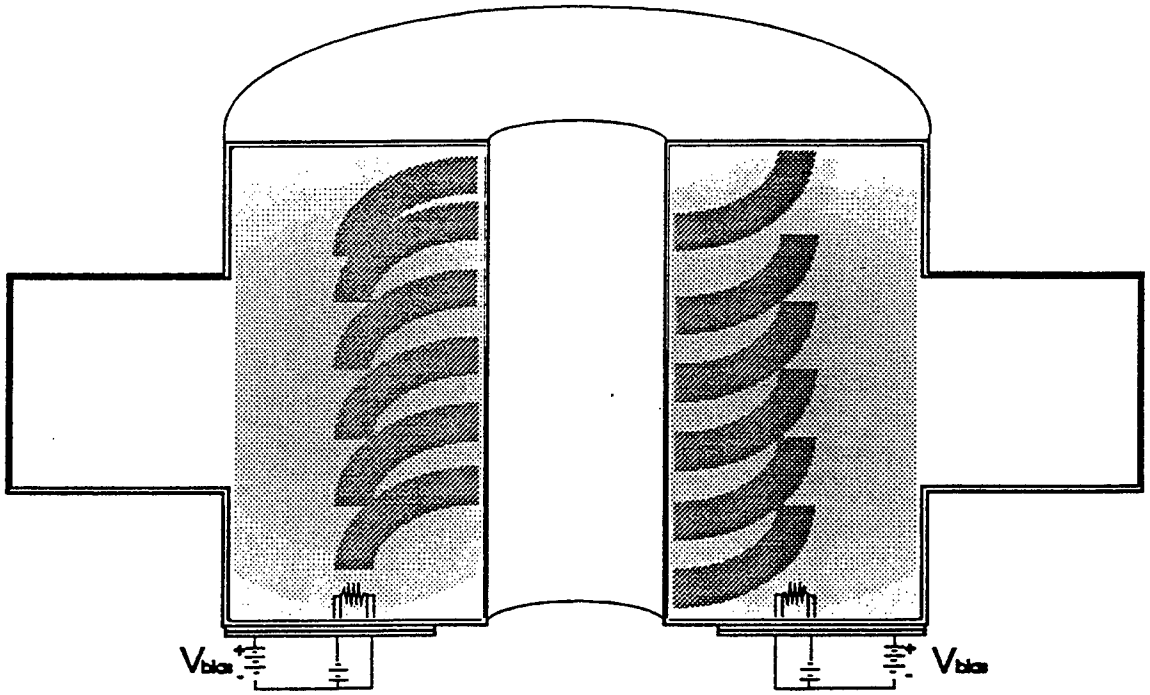


Figure 2: Cross-section of beam system showing two filaments emitting a helical beam.

parallel density gradient does not act as a source mechanism for driving plasma wave instabilities. One typically neglects the parallel gradient in comparison to the perpendicular gradient.

The first parameter ($\nabla n \perp \vec{B}_o$), is easily obtained in a toroidal plasma vessel. As for the second, ($f_{pe} \geq f_{ce}$), in the ionosphere the ratio can be as high as 8 to 1. Producing such a plasma is more difficult in the lab since it is the magnetic field that confines the plasma, allowing high densities. In the case of the RF-produced plasma, the magnetic field must be such that a resonance layer exists to absorb the radiation. Thus for RF plasmas in VTF the ratio is roughly one. The electron beam plasmas require a minimum toroidal field to prevent arcing to the grounded chamber walls. Good beam plasmas can produce a plasma to cyclotron frequency ratio of 3.

The third and fourth conditions tend to go hand in hand. The (relatively) high neutral pressure makes for low fractional ionization and high collision rates. Plasmas can be made with various gasses. Hydrogen has been traditionally used. For electron beam plasmas Argon ionizes more easily and consistently produces high current beams.

Lastly, the biased hot-cathodes produce parallel electric fields and strong perpendicular electric fields similar to those produced by neutral winds and other mechanisms in the ionosphere.

3 VTF Experiments

Experiments have been performed in VTF to explore the current convective instability. The beam plasmas contained a broad spectrum of turbulence: roughly DC to 2 GHz. A large portion of the excited waves can be explained with well known theories mentioned below. The waves with frequencies below the lower-hybrid resonance frequency are of primary interest and are discussed in more detail.

3.1 Steady State Spectra of Current Driven Modes

The data presented in this section were obtained from seconds long beam plasmas. A radially scanning potential probe provided signals to a spectrum analyzer whose outputs were then digitized. A typical potential fluctuation spectrum produced by the beam plasma is shown in Figure 3. Since there is no background spectra all waves are excited by the beam; and the intensities of the waves vary with the beam current. Notice that the radiated whistler and lower-hybrid wave spectrum resembles that of the ECHO 7 experiment reviewed in Section 4.1. In addition to the lower-hybrid and whistler waves, VTF beam plasmas produce low-frequency modes below f_{lhr}, the lower-hybrid resonance. These low-frequency modes will be discussed in detail in Section 5. The frequency at which the intensity drops off does not change with changing plasma conditions, suggesting a limitation of potential probes consistent with *Roth and Krawczonek* [1971].

3.2 Pulsed Experiments

Pulsed experiments were performed on VTF to explore the linear growth of the current driven waves. The motivation arose when it became clear that the steady state condition of the plasma was non-linear due to the strong electric fields of the driven plasma modes. In principle, by observing the growth of the instabilities, a more accurate comparison to the linear theories could be made.

Shown in Figure 4 is the timing setup for the pulsed experiments. Three traces are shown. The tallest trace shows the repelling bias voltage applied to the filament (arc voltage). Its rise time is typically 100 μsec. The next tallest signal is that of the spectrum analyzer trace while set to zero span, centered on 5 MHz. The third trace is that of the current emitted from the filaments (arc current). The first feature of interest is that the time scale of changes to the 5 MHz signal is

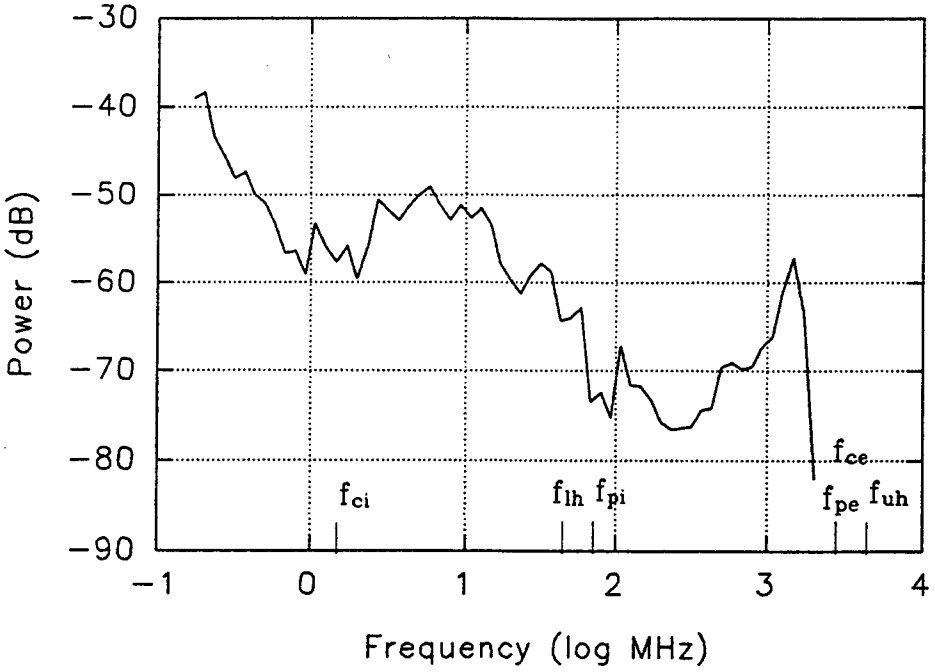


Figure 3: Log plot of beam plasma potential fluctuations.

comparable to the time changes in the arc current - and corresponding changes in plasma density. A second feature to note is the time scale of the horizontal axis - milliseconds. The growth rate of the waves, calculated in Section 5 is faster than changes in beam current. Therefore we conclude is that the waves see a steady state plasma as it slowly develops. So at any given moment the waves have already reached non-linear amplitudes.

Further insight can be obtained from the pulsed experiments. A comparison of the 5 MHz signal in Figure 4 to the 500 MHz signal shown in Figure 5 shows a clear difference in the development of the two modes. In the case of the 500 MHz signal, the wave quickly develops then dies away; by contrast, the 5 MHz signal gradually develops as the plasma density grows. One may speculate that the filaments first emit a fast beam which preferentially excite whistler radiation, then as the plasma develops and the current becomes more of a drifting Maxwellian - the high energy electrons slow down and heat the plasma - the low-frequency modes grow. Laboratory experiments by [*Kellogg et al.* 86] discussed in Section 4 exhibited a similar phenomena termed the "front edge effect".

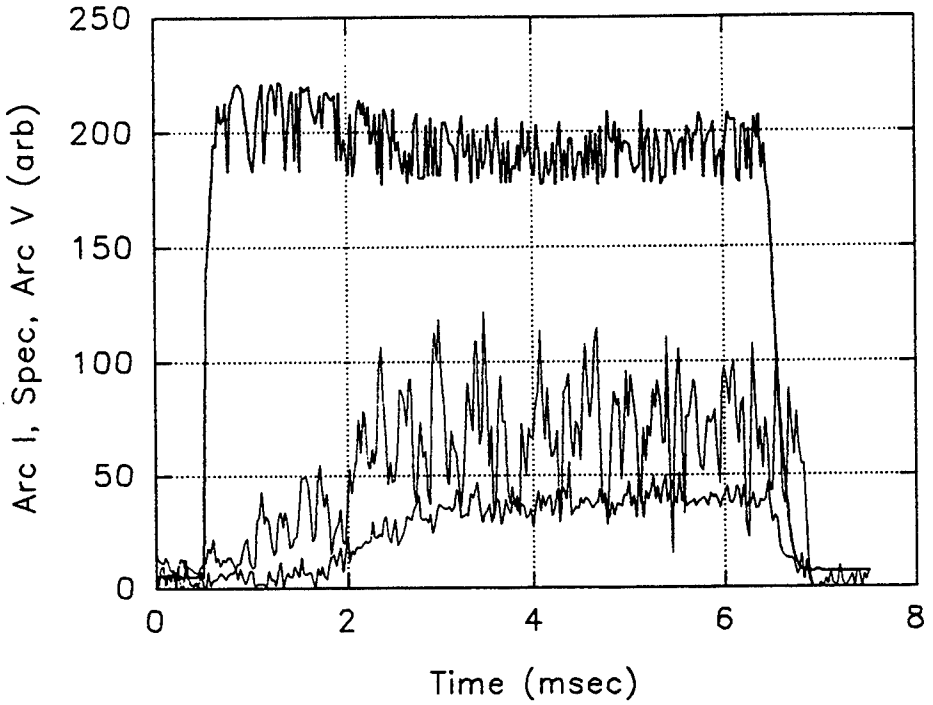


Figure 4: Timing of pulsed beam plasmas with the 5 MHz signal intensity displayed.

3.3 Wavenumber measurements

Two point potential correlation measurements near 1 MHz have been carried out in a beam plasma. Two separate types of measurements were taken: along the magnetic field to determine $k_{\|}$, and across for $k_{\perp}$. The first measurement showed the two signals were clearly in phase, suggesting a parallel phase velocity, $\frac{\omega}{k_{\|}}$, greater than that resolvable by the instrumentation. In this case the measurement predicts parallel phase velocities much faster than the ion acoustic speed. A similar measurement made with the same two tips spaced vertically, or across the magnetic field showed no time delay between the two signals suggesting a perpendicular phase velocity, $\frac{\omega}{k_{\perp}}$, that is also much faster than expected. A likely explanation of this unreasonably fast perpendicular phase velocity is that the waves are standing perpendicularly. Standing waves, described as $A\cos(kx)\cos(\omega t)$, have uniform phase in all space as opposed to travelling waves, $A\cos(kx - \omega t)$, which shift phase as they move in space. A more lengthy description of the measurable differences between standing and travelling waves can be found in [*Moriarty* 1992]. The large measured parallel phase velocities rule out the possibility of the low frequency modes being inverse Landau damped ion acoustic waves. Measurements of the standing wavelengths, which are in progress, will provide the necessary

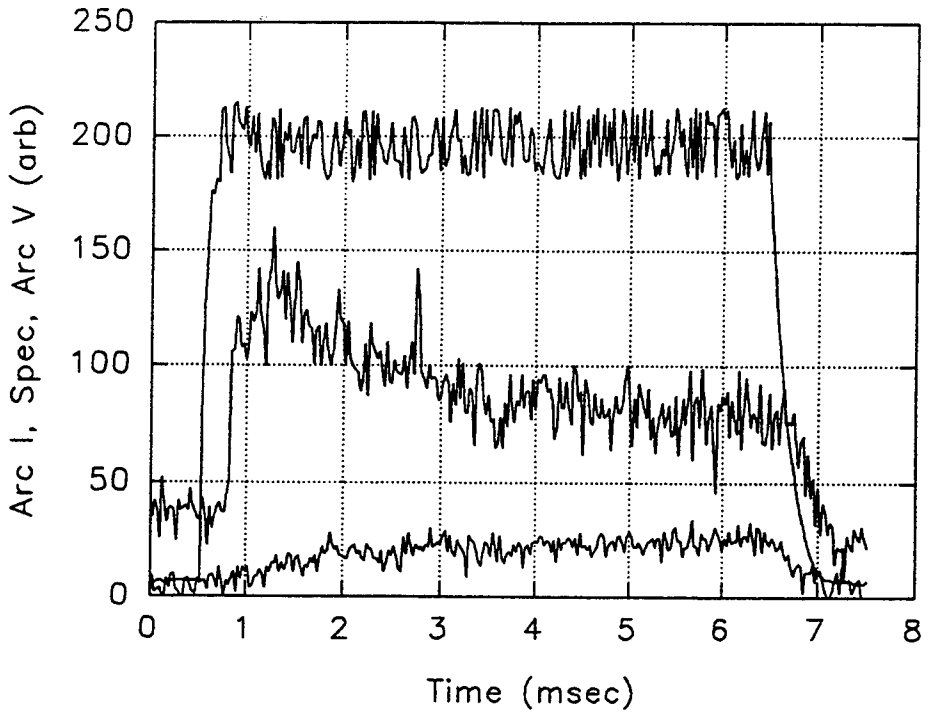


Figure 5: Timing of pulsed beam plasmas with the 500 MHz signal intensity displayed.

dispersion.

4 Current Driven Instabilities in Space

4.1 Rocket Experiments

Rockets have been launched into active aurora carrying payloads to inject electron beams and record the resultant wave emissions. Several rockets have been launched during the ECHO program. Two such missions are reviewed here: ECHO 6 and ECHO 7. Each rocket was launched from the Poker Flat Research Range in Alaska [*Winckler et al.*, 1984, *Winckler et al.*, 1989, *Winglee and Kellogg*, 1990, and *Ginet & Ernstmeyer*, 1991].

The ECHO 6 experiment in March, 1983, revealed some broadband wave emissions but had limited spectral resolution and resulted in mostly speculative conclusions. One particular shortcoming of rocket experiments was revealed. Because of the high velocities of the experimental apparatus there are two major limitations: spatial variations of the local plasma potential appear

as low frequency fluctuations via the $v \cdot \nabla\phi$ doppler effect: and the short interaction time of the injected beam in a single location prevents longer time scale waves from being excited, (e.g. ion waves). Because of this limitation little has been developed toward the understanding of naturally occurring ion and other ELF waves in the aurora.

The ECHO 7 program was a rocket-deployed electron beam-injection experiment. A payload containing two electron guns injected beams, one along the magnetic field and one at an angle to the field. A Plasma Diagnostic Package (PDP) recorded electric fields with a five tipped array of potential probes (Figure 6). The PDP in the figure has two orthogonal sets of electric probes in the spin plane of the payload. A fifth probe projects upward along the PDP spin axis. These probes consist of insulated rods terminating in spherical sensors coated with graphite. The main payload (denoted by MAIN) injected two independent electron beams: one (gun 2) at a pitch angle of 100° or 110° and the other (gun 1) with a pitch angle nearly parallel to the magnetic field. Gun 1 operated at 36 kV and 250 mA, and gun 2 was swept from 40 kV to 8 kV with a corresponding current in each 1 ms interval.

Presented in Figure 7 are typical electric field power spectra measured on the PDP during beam injection at a pitch angle of 110° [*Winckler et al.*, 1989; *Ginet and Ernstmeyer*, 1991]. The plotted spectra include four different measurements of E_x and E_y denoted by a dashed line, solid line, dot-dashed line, and dotted line. Here, E_x and E_y refer to the electric field components measured by the two orthogonal sets of electric dipole antennas. The dashed line represents the Fourier transform of the DC E_y component measured over a 200 ms interval every 0.4 ms. The solid line is the Fourier transform of the VLF broadband receiver measurement of E_y sampled every 0.05 ms over 100 ms. The dot-dashed line is the E_x component measured by the high swept frequency analyzer (HSFA) with a frequency resolution of 30 kHz and a sweep time of 102 ms between 60 kHz and 5.5 MHz. A composite background spectrum constructed in the same manner from data taken when there was no beam injection is shown by a dotted line.

Marked in Figure 7 on the top of the frame are the ion gyrofrequency (f_{ci}), the lower hybrid resonance frequency (f_{lh}), the ion plasma frequency (f_{pi}), the electron gyrofrequency (f_{ce}), the electron plasma frequency (f_{pe}), and the upper hybrid frequency (f_{uh}). Compare the data from space experiments with those recorded in our laboratory experiments as displayed in Figure 3. It appears that they are quite similar. But careful examination of them as elucidated below indicates

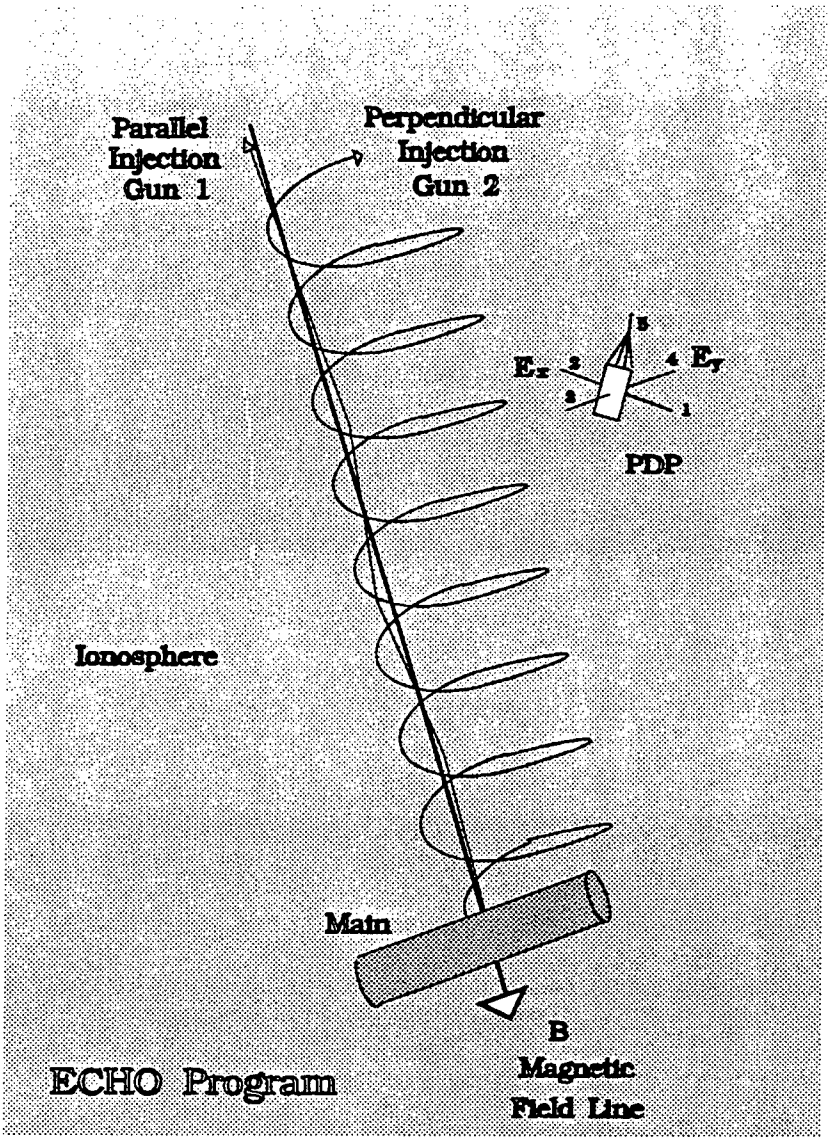


Figure 6: Cartoon sketch of the ECHO 7 experimental setup

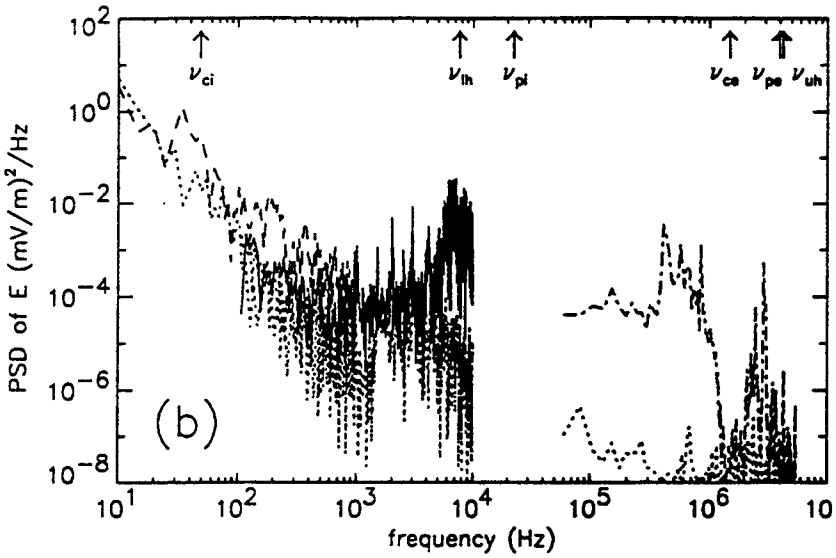


Figure 7: Emitted spectra of ECHO 7 experiment

that different mechanisms operate in the space experiments and in the laboratory experiments.

4.2 Comparison of Space Experiments with VTF Experiments

First, high-frequency modes with frequencies exceeding the electron gyrofrequency (f_{ce})–were not excited in our VTF plasmas, while they were present in the beam injection experiments in space. This difference arises from the fact that 36 keV electron beams were injected in space experiments, whereas 300 eV electron beams were produced in the VTF plasmas. Note that the thermal electron energy in the ionosphere and in the VTF are 0.1 eV and 5-8 eV, respectively. In fact, the 300 eV electron beams reduced their speeds to nearly the electron thermal speed at the location of the probe via quasilinear diffusion and the ionization of neutrals along the beam path to the top of the chamber. Thus, the low-energy electron beams are unable to excite Langmuir waves and upper hybrid waves in VTF. By contrast, whistler waves (viz. modes in the frequency range: $f_{lh} < f < f_{ce}$) were produced in both space experiments and laboratory experiments possibly by the convective beam amplification of incoherent Cerenkov noise [*Maggs*, 1976].

The other distinctive difference is the low-frequency modes whose frequencies are less than the lower-hybrid resonance (f_{lhr}). One can see from Figure 7 that the ionospheric plasmas had large thermal noises in the low-frequency band, $f < f_{lhr}$. The energetic electron beams injected into

the ionosphere cannot excite the low-frequency modes significantly, whereas prominent excitation of these modes occurs in the VTF plasmas.

The discrepancy in the low-frequency regime of the spectra stems from the different background plasma environments. In the space experiments. the sounding rocket carried the electron guns as its main payload and the electric field probes on a plasma diagnostic package (PDP). During the experiments, the PDP separated at 1.5 m/s relative to the main payload. Thus the electron beam-plasma interactions and their subsequent diagnoses were carried out in a rather small ionospheric region which can reasonably be considered a uniform plasma. By contrast, the VTF plasmas created by microwaves and/or electron beams have sharp density gradients across the magnetic field. In addition, the current convective instability requires ion motion such as found in a drifting Maxwellian plasma as opposed to the space experiments contain beams with very high energy.

5 Theoretical Analysis of Experiments

The excitation of whistler waves and lower-hybrid waves (Trivelpiece-Gould modes) can be explained by well known inverse Landau damping processes similar to beam interactions described by *Maggs*, [1976]. This section will focus on a lesser known process, the current convective instability.

The DNA Wideband satellite has detected scintillation phenomena caused by L shell aligned, field aligned, large-scale irregularities occurring in the high latitudes. [*Vickrey et al.*, 1980]. Associated with the irregularities are soft particle precipitation and horizontal plasma density gradients. The current convective instability has been suggested as a mechanism driving such irregularities [*Ossakow and Chaturvedi*, 1979]. The current convective instability occurs for field-aligned currents exceeding a threshold which depends on the local density gradient transverse to the magnetic field [*Kadomtsev and Nedospasov*, 1960]. A nonlinear saturation mechanism has been described [*Chaturvedi and Ossakow*, 1979], as well as effects of ion inertia, collisional ions, and electromagnetics [*Chaturvedi and Ossakow*, 1981]. Numerical simulations have been performed [*Keskinen et al.*, 1980] including those for large scale-lengths [*Keskinen and Ossakow*, 1982] and for small scale-lengths driven by large scalelength density gradients [*Keskinen and Ossakow*, 1983].

Presented in Figure 8 are the two geometries describing the current convective instability. In the auroral region the geomagnetic field, $\vec{B}_o$, is pointed nearly downward. The current and also

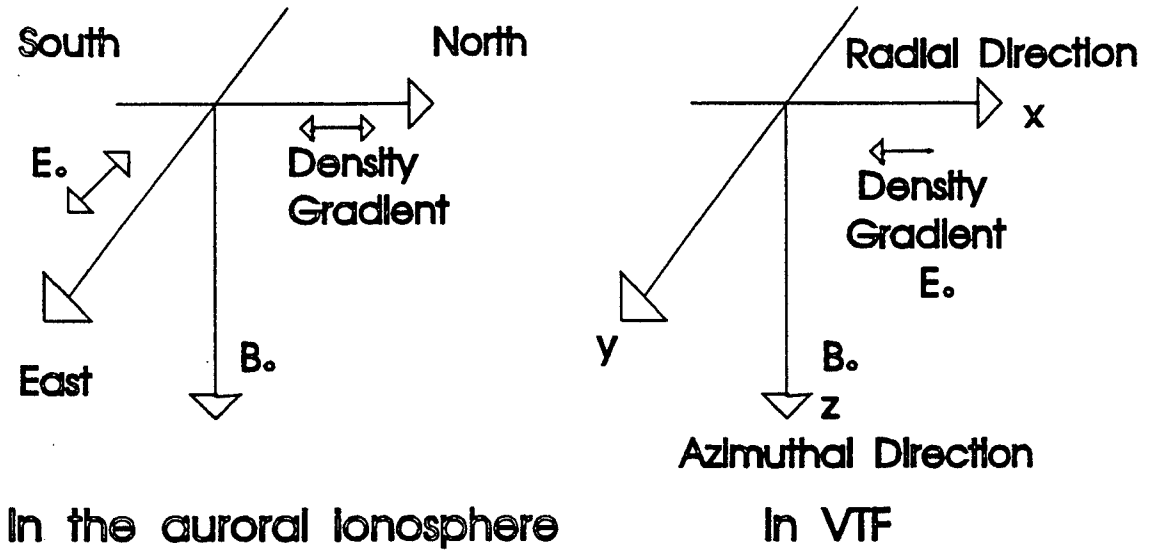


Figure 8: Geometry of the current convective instability in the ionosphere and in VTF.

the particle drift, $\vec{V}_{dr}$, is along the magnetic field. The plasma density gradient, ∇n, can be either northward or southward. Imposed on the plasma is an eastward or westward electric field, E_o. Depending on the sign of $(\vec{E}_o \times \vec{B}_o) \times \nabla n$ the usual E cross B instability may occur. According to *Chaturvedi and Ossakow*, [1979] the field aligned current may overcome the stabilizing E × B in one case or enhance it in the other case. The linear instability dispersion relation and growth rates are:

$$\omega_r = k_{\|} v_{dr\|} + \frac{\frac{k_{\|}^2}{k_x}\frac{\Omega_e}{\nu_e}\left(\frac{\nu_i}{\Omega_e}\frac{E_o}{B_o} + \frac{k_{\|}}{k_x}V_{dr}\right)}{\left(\frac{\Omega_i}{\nu_i} + \frac{\Omega_e}{\nu_e}\right)\frac{k_{\|}^2}{k_x^2} + \frac{\nu_i}{\Omega_i}}$$

$$\gamma = \frac{-\frac{1}{n_o}\frac{\partial n_o}{\partial y}\left(\frac{E_o}{B_o}\frac{\nu_i}{\Omega_i} + V_{dr}\frac{k_{\|}}{k_x}\right)}{\left(\frac{\omega_i}{\nu_i} + \frac{\Omega_e}{\nu_e}\right)\frac{k_{\|}^2}{k_x^2} + \frac{\nu_i}{\Omega_i}}$$

Figure 8 also shows the geometry of VTF which is similar to the geometry of the auroral region. The toroidal field, $\vec{B}_o$, defines the parallel direction along which the plasma current flows, $\vec{V}_{dr}$. Perpendicular to the magnetic field is the density gradient, ∇n_o. The primary difference between these two geometries is the direction of the externally applied DC electric field, $\vec{E}_o$. Unlike the auroral region, the VTF conditions have this electric field parallel to the density gradient. In VTF, this electric field will be written as $\vec{E}_{\perp}$ to distinguish it from the electric field, $\vec{E}_{\|}$, which drives the current along the magnetic field. The derivation of the current convective instability for this modified geometry is as follows.

Taking the electron momentum equation, ignoring the terms describing inertial and temperature

effects, crossing with $\vec{B}_o$ and substituting into itself, one gets

$$\vec{E} \times \vec{B}_o + B_o^2 \vec{V}_{e\|} - B_o^2 \vec{V}_e = \frac{m}{e}\nu_e \vec{E} + (\frac{m}{e}\nu_e)^2 \vec{V}_e$$

where for electrons $\nu_e = \nu_{ei} + \nu_{en} \simeq \nu_{ei}$. This can be rewritten as

$$\frac{\vec{E} \times \hat{z}}{B_o} - \vec{V}_{e\perp} = \frac{\nu_e}{B_o \Omega_e}\vec{E} + \frac{\nu_e^2}{\Omega_e^2}\vec{V}_e$$

since the imposed magnetic field is in the $\hat{z}$ direction. Neglecting the Pederson drift, considering the ratio ν_e/Ω_e for the plasmas in VTF and the auroral ionosphere, this results in the expression for the electron drift in the presence of perpendicular and parallel electric fields.

$$\vec{V}_e = \frac{\vec{E} \times \hat{z}}{B_o} - \frac{\Omega_e}{\nu_e B_o}\vec{E}_\|$$

The equation for the ion drift follows a similar derivation. Once again ignoring inertial and temperature effects. By retaining the Pederson drift since in VTF and Auroral plasmas the ratio ν_i/Ω_i is about three orders of magnitude larger than ν_e/Ω_e one finds

$$\vec{V}_i = \frac{\vec{E} \times \hat{z}}{B_o} + \frac{\nu_i}{B_o \Omega_i}\vec{E}_\perp + \frac{\Omega_i}{\nu_i B_o}\vec{E}_\| + \vec{V}_{io\|}$$

For ions, $\nu_i = \nu_{ie} + \nu_{in} \simeq \nu_{in}$.

For $n = n_o(y) + \tilde{n}$, $\vec{E} = E_\|\hat{z} + E_\perp \hat{y} - \nabla\tilde{\phi}$, $\vec{V}_{e,i} = \vec{V}_{e,io} + \tilde{\vec{V}}_{e,i}$, and using the usual Fourier-Laplace perturbation analysis with wavelike dependence, $\exp[j(k_x x + k_\| z - \omega t)]$, where $\omega = \omega_r + j\gamma$. The above equations can be rewritten,

$$\vec{V}_{io} = \frac{\Omega_i}{\nu_i B_o}\vec{E}_\| + \frac{\nu_i E_\perp}{\Omega_i B_o}\hat{y} + \frac{E_\perp}{B_o}\hat{x} = \vec{V}_{io\|} + \frac{\nu_i E_\perp}{\Omega_i B_o}\hat{y} + \frac{E_\perp}{B_o}\hat{x}$$

$$\vec{V}_{eo} = \frac{\Omega_e}{\nu_e B_o}\vec{E}_\| + \frac{E_\perp}{B_o}\hat{x} = \vec{V}_{eo\|} + \frac{E_\perp}{B_o}\hat{x}$$

for the quiescent velocities, and

$$\tilde{\vec{V}}_i = -\frac{e}{M\nu_i} jk_\| \tilde{\phi}\hat{z} - \frac{\nu_i}{B_o \Omega_i} jk_x \tilde{\phi}\hat{x} + \frac{jk_x}{B_o}\tilde{\phi}\hat{y}$$

$$\tilde{\vec{V}}_e = +\frac{e}{m\nu_e} jk_\| \tilde{\phi}\hat{z} + \frac{jk_x}{B_o}\tilde{\phi}\hat{y}$$

for the fluctuating velocities.

The linearized electron continuity equation can be written as:

$$\frac{\partial \tilde{n}}{\partial t} + \vec{V}_{eo} \cdot \nabla\tilde{n} + n_o \nabla \cdot \tilde{\vec{V}}_e + \tilde{\vec{V}}_e \cdot \nabla n_o = 0$$

Therefore

$$-j\omega\tilde{n} + (\vec{V}_{eo\|} + \frac{E_\perp}{B_o}\hat{x}) \cdot \nabla\tilde{n} - n_o\frac{ek_\|^2}{m\nu_e}\tilde{\phi} + \frac{jk_x}{B_o}\tilde{\phi}\frac{\partial n_o}{\partial y} = 0$$

or

$$j(\omega - k_\| V_{eo\|} - k_x\frac{E_\perp}{B_o})\tilde{n} - \left(\frac{jk_x}{B_o}\frac{\partial n_o}{\partial y} - \frac{ek_\|^2}{m\nu_e}n_o\right)\tilde{\phi} = 0$$

The next step is to linearize $\nabla \cdot \vec{J} = 0$, where $\vec{J} = \sum_\alpha n_\alpha q_\alpha \vec{V}_\alpha$.

$$\nabla \cdot \left\{\tilde{n}(\vec{V}_{io} - \vec{V}_{eo}) + n_o(\tilde{\vec{V}}_i - \tilde{\vec{V}}_e)\right\} = 0$$

$$\nabla \cdot \left\{\tilde{n}(\vec{V}_d + \frac{\nu_i}{\Omega_i}\frac{E_\perp}{B_o}\hat{y}) - n_o\left[ejk_\|\tilde{\phi}\hat{z}(\frac{1}{M\nu_i} + \frac{1}{m\nu_e}) + \frac{\nu_i jk_x}{B_o\Omega_i}\tilde{\phi}\hat{x}\right]\right\} = 0$$

then

$$jk_\| V_d\tilde{n} + \tilde{\phi}n_o\left[ek_\|^2(\frac{1}{M\nu_i} + \frac{1}{m\nu_e}) + \frac{\nu_i k_x^2}{B_o\Omega_i}\right] = 0$$

For nontrivial $\tilde{n}$ and $\tilde{\phi}$:

$$j(\omega - k_\| V_{eo\|} - k_x\frac{E_\perp}{B_o})n_o\left[ek_\|^2(\frac{1}{M\nu_i} + \frac{1}{m\nu_e}) + \frac{\nu_i k_x^2}{B_o\Omega_i}\right] + jk_\| V_d(j\frac{k_x}{B_o}\frac{\partial n_o}{\partial y} - \frac{ek_\|^2}{m\nu_e}n_o) = 0$$

therefore

$$\omega' = \frac{-j\frac{1}{n_o}\frac{\partial n_o}{\partial y}V_d\frac{k_\|}{k_x} + \frac{k_\|^2\Omega_e}{k_x\nu_e}\frac{k_\|}{k_x}V_d}{(\frac{\Omega_i}{\nu_i} + \frac{\Omega_e}{\nu_e})\frac{k_\|^2}{k_x^2} + \frac{\nu_i}{\Omega_i}}$$

where $\omega' = \omega - k_\| V_{eo\|} - k_\perp\frac{E_\perp}{B_o}$, and $\omega = \omega_r + j\gamma$. Finally, the dispersion relation:

$$\gamma = \frac{-\frac{1}{n_o}\frac{\partial n_o}{\partial y}V_d\frac{k_\|}{k_x}}{(\frac{\Omega_i}{\nu_i} + \frac{\Omega_e}{\nu_e})\frac{k_\|^2}{k_x^2} + \frac{\nu_i}{\Omega_i}}$$

and

$$\omega_r = k_\| V_{eo\|} + k_\perp\frac{E_\perp}{B_o} + \frac{\frac{k_\|^2}{k_x}\frac{\Omega_e}{\nu_e}\frac{k_\|}{k_x}V_d)}{(\frac{\Omega_i}{\nu_i} + \frac{\Omega_e}{\nu_e})\frac{k_\|^2}{k_x^2} + \frac{\nu_i}{\Omega_i}}$$

Using the following VTF parameters: $E_\perp$=100 V/m, $f_{ce} = 2.45$ GHz, $f_{ci} = 1.3$ MHz, $\nu_n = 10^3$, $\nu_e = 10^6$, $V_d = 5 \times 10^5$ m/sec, L $= n(\delta n)^{-1} = 0.1$ m, and $k_\|/k_\perp = 10^{-2}$ (for maximum growth rate), the growth time of the convective instability in VTF is 10 μsec. It indicates fast excitation of the current convective instability in the VTF plasmas.

Our modified current convective instability shows that for fixed wave number, (e.g. as measured by incoherent scatter radar or with a multi-point correlation), the real frequencies of excited current convective modes increase with the electron drift velocity and applied electric field. This feature has been qualitatively confirmed by spectral analysis and correlation measurements. The expression for

the growth rate contains a density gradient factor in addition to particle drift factor. As mentioned before, in the space experiments, the beam-plasma interactions occurred in a rather small region of the ionosphere. Hence, the beam-plasma interaction region in space experiments can be modelled as a uniform background plasma and, consequently, the current convective instability does not occur.

A saturation mechanism for various gradient driven instabilities has been applied to the current convective instability [*Chaturvedi and Ossakow*, 1979]. A two-dimensional nonlinearity in the continuity equation predicts the generation of linearly damped harmonics. The harmonics, which propagate anti-parallel to the density gradient, reach saturation amplitudes greater than the primary modes. Hence in the auroral ionosphere with northward density gradients and field aligned currents, the dominant modes of the fluctuation spectrum will be north-south. Applying this mechanism to the VTF beam plasmas predicts modes propagating radially outward. This characteristic has not yet been confirmed in our current laboratory experiments.

6 Discussions and Conclusions

VTF was constructed to simulate the ionospheric plasma environment. The excitation of high-frequency modes (i.e. Langmuir waves and upper hybrid waves) depends on the energy of the electron beams, as discussed before.

Comparison of VTF experiments with space experiments have shown that VTF can simulate several features of the electron beam injection experiments in space plasmas. However, it can more closely simulate the naturally occurring plasma turbulence in the auroral ionosphere. The latter was one of the original goals in constructing VTF.

The characteristics of the excited waves by injected electron beams in the ionosphere and in VTF can be summarized as follows. Langmuir waves, upper hybrid waves, whistlers, and lower hybrid waves can be excited by energetic electron beams in the ionospheric plasmas. In the VTF plasmas, the injected electron beams produce whistler waves and especially low-frequency modes with frequencies less than the lower hybrid resonance frequency (f_{lh}) and even the ion gyrofrequency (f_{ci}).

The accumulated research results reported here have demonstrated that the VTF plasma device is capable of simulating some characteristic features of ionospheric plasma turbulence. Although

the parameters of the VTF plasmas are quite different from those of ionospheric plasmas, this work has shown that laboratory experiments can adequately study some processes responsible for space plasma phenomena and thus complement the active experiments in space.

References

[1] Chaturvedi, P.K., and S.L. Ossakow, Nonlinear Stabilization of the Current Convective Instability in the Diffuse Aurora, *GRL*, 6, 12, 1979.

[2] Chaturvedi, P.K., and S.L. Ossakow, The Current Convective Instability as Applied to the Auroral Ionosphere, *JGR*, 86, A6, 1981.

[3] Ginet, G.P., and J. Ernstmeyer, VLF plasma waves generated by an electron beam in space, *Physics of Space Plasmas*, SPI Conference Proceedings and Reprint Series, Cambridge, Massachusetts, Number 11, 209, 1991.

[4] Kadomtsev, B.B. and A.V. Nedospasov, Instability of the Positive Column in a Magnetic Field and the "Anomalous Diffusion Effect," *J. Nucl. Energy, Part C*, 1, 230, 1960.

[5] Kellogg, P.J., S.J. Monson, R.H. Holzworth, and R.J. Jost, Beam-Generated Waves in a Large Plasma Chamber, *IEEE Transactions on Plasma Science*, PS-14(6), 891, 1986b.

[6] Keskinen, M.J., S.L. Ossakow, and B.E. McDonald, Nonlinear Evolution of Diffuse Auroral F Region Ionospheric Irregularities, *Geophysical Research Letters*, 7, 8, 1980.

[7] Keskinen, M.J., and S.L. Ossakow, Nonlinear Evolution of Plasma Enhancements in the Auroral Ionosphere 1. Long Wavelength Irregularities, *JGR*, 87, A1, 1982.

[8] Keskinen, M.J., and S.L. Ossakow, Nonlinear Evolution of Convecting Plasma Enhancements in the Auroral Ionosphere, 2. Small Scale Irregularities, *JGR*, 88, A1, 1983.

[9] Kindel, J.M., and C.F. Kennel, Topside current instabilities, *J. Geophys. Res.*, 76(13), 3055, 1971.

[10] Maggs, J.E., Coherent Generation of VLF Hiss, *J. Geophys. Res.*, 81(10), 1707, 1976.

[11] Moriarty, D.T., *Electron Cyclotron Range of Frequencies Propagation in Critically Dense Cold Magnetoplasmas*, MIT Nuclear Engineering Department Master's Thesis, 1992.

[12] Ossakow, S.L., and P.K. Chaturvedi, Current Convective Instability in the Diffuse Aurora, *GRL*, 6, 4, 1979.

[13] Riddolls, R.J., M.C. Lee, D.T. Moriarty, S.M. Murphy, D. Pooley, Laboratory study of anti-Stokes Langmuir waves in a magnetized plasma, *Proceedings, 22nd IEEE ICOPS, Madison, WI*, 1995.

[14] Roth, J.R., and W.M. Krawczonek, Paired Comparison Tests of the Relative Signal Detected by Capacitive and Floating Langmuir Probes in Turbulent Plasma from 0.2 to 10 MHz, *The Review of Scientific Instruments*. 42(5), 1971.

[15] Vickrey, J.F., C.L. Rino, and T.A. Potemra, Chatanika/Triad observations of unstable ionization enhancements in the auroral F region, *Geophys. Res. Lett.*, 7, 789, 1980.

[16] Winglee, R.M., and P.J. Kellogg, Electron beam injection during active experiments: 1. Electromagnetic wave emissions, *J. Geophys. Res.*, 95, 6167, 1990.

[17] Winckler, J.R., J.R. Steffen, P.R. Malcolm, K.N. Erickson, Y. Abe, and R.L. Swanson, Ion resonances and ELF wave production by an electron beam injected into the ionosphere: ECHO 6, *J. Geophys. Res.*, 89, 7565, 1984.

[18] Winckler, J.R., R.R. Malcom, R.L. Arnoldy, W.J. Burke, K.N. Erickson, J. Ernstmeyer, R.C. Franz, T.J. Hallinan, P.J. Kellogg, S.J. Monson, K.A. Lynch, G. Murphy, and R.J. Nemzek, ECHO 7: An electron beam experiment in the magnetosphere, *EOS Trans. AGU*, 70, 657, 1989.

ELECTRON BEAM-PLASMA INTERACTIONS WITH THE VERSATILE TOROIDAL FACILITY (VTF) FOR IONOSPHERIC SIMULATION

Suzanne Marie Murphy
Graduate Student
Electrical Engineering

Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

Final Report for
Graduate Student Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Hanscom Air Force Base, Lexington, MA
and
Phillips Laboratory

September, 1995

ELECTRON BEAM-PLASMA INTERACTIONS WITH THE VERSATILE TOROIDAL FACILITY (VTF) FOR IONOSPHERIC SIMULATION

Suzanne Marie Murphy
Graduate Student
Department of Electrical Engineering
Massachusetts Institute of Technology

Abstract

An investigation of electron beam-plasma interactions in the laboratory is conducted to cross-check the 1992 TSS-1 space shuttle experiments. Because active experiments in space plasmas are often expensive and difficult to conduct on a frequent basis, there arises a need for laboratory simulations. The Versatile Toroidal Facility (VTF) at the MIT Plasma Fusion Center has been used to generate a large magnetized plasma with overdense conditions as a method of simulating shuttle and ionospheric heating experiments. The capabilities of the Taylor discharge method as well as electron beams emitted from LaB_6 filaments are explored to provide an overdense background plasma. Electron beams are also used to represent to the electron guns in the shuttle experiments. A vertical langmuir probe is designed and built to characterize the plasma behavior in close proximity to the electron beam. A radial side probe is used to measure data at mid-plane for comparison. Limitations on the current studies are described and future work for further investigation is discussed.

Contents

ELECTRON BEAM-PLASMA INTERACTIONS WITH THE VERSATILE TOROIDAL FACILITY (VTF) FOR IONOSPHERIC SIMULATION

Suzanne Marie Murphy

1 Introduction

The Versatile Toroidal Facility (VTF) at the MIT Plasma Fusion Center has been used for several years to study wave propagation and turbulence in plasmas. Because of the many ionospheric heating and shuttle experiments there has grown a need for laboratory simulations of ionospheric conditions. VTF is capable of filling that need, specifically by simulating electron beam experiments that were conducted with the space shuttle Tethered Satellite System (TSS) in 1992. The Taylor discharge method is explored in conjunction with LaB_6 filaments to produce a steady state plasma, analogous to the background plasma in the ionosphere. An electron beam is pulsed to simulate the beam injected in the space plasmas. The results of these experiments will be compared with the data obtained in the TSS experiments to determine how well plasma turbulence generated inside the VTF can be used to study ionospheric plasmas.

2 Space and Laboratory Experiments

2.1 Space Shuttle Experiments

The Tethered Satellite System (TSS) was launched aboard the Space Shuttle STS-46 Atlantis which flew from July 31 to August 6, 1992, as a joint effort between the United States and Italy. The paramaters and equipment used in the shuttle experiments are described by David Hardy (1995). The parameters Of the many experiments that were carried out on this mission, this project focuses on the Shuttle Potential and Return Electron Experiment (SPREE), which utilizes electrostatic analyzers in the shuttle payload bay to make measurements of particle energy and power spectra.

The shuttle flew at an altitude of 300 km in a 28.5° inclination orbit. SPREE consisted of two Fast Pulsed Electron Generators (FPEG) which injected a 100 mA, 1 keV electron beam at angles nearly perpendicular to the earth's magnetic field and with turn on and turn off times of approximately 100 ns. The beam electrons were emitted through an aperture of 0.56 cm radius, with a magnetically controlled divergence of 7.5°. Two electrostatic analyzers (ESA) were also located on the opposite side of the shuttle payload. These analyzers simultaneously measured ion and electron fluxes with energies from 10 eV to 10 keV and within a 100° x 8° angular fan. This angular region was divided into 10 zones of 10° width. Ion and electron spectra were measured independently in each zone. The two analyzers, labeled A and B, were mounted on rotary tables and placed back to back on the shuttle payload. The motion of these analyzers were synchronized so that both ion and electron spectra were sampled over a solid angle of 2π steradians every 30 s. To prevent saturated measurements, the geometric factors of ESA B were set to be around 100 times smaller than those for ESA A. The energy resolution for each of the electron channels is $\Delta E/E \approx 7\%$. The arrival times of individual electrons were also measured to detect particle bunching and hence, signify resonant wave-particle interactions [Dobrowolny, 1994].

Many sources of error in the SPREE measurements had to be taken into account. The current collection may have been affected by induced localized discharges due to Orbiter outgassing, thruster actions, water dumps and the presence of neutral gas in the vicinity of the shuttle. In addition, the size of the sheath that formed around the shuttle may also have had an effect on the particle measurements. The experimental setup is more complex than the conditions laid by Langmuir for a sphere at high voltage inside an unmagnetized plasma. The TSS setup is more complex because the ionosphere is magnetized and the shuttle is moving in a meso-sonic plasma since its velocity is supersonic with respect to ions and subsonic with respect to electrons [Dobrowolny, 1994].

The TSS experiments were done in the F-region of the ionosphere which extends from 100 to 350 km. This region in altitude contains overdense plasma, which exists when the electron plasma frequency, ω_{pe}, and the

electron cyclotron frequency, ω_{ce}, obey the relation

$$\omega_{pe} > \omega_{ce} \tag{1}$$

At the shuttle's altitude of 300 km, these frequencies lie in the following ranges: $f_{ce} = 0.7$ to 1.4 MHz and $f_{pe} = 2.8$ to 8.9 MHz. For the shuttle experiments, the electron cyclotron frequency was taken to be $f_{ce} = 0.8$ MHz [Hardy, et. al., 1995]. The ratio of these frequencies thus lies in the range $3.5 < \frac{f_{pe}}{f_{ce}} < 11.1$.

Table 1: Table of Parameters for Shuttle and VTF.

	VTF	Space Shuttle
Background Density (cm^{-3})	0	10^4
Beam Density (cm^{-3})	10^{12}	1 to $5{\cdot}10^4$
Acceleration Potential (V)	360	1000
Neutral Pressure (T)	$9.8 \cdot 10^{-5}$	10^{-6}
Magnetic Field (G)	1200-1800	0.3

2.2 VTF Experimental Setup and Diagnostics

The VTF is set up with a magnetic field inside the chamber produced by 18 toroidal magnets with 4 current coils running through each for a total of 72 toroidal field coils. There are 4 vertical field coils, two running outside the top of the chamber and two outside the bottom. The current through these coils produces an upward directed magnetic field inside the chamber. For a current, I_{TF}, going through the toroidal coils, we can use Amperes law to determine the resulting toroidal magnetic field:

$$\oint B \cdot ds = \mu_o \int J \cdot da \tag{2}$$

$$B = \frac{\mu_o 72 I_{TF}}{2\pi r} \tag{3}$$

The vertical field is described by the following equation:

$$B_z = \frac{\mu_o N I a^2}{2(a^2 + z^2)^{\frac{3}{2}}} \tag{4}$$

where N is 4, a is sthe radius of the coils, and z is the height above the plane of the loop, at the center.

The experiments conducted in this project will utilize one of the four available filaments as well as the Taylor discharge unit which is described in the following section. Figure 1 shows the position of the vertical probe and the four filaments in the chamber. The filaments, are labeled: South Top, North Middle, South Bottom, and North Top. The names refer to the power supplies being used for each filament. The terms *filament voltage* and *filament current* refer to the 11 V that is applied to heat up the filament, and the resulting current, respectively. The term *arc voltage* refers to the negative bias voltage that is applied across the filament to cause electron emission. *arc current* refers to the current that is emitted after the arc voltage has been applied.

Both Hydrogen and Argon were available to use as the neutral fill gas in VTF. However, the Argon plasmas tended to be more dense and consistent than the Hydrogen plasmas.

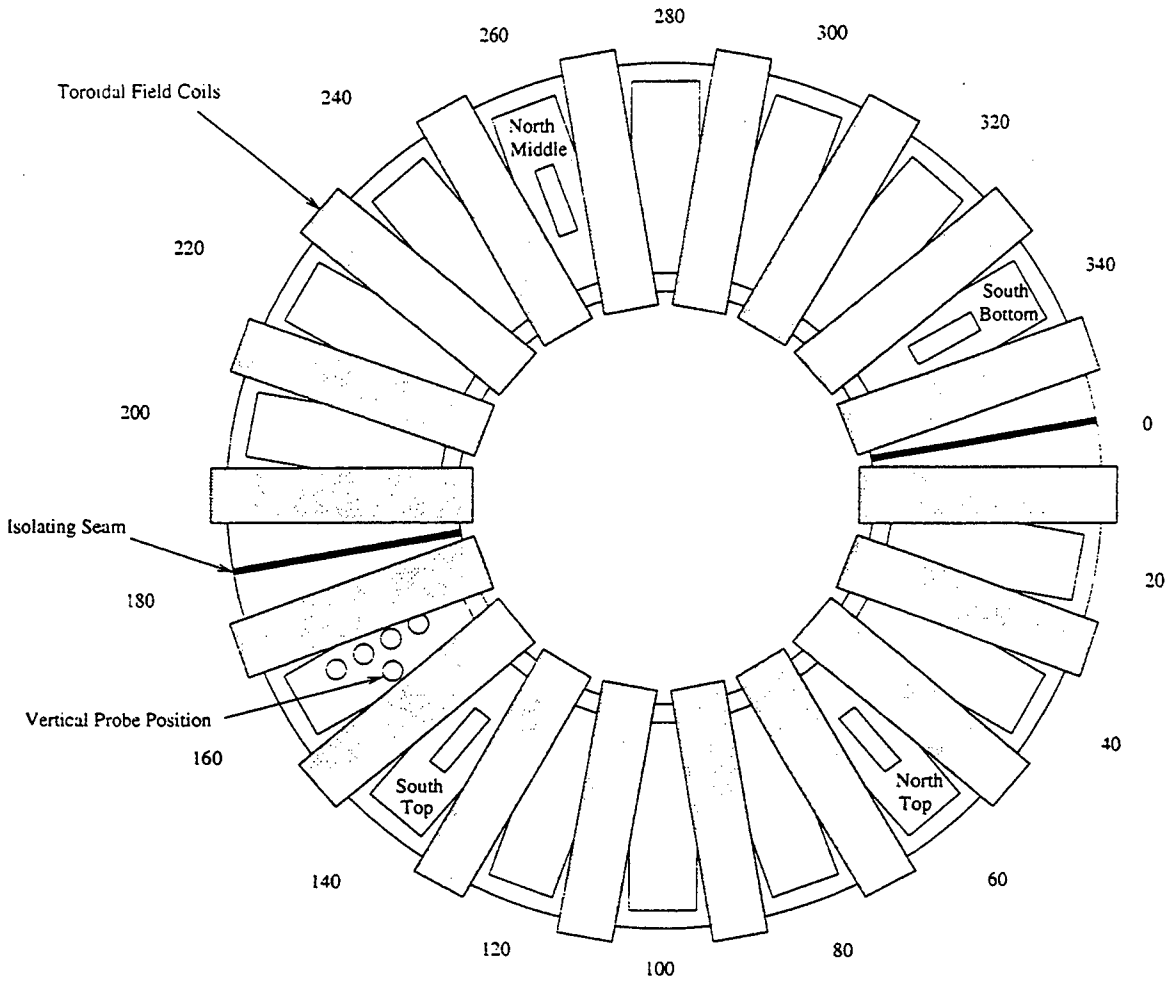


Figure 1: Map of the VTF Chamber Floor Including Positions of Vertical Probe and the Four Filaments: South Top, North Middle, South Bottom, and North Top. (Drawing Not to Scale)

2.2.1 Langmuir Probe Operation

Two Langmuir probes were used to make measurements in the VTF. The first is a side probe which sweeps across the radius of the chamber. The tip of this probe consists of a flat, square piece of copper, 0.63 cm on each side, for a cross-sectional area of 0.40 cm^2. The tip is oriented such that the flat surface is normal to the beam direction. The other probe is a vertical probe, placed at the bottom of the chamber. The design of this probe is further explained in section 2.2.2.

The first experiments performed with the radial probe consisted of density measurements. The Langmuir probe was swept across the radius from the inner wall to the outer wall. The probe was negatively biased to -100 V using a power supply. In this way the ion saturation current is measured, which can be used to calculate plasma density in the vicinity of the probe. The I-V characteristics of a Langmuir probe are described in detail by Hutchinson (1987). He derives the current measured by a Langmuir probe to be:

$$I = n_\infty e A_p \sqrt{\frac{T_e}{m_i}} \left[\frac{1}{2} \sqrt{\frac{2m_i}{\pi m_e}} exp\left(\frac{e(V_o - V_f)}{T_e} \right) - \frac{A_s}{A_p} exp\left(-\frac{1}{2} \right) \right] \tag{5}$$

where n_∞ is the density at steady-state, V_o is the probe potential, A_s is the area of the sheath, and A_p is the area of the probe. When the probe is biased to a voltage much lower than the floating potential, V_f, the electron current is negligible, and the ion saturation current density is found to be:

$$J_i = n_\infty e \, exp\left(-\frac{1}{2} \right) \sqrt{\frac{T_e}{m_i}} \tag{6}$$

It should be recognized that the equation for ion saturation current is derived for the case of an unmagnetized plasma, however most of the current being measured is traveling parallel to the magnetic field so this formula will be adequate. We then find the current by multiplying by the projected area of the Langmuir probe tip, A_p, seen by this current.

$$I_s = J_i A_p \tag{7}$$

If we know the electron temperature, in eV, then the current measurement can be used to find n_∞, the steady-state plasma density:

$$n_\infty = \frac{I_s}{e A_p} exp\left(\frac{1}{2} \right) \sqrt{\frac{M_i}{T_e}} \tag{8}$$

As mentioned earlier, the area of the side probe is A_p =0.4 cm^2. The area of the vertical probe tip seen by an approaching beam is 0.63 cm by 0.0625 cm which is equal to $3.9 \cdot 10^{-2}$ cm^2.

2.2.2 Design and Operation of a Vertical Probe

A vertical Langmuir probe was designed, built and installed on the bottom of the chamber. This probe was positioned as close to the electron beam as possible to characterize the plasma right after the electrons

are emitted from the filament. The VTF LaB_6 filaments emit electrons which are then accelerated by the applied 360 V which produces an electric field in the chamber. The velocity of electrons coming off the filament is initially zero and then increases until they hit the collector plate at the top of the machine.

3 Simulation of Ionospheric Conditions

3.1 Taylor Discharge Method

It has been found that a resonant LC circuit can be used to generate a plasma. This scheme for producing inductive-drive plasmas is known as Taylor discharge, developed by Robert Taylor, now at UCLA. With this method a plasma can be formed in a toroidal machine without using high voltage power supplies. The scheme involves hooking up the inner transformer coil of the chamber in series with a large capacitor and then driving the resulting LC circuit near resonance. The large reactive AC current in the inner transformer coil induces current in the plasma, which transfers energy via ohmic heating.

Finding the exact resonant frequency has proven to be difficult. Before any plasma has formed, the only load in this two-transformer model is the inductance of the secondary coil. When the plasma forms, the impedance of the load cannot be expressed as a linear expression. In general, the resistive part of the plasma load is inversely proportional to density, which causes the resistance to decrease as the plasma is formed. We also know that the density is proportional to the power that is transferred to the plasma. By substituting into the expression relating coil voltage and current, $V = IR$, we find that the voltage across the coils is constant, which makes the modeling of this system difficult.

The Taylor discharge circuit in the VTF, constructed by Ryan Riddolls, uses one power supply that can provide up to 40 kW of DC current at 360 V [Beals, 1994]. In order to use the power supply to drive the LC circuit near resonance, the DC power from the supply must be chopped to AC power near the LC resonant frequency and coupled to the resonant circuit to achieve a high coil voltage.

The method used to achieve this DC to AC conversion is somewhat similar to the series resonant converter. The 45-turn inner transformer coil is tapped at its first and ninth turn. A full-bridge inverter is then attached across these taps to create a square voltage waveform from the DC power. This square wave then couples into the remaining 37 turns of the transformer coil through the air core. The higher harmonics of the square wave are attenuated by the narrow-peaked transfer function of the LC circuit. When the amplitude of the sinusoidal voltage waveform across the full 45 turns of the inner transformer coil is high enough the loop voltage amplitude around the chamber reaches the breakdown voltage of the fill gas in the vacuum chamber and the plasma is formed.

We should note that when the Taylor discharge device is being run at $\approx$ 1 kHz, the millisecond confinement time of the VTF machine ensures that the plasma is extinguished during the part of the AC cycle when the loop voltage is below the breakdown voltage. If one were to change the resonant frequency, for example, by decreasing the size of the capacitor in the LC circuit, then the plasma would become continuous as the time for one period of the AC cycle became much less than the confinement time of the VTF plasma.

One major concern with respect to toroidal-shaped plasma confinement machines is that the driving of large currents through the plasma might create a distortion of the toroidal magnetic field geometry. The kink safety criterion is a rough assessment of this type of distortion. The criterion defines a quantity $q(r)$ and specifies the minimum value of this term such that a kink instability will not occur in a helical plasma:

$$q(r) = \frac{rB_\phi}{RB_\theta} \simeq 1 \tag{9}$$

where B_ϕ is the toroidal field, B_θ is the poloidal field created by the current in the plasma, R is the major radius of the toroidal plasma, and r is the radius of the assumed circular cross section of the plasma.

We recall Ampere's law for a toroidal field geometry:

$$B_\theta = \frac{\mu_0 \int J \cdot da}{2\pi r}.$$

Assuming uniform current density over the circular area of integration, B_θ reduces to

$$B_\theta = \frac{\mu_0 J r}{2} \tag{10}$$

and $q(r)$ becomes a constant with respect to r:

$$q_0 = \frac{2B_\phi}{\mu_0 R J} \tag{11}$$

The major radius, R, is equal to 92 cm and the toroidal field is 0.08 T. J is then calculated to be $1.38 \cdot 10^5 \frac{A}{m^2}$.

Now we assume a 0.5 m^2 plasma cross-sectional area, and we will continue to assume that the current density is uniformly distributed as we did in the calculation of maximum current density above. Thus the maximum current flowing in the Taylor discharge plasma without a kink instability occurring is $I_{max} = 68.6$ kA.

3.2 Beam-Plasma Interaction in VTF

To obtain an overdense plasma in the VTF we will be implementing the Taylor Discharge method as described in the previous section. We can now determine the parameters in the Taylor Discharge unit necessary to simulate the plasma conditions during the TSS shuttle mission. Previously, it has been found that the toroidal field must be at least 800 G before the density becomes independent of the field strength [Beals, 1994]. Using the formula for the electron cyclotron frequency,

$$\omega_{ce} = \frac{qB}{m_e} \tag{12}$$

we obtain

$$f_{ce} = \frac{\omega_{ce}}{2\pi} = 2.24GHz \tag{13}$$

The plasma conditions during the TSS mission demonstrate that $\frac{\omega_{pe}}{\omega_{ce}} \approx 3$. To match this condition in the VTF we need:

$$f_{pe} = \frac{\omega_{pe}}{2\pi} = 6.71\,GHz$$

Using the formula for plasma frequency,

$$\omega_{pe} = \sqrt{\frac{n_o e^2}{\epsilon_o m_e}} \tag{14}$$

we calculate the minimum necessary density to be:

$$n_o = 5.57 \cdot 10^{11} cm^{-3}$$

Higher density plasmas can be obtained using LaB_6 filaments which generate electron beam plasmas through thermionic emissions. When a filament is negatively biased, it emits electrons that travel upward inside the vessel. As these electrons collide with neutral particles a plasma is formed along the helical magnetic field in the VTF. One important concept related to thermionic emission is the work function. There are surface restraints that prevent most of the free electrons in a metal from leaving. These are the electrostatic forces produced by the charges in the atoms of the metal. A free electron must have a minimum kinetic energy to free itself from these forces. The work function is the work per unit charge required to free an electron from the influence of the charges in the metal. The emission equation is defined as follows:

$$J = AT^2 e^{\frac{-b_o}{T}} \tag{15}$$

where

$$A = 120.4\,\frac{Amp}{cm^2 K^2}$$

$$b_o = 11600\phi_o\,K$$

and ϕ_o is the work function.

One important consideration is the Schottky Effect which occurs when the emitting surface is subjected to a positive potential gradient. The applied field effectively reduces the work function. We are then left with a constant gradient as well as the normal potential barrier. These two forces produce a new potential barrier which has a maximum at $d_c = \frac{1}{2}\sqrt{\frac{e}{4\pi\epsilon_o E}}$, the critical escape distance. After an electron exceeds d_c, the electrostatic forces are outward so the electron keeps moving and is accelerated toward the collector

plate. The work function is effectively reduced by $\sqrt{\frac{eE}{4\pi\epsilon_o}}$. Half of this reduction is due to the image force being overcome, whereas the other half is due to the potential at d_c being reduced by the quantity, $d_c E$. Factors that may inhibit emission include when the free metal has been exhausted, when oxygen is present with a pressure, $P > 10^{-4}$ T, and also when $P > 10^{-3}$ T [Gewartowski, 1968].

When the two FPEGs injected electrons into the ionosphere it resulted in high energy electrons superimposed on the background maxwellian velocity distribution. Similarly, the Taylor Discharge unit on the VTF will produce a drifting maxwellian velocity distribution in the laboratory plasma. The short-pulsed electron beam emits higher energy electrons which form high velocity tail electrons. Presently, the electron beam can run with pulses as short as 10 ms. The slow turn on time of the emitter limits how fast we can pulse this beam. However, this pulse length is within the range of pulses used in the TSS experiments, which use 25 ms, 3.3 ms, and 51 μs. Because the waves can only grow as quickly as the plasma can be formed, the plasma is essentially always in steady state. But as described by Kellogg, et. al. (1986), the transient waves excited by electron beam pulses only exist for a few milliseconds of the pulse. Thus, our pulsed electron beams in conjunction with a Taylor discharge plasma, will adequately simulate the conditions observed in the ionospheric experiments for the steady state part of the FPEG pulses.

Unlike the FPEGs, the LaB_6 filaments in the VTF are not true electron guns. Rather, through thermionic emission the electrons are excited to the outer surface of the filament. Then an electric field is applied such that the electrons are repelled and accelerated along the helical magnetic field to a collecting plate at the top of the chamber. The path of electrons through the chamber causes ionization of the ambient neutral atoms. Because the electrons have a higher average energy, the velocity distribution becomes a drifting maxwellian in the vicinity of the beam path. Many of the electrons travel the entire way to the top of the chamber, but some diffusion also occurs out of the main stream to the chamber walls.

Table 2: Table of Results From Density Measurements.

	Taylor Discharge	T.D. + 1 beam	T.D. + 2 beams
Peak Density (m^{-3})	$8.5 \cdot 10^{16}$	$3.25 \cdot 10^{17}$	$1.26 \cdot 10^{18}$
TF Current (A)	7300	7300	12000
f_{pe} (GHz)	2.62	5.13	10.1
f_{ce} (GHz)	3.68	3.68	4.56
Peak Radius (cm)	80	80	106
Ratio $\frac{f_{pe}}{f_{ce}}$	.7	1.39	2.2

4 Analysis of Results

4.1 Radial Probe Measurements

4.1.1 Plasma Density

The density experiments were conducted for combinations of different plasma sources, namely, Taylor discharge, and one or two electron beams. The density profiles obtained for a Taylor discharge plasma is shown in Figure 2. The radial profiles for Taylor discharge plasma in conjunction with one and two electron beams are shown in Figure 3 and Figure 4, respectively. Note that these density values do not take into account the multiple ionization described in Section 2.2. The oscillations in each profile are due to the 1200 Hz of the Taylor discharge circuit. An important observation is that with Taylor discharge, and two electron beams running we achieved a density of $1.26 \cdot 10^{12} cm^{-3}$. Using the formula for f_{pe} and f_{ce} we find that $\frac{f_{pe}}{f_{ce}} = 2.2$ which constitutes an overdense plasma. The resulting ratios and parameters for each set of conditions is shown in Table 2. It is important to note that the peak for the Taylor discharge plasma alone occurs very close to the inner wall.

In Figure 3 two peaks are easily distinguished, due to Taylor Discharge and the electron beam, respectively. However, in Figure 4 the density produced by two electron beams outshadows the Taylor discharge contribution. It is thus possible to produce an overdense plasma by only using electron beams. Once better coupling is achieved through a series supply and iron core, we expect to produce a much greater density with Taylor discharge, thus increasing the ratio of $\frac{f_{pe}}{f_{ce}}$.

4.1.2 Spectra Measurements

Frequency spectra were measured using the radial probe, while pulsing the North Middle electron beam. The frequency power was measured for six distinct frequencies: 1, 5, 10, 50, 100, and 500 MHz. Eight pulses of 10 ms duration were equally spaced across the 65.1 cm sweep of the probe. It was found that the arc voltage turns on $\approx$ 30 ms after the input signal is sent. Because it is not precisely a 30 ms delay, the spectra data was digitized for 20 ms, within which the arc voltage came on for 10 ms.

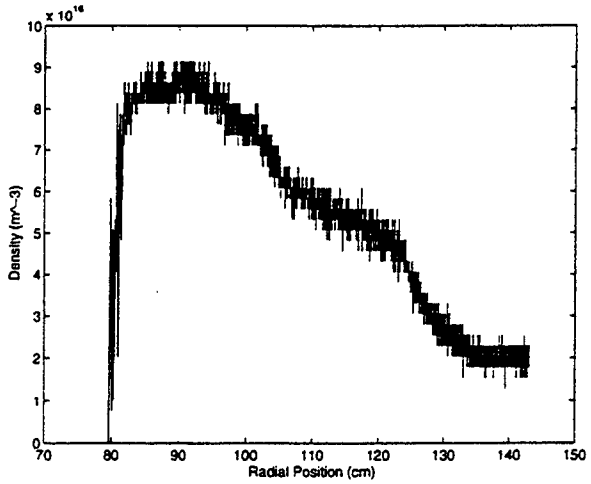


Figure 2: Radial Density Profile for a Taylor Discharge Plasma.

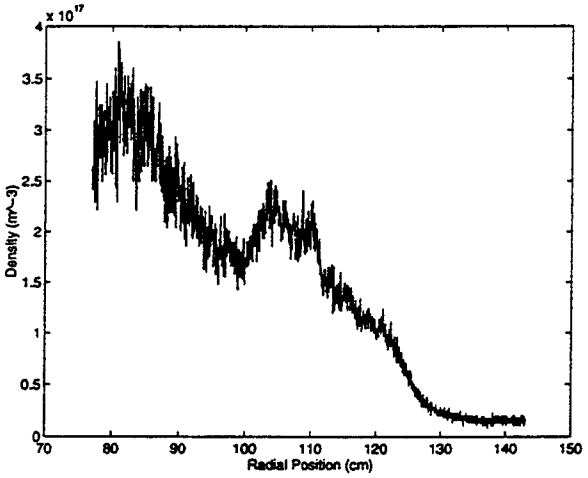


Figure 3: Radial Density Profile for Plasma Formed by Taylor Discharge and North Middle Filament.

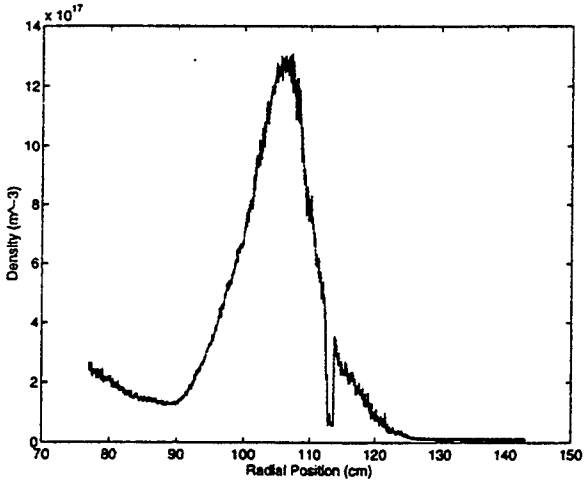


Figure 4: Radial Density Profile for Plasma Formed by Taylor Discharge and North Middle and South Top Filaments.

The maximum power level at 1 MHz is -49.8 dBm. The level is a little higher at 5 MHz with a maximum of -46.7 dBm, whereas at 10 MHz the maximum power goes down slightly to -48.4 dBm. The power then increases with maximum values of -45.69 dBm, -40.72 dBm, and -29.53 dBm, occurring at 50 MHz, 100 MHz, and 500 MHz, respectively. As shown in Figure 5 the peak for 500 MHz, as well as the other frequencies, occurs while the plasma is already in steady state.

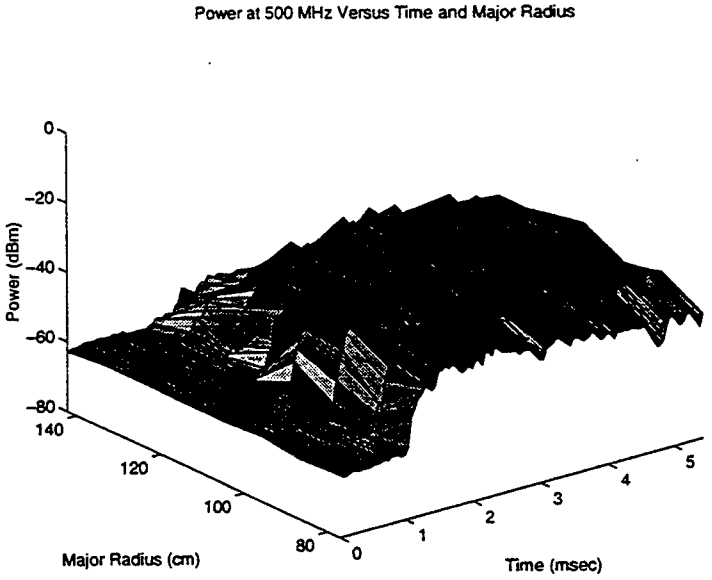


Figure 5: Power at 500 MHz Versus Time and Major Radius Measured by the Radial Probe

4.2 Vertical Probe Measurements

4.2.1 Plasma Density

Density measurements were done using the vertical probe. The probe has a fixed radial position but is able to be moved upward to 47.5 cm above the chamber floor. The plasma density at this location near the chamber floor was measured while running the South Top electron beam 27 cm away. The profiles measured by two different probe tips is shown in Figure 6. The last values measured by each tip correspond to the height of the electron beam.

4.2.2 Spectra Measurements

Using the vertical probe, the spectra was measured from 10 MHz to 10 GHz for Taylor discharge plasma, with and without electron beams, but no prominent peaks or enhancements were detected. Close to the beam, we hypothesize that waves have not had time to grow or cause instabilities. With a true electron gun we would expect to see wave activity around f_{ce} as well as lower hybrid instabilities which can occur with the existence of tail electrons in the energy distribution.

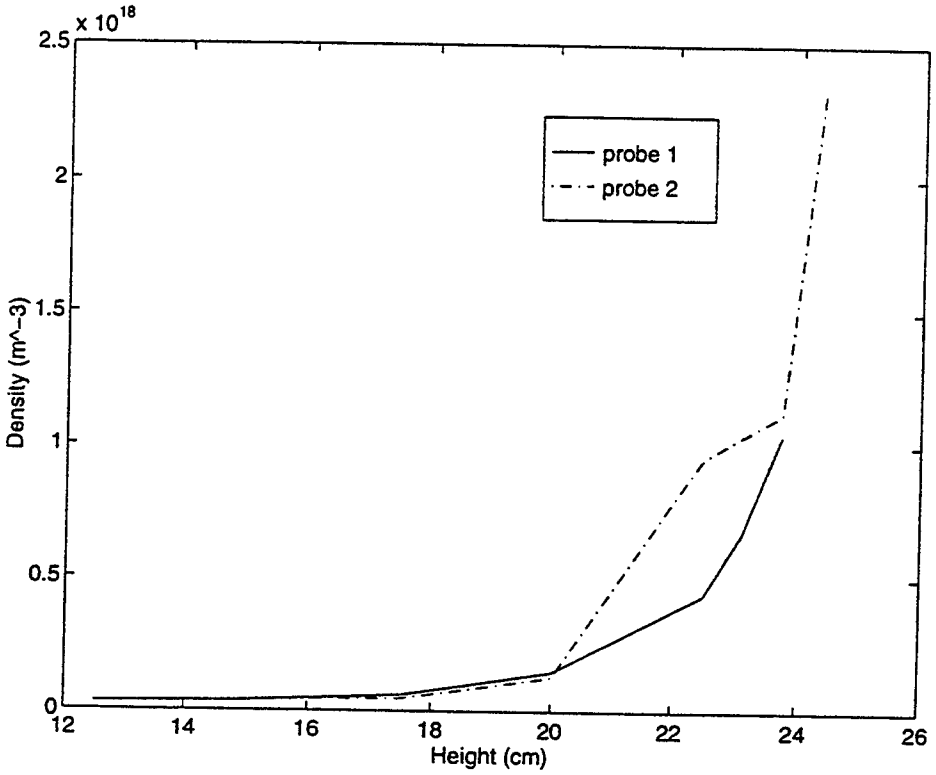


Figure 6: Plasma Density Measured by Vertical Probe With South Top Filament On

5 Conclusions

The results of this work help to characterize the plasma properties in the vicinity of the beam and how these properties differ at different locations within the chamber. The Taylor discharge unit in its current configuration was not able to produce an overdense plasma alone. Modifications to the circuitry should be explored and implemented to insure that more of the power goes into the plasma. The current setup and diagnostics of the VTF also pose certain constraints in determining how well the shuttle experiments can be simulated. The LaB_6 filaments are able to provide an overdense plasma, however without an electron gun it is more difficult to achieve an energy distribution with high energy tail electrons.

Pulsing of the beams is limited by the turn on time of the arc voltage as well as the delay of the arc current turning on. Aside from the initial 30 ms arc voltage switch delay, it takes $\approx$ 100 μs for the arc voltage to rise to its full value once the supply has been physically turned on. The arc current then turns on between 2 and 5 ms later. This delay makes it difficult to run shorter pulses because there might not be enough time for the arc current to come on and for the plasma to be formed.

The probe measurements are also limited in how much of the chamber is scannable. The existing side probe can make a radial scan along the midplane of the chamber. The vertical probe on the bottom which was recently installed can move up to 47.5 cm above the chamber floor, also at mid-radius.

Because of the constraints on producing an environment analogous to that produced by the FPEGs it would be beneficial to install an electron gun at the bottom of the chamber and to place the vertical probe close enough to detect the tail electrons in the velocity distribution. An overdense background plasma can be provided by the Taylor discharge method together with up to 3 other beams, or Taylor discharge by itself with an iron core added to replace the present air core.

An energy analyzer similar to the ESAs of the shuttle experiments, would also provide valuable measurements. Then the energy distribution could be measured close to the electron gun and also at the side probe. The effects of changing the beam current could then be studied and compared with the results of Vedenov and Ryutov (1975). Another valuable addition would be an array of mobile probes either vertical or radial. This type of array would allow us to obtain a two-dimensional density profile like the one obtained by Kellogg, et. al.(1986).

In summary, the VTF is suitable for simulating ionospheric plasma. Overdense plasma conditions have been achieved with $\frac{f_{pe}}{f_{ce}} = 2.2$. Installing precise electron guns would provide more control and flexibility in the beam pulses and consequently a more accurate representation of the active experiments conducted in space.

References

[1] Beals, Dexter F. *Characterization of a Hot Cathode Helimak Plasma,* Masters Thesis, MIT, 1994.

[2] Cartwright, D. C. and P. J. Kellogg. *Observations of Radiation From an Electron Beam Artificially Injected into the Ionosphere,* Journal of Geophysical Research, Vol. 79, 1974.

[3] Chen, F. F. *Introduction to Plasma Physics,* Plenum, New York, 1977.

[4] Dobrowolny, M. *Il Nuovo Cimento,* Societa Italiana di Fisica, Vol. 17C, N. 1, 1994.

[5] Gewartowski, J. W. and H. A. Watson. *Principles of Electron Tubes,* D. Van Nostrand Company, Inc., Princeton, NJ, 1965.

[6] Gough, M. Paul, David A. Hardy, Marilyn R. Oberhardt, William J. Burke, Louise C. Gentile. *Correlator Measurements of MHz Wave-Particle Interactions During TSS 1 Electron Beam Operations,* U.S. Air Force Office of Scientific Research, 1995.

[7] Hardy, David A., Marilyn R. Oberhardt, William J. Burke, Donald C. Thompson, W. John Raitt, Louise C. Gentile. *Electron Beam Propagation Observed During TSS 1,* U.S. Air Force Office of Scientific Research, 1995.

[8] Hutchinson, Ian. *Principles of Plasma Diagnostics,* Cambridge University Press, Cambridge, 1987.

[9] Jost, R. J., H. R. Anderson, and J. O. McGarity. *Electron Energy Distributions Measured During Electron Beam/Plasma Interactions,* Geophysical Research Letters, Vol. 7, No. 7, July 1980.

[10] Kellogg, Paul J., Steven J. Monson, Robert H. Holzworth, and R. Jerry Jost. *Beam-Generated Waves in a Large Plasma Chamber,* IEEE Transactions on Plasma Science, Vol. PS-14, No. 6, December 1986.

[11] Spangenberg, Karl R. *Vacuum Tubes,* First Edition, McGraw-Hill, New York, 1948.

[12] Vedenov, A. A. and D. D. Ryutov. *Quasilinear Effects in Two-Stream Instabilities,* Reviews of Plasma Physics, Vol. 6, 1975.

[13] Walker, D. N. and E. P. Szuszczewicz. *Electrostatic Wave Observation During a Space Simulation Beam-Plasma Discharge,* Journal of Geophysical Research, Vol. 90, 1985.

USING LASER DYNAMICS TO PROBE INTRACAVITY COLLISIONAL RELAXATIONS

Jeff Nicholson

Graduate Student

Wolfgang Rudolph

Associate Professor

Department of Physics and Astronomy

The University of New Mexico

Albuquerque, NM 87131

Report for:

Faculty Summer Research Program

Phillips Laboratory

Sponsored by:

Air Force Office of Scientific Research

Phillips Laboratory, Albuquerque

September 1995

USING LASER DYNAMICS TO PROBE INTRACAVITY COLLISIONAL RELAXATIONS

Jeff Nicholson

Graduate Student

Wolfgang Rudolph

Associate Professor

Department of Physics and Astronomy

The University of New Mexico

Albuquerque, NM 87131

Abstract

The pulse parameters of a photolytic, gain-switched iodine laser are measured for different pressures of active gas and for various buffer gases. A rate equation model of the laser dynamics is fit to experimental data, to obtain information about collision-controlled relaxation processes of atomic iodine.

USING LASER DYNAMICS TO PROBE INTRACAVITY COLLISIONAL RELAXATIONS

Jeff Nicholson and Wolfgang Rudolph

1 Introduction

Photolysis of a molecule by a short laser pulse results in a nonthermal velocity distribution. Due to collisions changing the direction and the magnitude of the velocity vector the initial distribution relaxes to an isotropic Maxwellian distribution. It was found that the velocity relaxation comprises two components — a relaxation of the speed to a Maxwellian distribution at a certain equilibrium temperature and a relaxation of the direction of the velocity vector to an isotropic distribution. Being controlled by collisions, the relaxation rates of these two processes can be different and depend on the collision partners (photolytic products and buffer gas). Heavy collision partners tend to produce a directional relaxation which can be several times faster than the corresponding time constant for the speed [1]. While this effect follows from hard-sphere collision theory, the effect of the structure (e.g., mono-atomic noble gas versus polar multi-atomic molecule) is much more complex and not very well understood. The latter is due partly to the lack of suitable experimental data.

Information about the velocity can be gained by measuring the spectral absorption profile of one of the products of the photolysis as a function of time. A corresponding gain measurement in such a pump-probe experiment can be performed if one of the reaction products emerges in an excited state. A number of such experiments have been executed in the past [2], [3] and yielded collision rates as a function of collision partner (buffer gas) and the temporal behavior of the velocity profile as it developed from a nonequilibrium to an equilibrium distribution. A typical measurement required a narrow bandwidth probe laser and a time resolving data acquisition system such as a boxcar integrator or a digital storage oscilloscope.

Recently the temporal shape of a gain switched laser pulse was used to derive relaxation parameters of the active gas after short pulse photolysis [4]. The pump pulse, being shorter than all relaxation processes of interest here, prepares the sample in a certain state. The time evolution of the gain coefficient can be probed by the evolving laser pulse itself provided the cavity lifetime is sufficiently short. This technique with its high signal to noise ratio, in principle, allows determination of the velocity relaxation parameters after a single photolysis event.

Immediately after photolysis off CF_3I with a short (20 ns) pump pulse at 248 nm propagating along the laser axis the gain profile differs considerably from a Doppler profile at room temperature. Since the pump photon energy is much greater than the energy needed for dissociation the mean energy of the photofragments corresponds to a higher temperature and larger mean velocity then when at room temperature. Since the pump laser is polarized perpendicularly to the resonator axis the nascent velocity distribution is peaked in

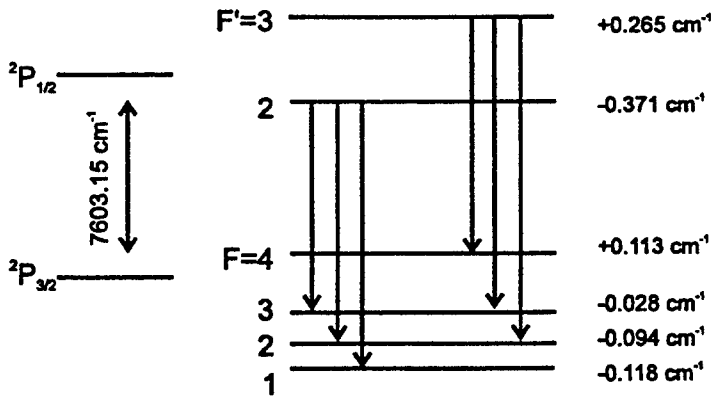


Figure 1: Energy levels and allowed transitions of atomic iodine

this direction making the gain at resonance greater as compared to an isotropic distribution. At the same time, however, the speed distribution is broader than at thermal equilibrium (room temperature). Therefore, the net gain at resonance is actually smaller as compared with the equilibrium situation.

While the laser radiation develops from spontaneous emission, the gain media cools thus narrowing the Doppler profile and increasing the gain available to the laser. At a certain intensity of the laser the gain at resonance is depleted, burning a hole in the velocity distribution, and leading to a decrease in laser power. Because of cross-relaxation processes filling the hole, the gain and thus the laser output can increase again which manifests itself as a distinct shoulder or a second pulse.

We will show in detail how the various relaxation times can be obtained for different buffer gases and pressures from fitting the experimental laser pulse profiles with a comprehensive rate equation model of the laser.

2 Experimental

When pumped at 248 nm, CF_3I dissociates with approximately 92% efficiency into excited state iodine (I^*) and CF_3. The other 8% of the iodine goes into the ground state [5]. The energy levels of atomic iodine are shown in Fig. 1. The strongest transition and the one that lases under typical experimental conditions is the F=3 to F=4 transition corresponding to a wavelength of 1.315 μm.

The experiment consisted of a photolytic iodine laser pumped by a KrF laser (see Fig 2). The gain media used was 0.8 torr CF_3I in an 85 cm cell with fused silica windows at Brewster's angle for 1.3 μm. The cell was pumped longitudinally by a 248 nm pulse with a width of approximately 20 ns and an energy of 200 mJ. The pump beam was focused to a 10mm by 4mm spot in the cell. At 0.8 torr of CF_3I with these pumping

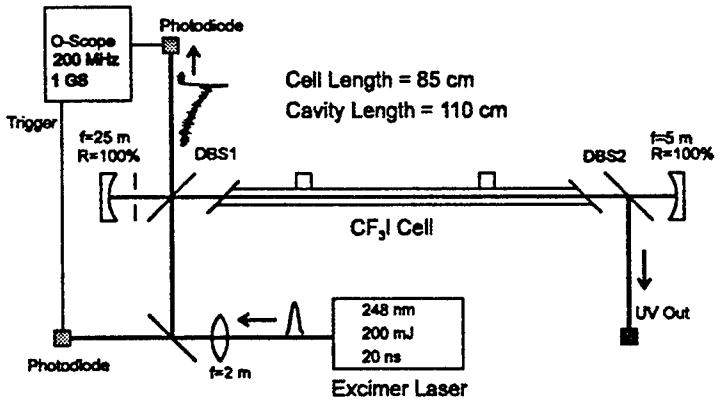


Figure 2: Schematic diagram of the photolytic, gain-switched iodine laser used to study collisional relaxation processes.

conditions, 15% of the CF_3I was dissociated into CF_3 plus I^*.

In addition to the CF_3I in the cell, buffer gases were introduced into the gain media to change collision times without changing the number of active molecules. Buffer gases used were He, Ne, Kr, and Xe. CO_2 and N_2 were also used in order to investigate the difference between the hard sphere collisions of noble gases and inelastic collisions of polyatomic molecules.

The cavity consisted of two 100% reflecting mirrors separated by 1.1m, with radius of 5 m and 25 m. The pump beam was coupled into the cavity with a dichroic beam splitter and coupled out with a second dichroic beam splitter. The output of the iodine laser was taken from a small reflection off the first dichroic beam splitter, while the second beam splitter was set at Brewster's angle for 1.3 μm. Iodine pulse energies were on the order of several μJ.

After laser action, only about 90% of the molecular fragments recombine to form CF_3I, the rest of the fragments going into processes such as dimerization [5]. I_2 is a strong quencher of the $^2P_{1/2}$ state, leading to a degradation of the iodine laser output after multiple shots of the excimer laser. Therefore, the iodine cell was refilled after every fourth or fifth excimer pulse. Note, that while these processes proved to be important after multiple shots, they do not play a significant role in single shot pulse dynamics on a time scale < 10 μs because of their slow (e. g. $\sim 4 * 10^{-31} cm^6 mol^{-2} s^{-1}$ for dimerization) rate.

The iodine pulse was detected with a fast (< 500 ps risetime) InGaAs photodiode and viewed on a digital oscilloscope (Textronix TDS 350). A part of the excimer pulse was also incident on a photodiode. This signal was used to trigger the oscilloscope and give an absolute measure of the delay between the arrival of the excimer pulse and onset of the iodine pulse.

A typical iodine pulse with 0.8 torr CF_3I in the gain cell is shown in Fig. 3. The excimer pulse occurs

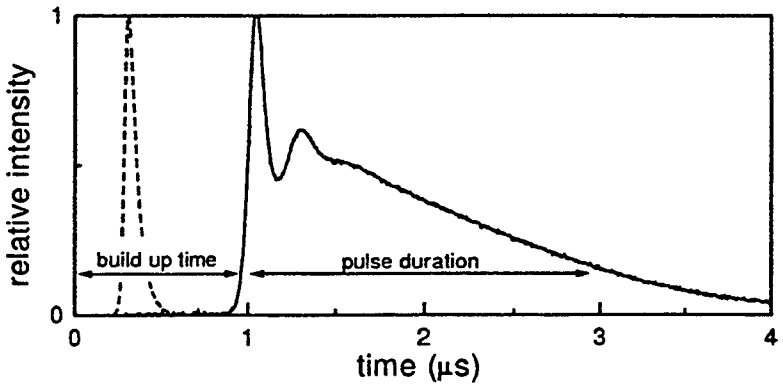


Figure 3: A typical pulse from a low pressure photolytic iodine laser. CF_3I pressure was 0.8 torr. The dotted line is the result from a simple rate equation model without velocity relaxation mechanisms. The pump pulse occurs at t=0.

at $t = 0$, and is essentially a delta function on this time scale. The dotted pulse is the result obtained from a simple three level rate equation model that does not include any velocity relaxations. The experimental pulse shape has a much longer delay between pump pulse and onset of lasing, a much slower decay time, and a distinct second peak. All of these features are the result of the finite relaxation times for the velocity distribution. Fig. 3 also shows how the build up time and pulse duration were defined. These quantities were measured at 10% of the pulse peak intensity.

Simultaneous lasing of multiple longitudinal modes is frequently observed in gain switched iodine lasers [6],[7]. Fig. 4 shows two iodine pulses taken under identical experimental conditions that show clear longitudinal mode beating. However, the relative strengths and build up times are different for the two modes for the two pulses, obscuring or enhancing the pulse features shown Fig 3. Therefore, an intracavity aperture was used to force the laser to oscillate both with a single longitudinal mode and in TEM_{00} mode. An aperture diameter of less than 2 mm was used as compared to a TEM_{00} waist of 2.2 mm. This eliminated mode beating in approximately 50% of the iodine pulses. The aperture also forced the laser frequency to be much less than half the free spectral range of the cavity (70 MHz) away from line center.

Although the 3-4 transition is the strongest transition in atomic iodine, under certain experimental conditions both the 3-4 and 2-2 transition have lased [8],[9]. In order to be sure only the 3-4 transition was lasing, a streak camera (resolution < 5 ps) was used to look for the fast modulation that would result from simultaneous lasing and beating of two hyperfine modes. This mode beating would have a frequency of 13 GHz, the difference frequency between the 2-2 and 3-4 transitions. However, there was no evidence of a 13 GHz signal, in any of the measurements made with the streak camera.

In addition, the rate equation model, described in Section 3.2 was used to investigate the possibility of

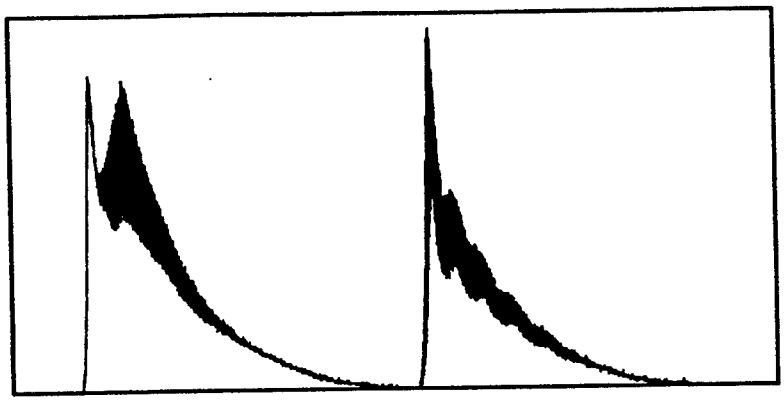


Figure 4: Two examples of iodine pulses with clear longitudinal mode beating

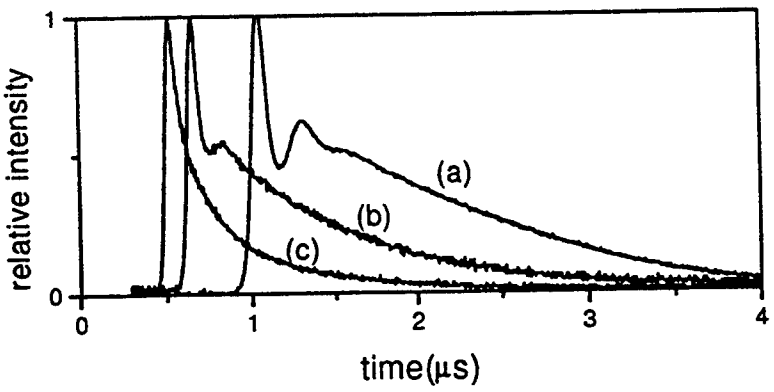


Figure 5: The change in pulse shape as Xe is added to the gain cell. 0.8 torr CF_3I + (a) 0 torr Xe, (b) 0.8 torr Xe, (c) 2.4 torr Xe. The peak intensities have been normalized to one for comparison purposes

the 2-2 line lasing. The cross section of the 3-4 transition is twice as large as the 2-2 transition. When only the 2-2 transition is included in the model, with the 3-4 transition artificially suppressed, the transition is able to lase, but the build up time is 2 μs greater and the peak intensity is one fourth of the values obtained when only the 3-4 line is included in the model. However, when both (single longitudinal mode) hyperfine transitions are included in the model, only the 3-4 line is able to lase. The 3-4 line builds up intensity quickly compared to the 2-2 line, and depletes the F=3 level. Then a fast population transfer due to collisions with ground state iodine [10] equalizes the upper hyperfine levels. This depletes the population in the upper level of the 2-2 line, preventing it from lasing. From this, we concluded that the experimental iodine pulses contained the 3-4 line only.

Fig. 5 shows the change in pulse shape as a buffer gas (in this case, Xe) is introduced into the gain cell.

As the buffer gas' partial pressure is increased, the build up time of the pulse decreases, the second peak disappears, and the pulse width decreases. The pulse energy remains approximately constant with increasing buffer gas pressure.

A summary of the effects of buffer gas on pulse parameters is shown in Fig. 6. Because the number of active molecules remains constant, the change in pulse shapes can be attributed to changing the collisional dynamics caused by adding buffer gas. Fig. 6(a) shows Noble gases; Fig. 6(b) shows polyatomic molecules. Xe is included in (b) for comparison purposes.

As the partial pressure of buffer gas in the gain cell increases, collision rates increase, and build up time and pulse duration decrease, and the second peak becomes smaller and disappears. However, more massive buffer gases are more effective than lighter buffer gases in thermalizing the velocity distribution. In fact, the changes in the velocity relaxation times caused by helium are too small to account for the observed changes in the pulse shape, but can be explained by the small increase in the homogeneous linewidth caused by pressure broadening (see Section 4). The increase in homogeneous linewidth is insufficient to explain the changes in pulse shape caused by the more massive buffer gases, however.

3 Theoretical Model

3.1 Collision Dynamics

Immediately after photodissociation, the velocity distribution of the photofragments is separable [11]; it can be written as a product of a function of speed,f and an angular function,I,

$$F(\vec{v},t) = f(v,t)I(\theta,\phi,t) \tag{1}$$

where the angular distribution function $I(\theta,\phi,t)$ has the form

$$I(\theta,\phi,t) = \frac{1}{4\pi}\,[1+\beta(t)\;P_2(\cos\theta)]. \tag{2}$$

β is the anisotropy parameter, and θ and ϕ are the angles between the fragment's velocity vector and the electric field vector of the dissociating beam. Depending on whether the molecule undergoes a parallel or perpendicular transition during dissociation, β can take the value 2 or -1 respectively. CF_3I dissociating into CF_3 and I^* is a parallel transition with $\beta = 1.4$ [12].

In previous Doppler spectroscopy experiments, a probe beam was used to measure the Doppler profile. In general the propagation direction of the probe beam is different from that of the dissociating beam. However, in this experiment, the probe beam is the laser field, and because the cell is longitudinally pumped, the pump and probe beams are parallel. The relevant coordinate axis in relation to the resonator axis is shown in Fig. 7. Immediately after photodissociation, the I^* atoms tend to move perpendicular to the laser axis.

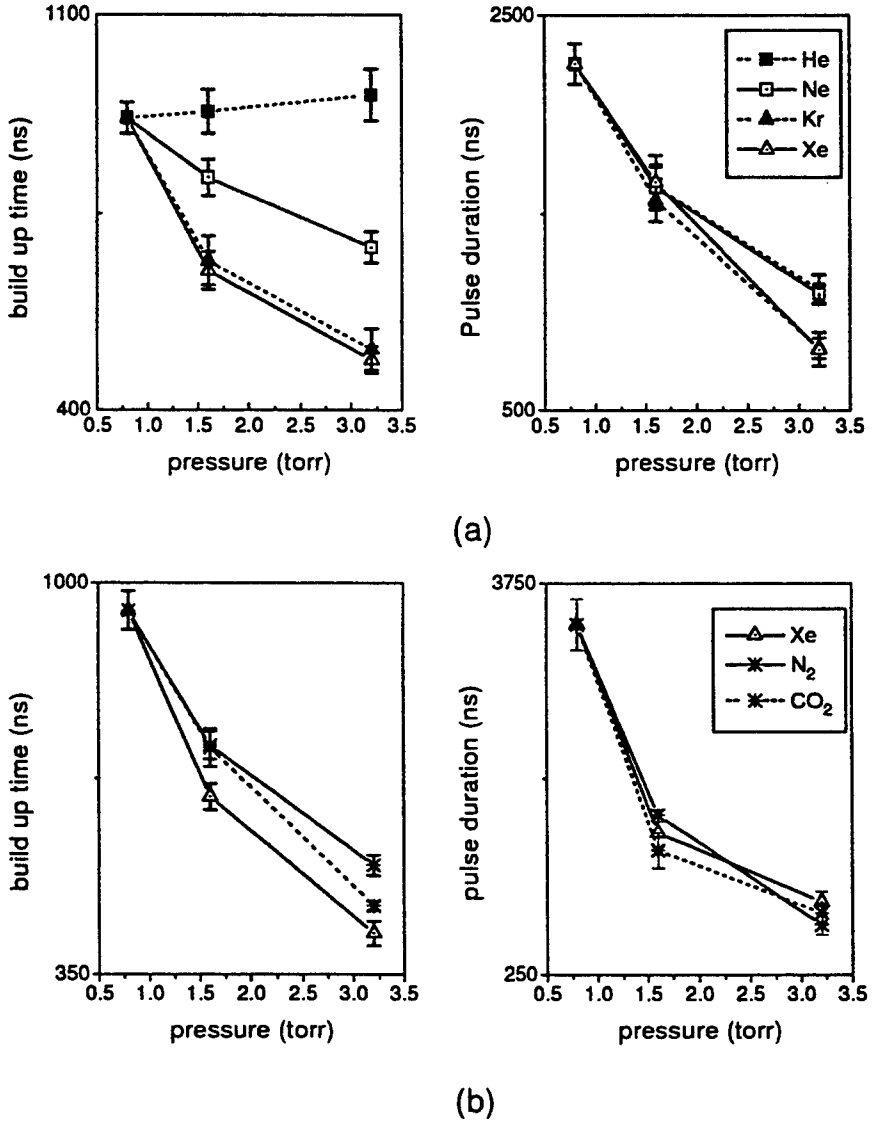


Figure 6: Summary of the effects of the addition of buffer gas on the build up time and pulse duration. (a) Noble gases. (b) Polyatomic molecules.

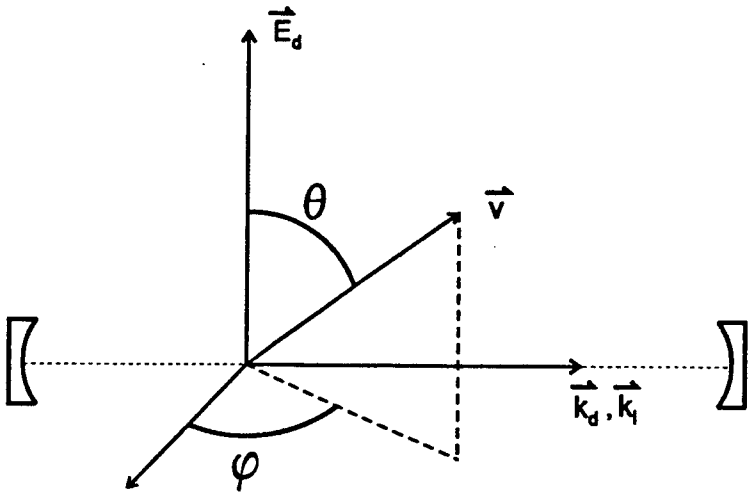


Figure 7: Coordinate axis in relation to the iodine resonator axis. The subscript d refers to the dissociating beam, and l refers to the iodine laser beam.

It was recently shown that as $F(\vec{v},t)$ relaxes to thermal equilibrium, it remains approximately separable [13]. The mean velocity relaxes exponentially to its equilibrium value with a time constant τ_v

$$\bar{v}(t) = \bar{v}_{\text{fin}} + (\bar{v}_{\text{in}} - \bar{v}_{\text{fin}})e^{-t/\tau_v} \tag{3}$$

where $\bar{v}_{\text{in}}$ and $\bar{v}_{\text{fin}}$ are the initial and final mean velocities respectively. β relaxes to zero (isotropy) with a time constant τ_b

$$\beta(t) = \beta_{\text{in}} e^{-t/\tau_b} \tag{4}$$

A full analysis using the Boltzmann collision operator shows that the ratio τ_v/τ_b becomes larger for more massive collision partners [1]. The approximate separability and the relative ratios of the time constants has been experimentally verified for a few examples using dissociation of C_3F_7I at 266 nm [2].

Stimulated emission has a much smaller effect then collisions on the separability of the velocity distribution. If we assume, for the moment, that the laser field is single frequency, and there is no homogeneous broadening of the laser transition, then only atoms moving perpendicular to the laser axis are removed from the velocity distribution of the excited state atoms. If we begin with a separable velocity distribution that has the form of Eq. (1) then change in the velocity distribution due to stimulated emission can be written as

$$f(v,t)I(\theta,\phi,t) \rightarrow f(v,t)(1 - a\delta(\theta))I(\theta,\phi,t) = f(v,t)I'(\theta,\phi,t) \tag{5}$$

where a is the fraction of total atoms removed from the excited state due to stimulated emission. $F(\vec{v},t)$ is still separable, although I is no longer of the form of Eq. (2). Although this is strictly true for a monochromatic

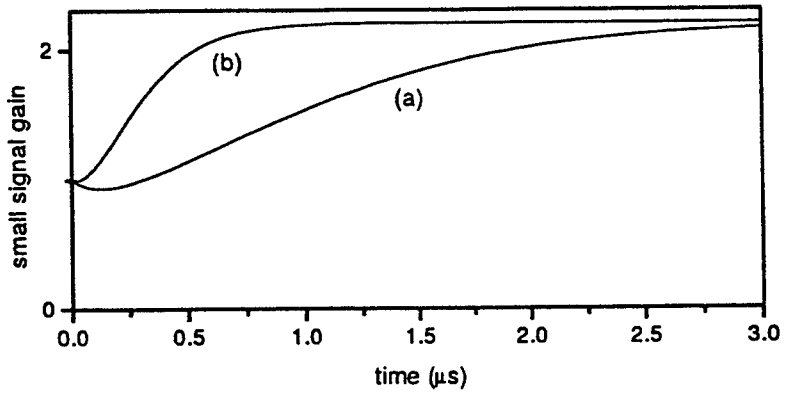


Figure 8: Small signal gain at line center. (a) $\tau_v = 650$ns, $\tau_b = 120$ns (b) $\tau_v = 180$ns, $\tau_b = 60$ns

laser field and no homogeneous broadening, the linewidth of the laser and the homogeneous linewidth are so much smaller than the inhomogeneous linewidth ($\approx$ 10 MHz vs 1 GHz) that it remains a good assumption. Thus, the dominant reasons for the loss of separability remain those discussed in Ref. [2], such as collisions between I^* and buffer gas.

The Doppler profile is obtained from the velocity distribution by summing all the fragments that have the same velocity component in the direction of the probe beam

$$\Gamma_D(w, \theta', \phi', t) = \int d^3v \; F(\vec{v}, t) \; \delta(w - \vec{v} \cdot \vec{k}_p) \tag{6}$$

where w is the velocity component in the direction of the probe beam. It is related to the Doppler frequency shift by $\nu - \nu_0 = \nu_0 \frac{w}{c}$ where ν_0 is the center frequency of the transition and c is the speed of light. The primed angles refer to the probe direction. For the particular case where the pump and probe are parallel, using Eqs (1) and (2), Eq. (6) reduces to [3]

$$\Gamma_D(\nu - \nu_0, t) = \frac{c}{\nu_0} \int_{|\frac{c}{\nu_0}(\nu - \nu_0)|}^{\infty} \frac{v}{2} \left[1 - \frac{\beta(t)}{4} \left(3 \left(\frac{c(\nu - \nu_0)}{\nu_0 v} \right)^2 - 1 \right) \right] \; f(v, t) \; dv. \tag{7}$$

From here on, we write Γ_D as a function of ν rather than w. Eq. (7) can be solved analytically for line center ($\nu = \nu_0$), assuming a Maxwellian velocity distribution for $f(v, t)$:

$$\Gamma_D(0, t) = \frac{2c}{\nu_0 \pi} \left(1 + \frac{\beta(t)}{4} \right) \frac{1}{\bar{v}(t)}. \tag{8}$$

A plot of Eq. (8) is shown in Fig. 8 for two values of τ_v and τ_b. If we neglect for the moment the small amount of ground state iodine produced after dissociation, $\Gamma_D(0, t)$ is proportional to the small signal gain at line center. This clearly shows that as the relaxation times decrease, the gain available to the iodine laser reaches equilibrium faster, shortening the build up time of the iodine pulse. In addition, the build-up time of the laser pulse is due primarily to τ_v and only to a lesser extent τ_b.

In order to connect the relaxation of the Doppler profile $\Gamma_D(\nu-\nu_0,t)$ to the relaxation of the velocity distribution $F(\vec{v},t)$, take the time derivative of Eq. (6)

$$\frac{d\Gamma_D(\nu-\nu_0,t)}{dt}=\int d^3v\left[\frac{dI(\theta,\phi,t)}{dt}f(v,t)+\frac{df(v,t)}{dt}I(\theta,\phi,t)\right]\ \delta\left((\nu-\nu_0)-\frac{\nu_0}{c}\vec{v}\cdot\vec{k}_p\right). \tag{9}$$

Therefore, for small enough time steps, the contribution to the time rate of change of $\Gamma_D(\nu-\nu_0)$ can be separated into two terms. One term controls how $\Gamma_D(\nu-\nu_0,t)$ changes due to a relaxation of the angular anisotropy function $I(\theta,\phi,t)$, and the other term controls how $\Gamma_D(\nu-\nu_0,t)$ changes due to a relaxation of the speed distribution $f(v,t)$.

The integral that arises from the first term in the brackets in Eq. (9) describes how the Doppler profile returns to equilibrium due to a relaxation of the anisotropy only. If $I(\theta,\phi,t)$ relaxes to equilibrium exponentially, (as is the case for the anisotropy function in Eq. (2)), then the relaxation of $\Gamma_D(\nu-\nu_0,t)$ to equilibrium due to the anisotropy relaxation in $F(\vec{v},t)$ can be written

$$\frac{d\Gamma_D(\nu-\nu_0,t)}{dt}=\frac{\Gamma_D^{eq}(\nu-\nu_0,t)-\Gamma_D(\nu-\nu_0,t)}{\tau_b}. \tag{10}$$

$\Gamma_D^{eq}(\nu-\nu_0,t)$ is the Doppler profile whose velocity distribution has the same mean velocity as $\Gamma_D(\nu-\nu_0,t)$ but no anisotropy; this is just a Gaussian profile with a (time dependent) full width at half maximum $\Delta\nu_D(t)$. $\Delta\nu_D(t)$ is simply related to the mean velocity by $\Delta\nu_D(t)=\nu_0\sqrt{\pi\ln 2}/c\bar{v}(t)$. Because the Doppler width is proportional to the mean velocity, $\Delta\nu_D$ also relaxes exponentially with the time constant τ_v

$$\Delta\nu_D(t)=\Delta\nu_{\text{fin}}+(\Delta\nu_{\text{in}}-\Delta\nu_{\text{fin}})e^{-t/\tau_v} \tag{11}$$

where $\Delta\nu_{\text{in}}$ and $\Delta\nu_{\text{fin}}$ are the initial and final Doppler width respectively.

Eq. (10) is sufficiently general to describe any relaxation of an anisotropy in the velocity distribution. It will serve to relax the initial angular anisotropy caused by photodissociation of CF_3I. In addition, stimulated emission from radiation at the center of the inhomogeneously broadened transition will remove atoms moving perpendicular to the laser axis and burn a hole in the Doppler profile. Eq. (10) will also relax such an induced angular anisotropy caused by the build up of laser radiation.

The second term in Eq. (9) shows how the Doppler profile relaxes due to the cooling of the velocity distribution. If we assume that immediately after dissociation, $f(v,t)$ is a Maxwellian distribution with a high temperature, $f(v,t)$ remains a Maxwellian distribution as the gas cools.

$$f(v,t)\propto\exp\left(-\frac{4v^2}{\pi\bar{v}^2(t)}\right) \tag{12}$$

Then, at time $t+dt$, $f(v,t+dt)$ is

$$f(v,t+dt)\propto f\left(\frac{\bar{v}(t)}{\bar{v}(t+dt)}v,t\right) \tag{13}$$

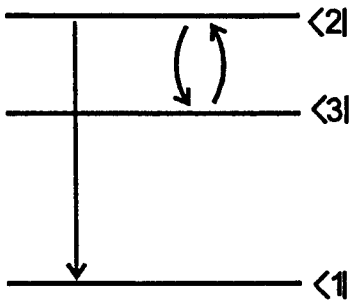


Figure 9: Diagram for levels and transitions used in the rate equation model. The laser transition is from $\langle 2|$ to $\langle 1|$. Population in $\langle 2|$ and $\langle 3|$ is equalized through collisions with population in $\langle 1|$.

Using Eq (6) and keeping $I(\theta, \phi, t)$ constant for small dt, we can relate the Doppler profile at time $t + dt$ to the Doppler profile at time t by

$$\Gamma_D(\nu - \nu_0, t + dt) = \alpha(t)\Gamma_D\left(\frac{\Delta\nu_D(t)}{\Delta\nu_D(t + dt)}(\nu - \nu_0), t\right) \tag{14}$$

where the proportionality constant $\alpha(t)$ is used to insure

$$\int \Gamma_D(\nu - \nu_0, t)d\nu = 1 \tag{15}$$

Therefore the functional form of $\Gamma_D(\nu - \nu_0, t)$ does not change its shape; only the width of the function changes. This gives a method for determining how $\Gamma_D(\nu - \nu_0, t)$ changes due to relaxation of $\Delta\nu_D(t)$.

3.2 Laser Rate Equations

The energy level diagram used for the laser model is shown in Fig. 9. Because of the fast relaxation processes between the ground state hyperfine levels (time constant $<$ 1 ns [8]) we lump the lower levels shown in Fig. 1 into an effective ground state $\langle 1|$. Lasing can then occur between $\langle 2|$ and $\langle 3|$ and this ground state. The result of the photolysis is a population in the two excited states of iodine distributed according to the degeneracies of the states [10]. There is a relaxation process involving collisions with unexcited atomic iodine between levels $\langle 2|$ and $\langle 3|$ with a rate constant $k_{2,3} \approx 1.5 * 10^{-9}\text{cm}^3\text{mol}^{-1}\text{s}^{-1}$ [10] which tends to equalize the population once an asymmetry has been created:

$$\frac{d}{dt}N_2 = k_{2,3}N_1\left(N_3 - \frac{g_3}{g_2}N_2\right) \tag{16}$$

where N_i and g_i $(i = 1, 2, 3)$ is the occupation number and degeneracy, respectively of level i, and $N_2 + N_3$ is the total number density of excited iodine. Because of this cross relaxation and the two times larger interaction cross-section, under typical conditions used in our experiments, lasing occurs only between states $\langle 2|$ and $\langle 1|$.

To formulate the rate equations for the laser we assume that only one single longitudinal mode oscillates and is in resonance with the center of the inhomogeneously broadened gain transition $\langle 2| \rightarrow \langle 1|$ at a frequency $\nu = \nu_0$. The laser spectrum is supposed to be (much) smaller than the homogeneous line width $\Delta\nu_h$.

Let us define occupation numbers of a subensemble of particles with a resonance frequency ν per frequency interval $[\nu - \frac{d\nu}{2}, \nu + \frac{d\nu}{2}]$, $\bar{N}_i(\nu - \nu_0)$, such that

$$\bar{N}_i(\nu) = \Gamma_D(\nu - \nu_0)\ N_i \tag{17}$$

and

$$N_i = \int_0^\infty \bar{N}_i(\nu)\ d\nu \tag{18}$$

The rate equations for these occupation numbers, including relaxation between the upper hyperfine levels, and relaxation of the velocity distribution anisotropy as described in Eq. (10), can be written as:

$$\begin{aligned} \frac{d}{dt}\bar{N}_2 &= -q_{21}\Gamma_h(\nu-\nu_0)F\left(\bar{N}_2 - \frac{g_2}{g_1}\bar{N}_1\right) + \frac{1}{\tau_b}\left[\bar{N}_2^{eq} - \bar{N}_2\right] \\ &\quad +\Gamma_D(\nu-\nu_0)k_{23}N_1\left(N_3 - \frac{g_3}{g_2}N_2\right) \end{aligned} \tag{19}$$

$$\frac{d}{dt}\bar{N}_1 = q_{21}\Gamma_h(\nu-\nu_0)F\left(\bar{N}_2 - \frac{g_2}{g_1}\bar{N}_1\right) + \frac{1}{\tau_b}\left[\bar{N}_1^{eq} - \bar{N}_1\right] \tag{20}$$

$$\frac{d}{dt}N_3 = -\Gamma_D(\nu-\nu_0)k_{23}N_1\left(N_3 - \frac{g_3}{g_2}N_2\right) \tag{21}$$

$$\tag{22}$$

where

$$\Gamma_h(\nu-\nu_0) = \frac{\Delta\nu_h}{2\pi\left[(\nu-\nu_0)^2 + (\Delta\nu_h/2)^2\right]} \tag{23}$$

is the homogeneous line shape factor, $\Gamma_D(\nu-\nu_0)$ is the (time dependent) inhomogeneous lineshape function, and $q_{21} = \frac{7}{12}\frac{A_{21}\lambda_0^2}{8\pi}$ with A_{21} being the Einstein coefficient. The time τ_b describes the relaxation of the occupation numbers towards an equilibrium value $\bar{N}_i^{eq} = \Gamma_D^{eq}(\nu-\nu_0)\ N_i$, where $\Gamma_D^{eq}(\nu-\nu_0, t)$ is determined by the instantaneous Doppler width, which relaxes with time constant τ_v.

The corresponding rate equation for the photon flux F at $\nu = \nu_0$ reads:

$$\frac{d}{dt}F = -cq_{21}F\int_0^\infty \Gamma_h(\nu-\nu_0)\left[\bar{N}_2(\nu) - \frac{g_2}{g_1}\bar{N}_1(\nu)\right]d\nu - \frac{1}{\tau_r}F \tag{24}$$

where τ_r is the cavity life time.

The above rate equations do not contain all the necessary information to fully describe the relaxation of the velocity distribution. They do have the relaxation of the angular anisotropy with time constant τ_b, as detailed in Eq. (10). A separate procedure was utilized in the computer model to implement the affect of cooling on the occupation numbers as described at the end of the previous section. Every time step, the horizontal scale of $\Gamma_D(\nu-\nu_0, t)$ was changed according to Eqs. (14) and (15).

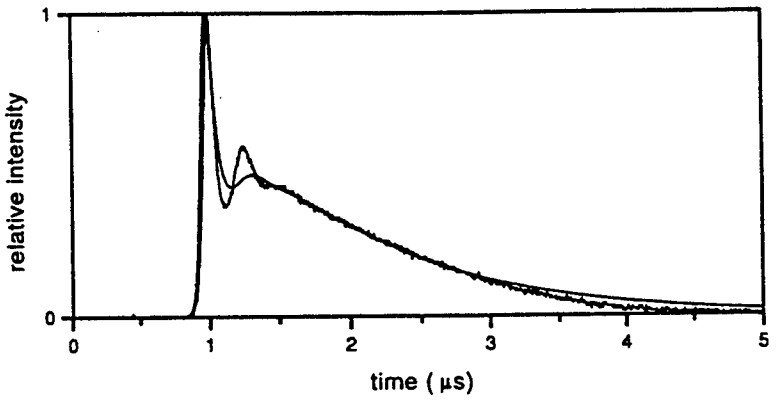


Figure 10: Experimental and theoretical pulse shapes for 0.8 torr CF_3I. The time constants used were $\tau_v = 650$ ns and $\tau_b = 120$ ns

The pumped region of the gain cell was a column 4 mm by 10 mm. Within this region approximately 15% of the CF_3I was dissociated into $CF_3 + I^*$. This column was surrounded by a bath of undissociated CF_3I. A diffusion equation calculation was done in order to see how long it would take the unexcited CF_3I to diffuse into the center region of the cell that comprised the mode volume of the iodine laser. We estimated it would take approximately 2.5 μs after the arrival of the pump pulse for the cold CF_3I to diffuse into the center region. This would have an effect on the tail of the pulse only (see Fig. 3), and therefore, diffusion was neglected in later calculations.

The initial values for the occupation numbers were set using Eq. (7). A value of $\beta = 1.4$ was assumed. The mean translational energy of the fragments of CF_3I after dissociation at 248 nm has been measured to be 10200 cm^{-1} for the $CF_3 + I^*$ channel [14]. This gives an initial Doppler width of 950 MHz for I^*. With a large bath of buffer gas, the Doppler width would relax to its room temperature value of 250 MHz. However, under our experimental conditions, we did not have a bath at room temperature. Instead, the excess energy in the dissociating photons is redistributed among the species in the gain cell according to their specific heats. Thus, with 15% of the 0.8 torr CF_3I being dissociated, the final Doppler width of the I^* was 330 MHz. With 2.4 torr of buffer gas in the cell, the final Doppler width was 290 MHz. The linewidth of the homogeneous transition $\Delta\nu_h$ used in Eq. (23) was calculated using published pressure broadening coefficients for the species present in the gain cell [15].

4 Comparison of Experiment and Theory

The rate equation model was used first to fit the experimental pulse shapes obtained with only CF_3I in the cell. The experimental and theoretical pulse shapes are shown in Fig. 10. The time constants used in

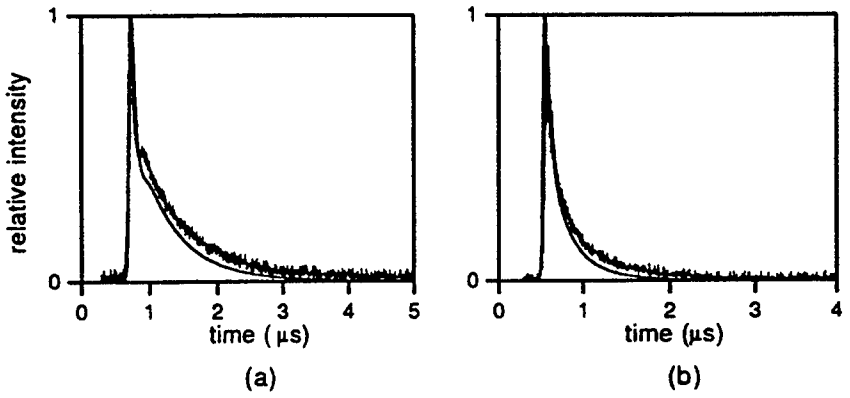


Figure 11: Experimental and theoretical pulse shapes for 0.8 torr CF_3I plus (a) 0.8 torr Xe, (b) 2.4 torr Xe. The contribution to the relaxation times $\tau_{v,b}$ due to Xe was decreased by a factor of three from (a) to (b)

the rate equation model were $\tau_v = 650$ ns and $\tau_b = 120$ ns.

The relaxation times τ_v and τ_b can be expressed as rate constants k_v and k_b respectively. In other words, the contribution to τ_v and τ_b is inversely proportional to the partial pressure, and the relaxation times can be written as

$$\frac{1}{\tau_v} = \sum_i \frac{1}{\tau_{v,i}} \tag{25}$$

with a similar equation for τ_b. The summation is over the various species present in the gain cell, i.e. CF_3I, CF_3, I, and the buffer gas.

Next, the data for 0.8 torr CF_3I plus the various buffer gas was fit. The contribution of the buffer gas to the relaxation times can be obtained using Eq. (25) and the relaxation times from the fit for pure CF_3I in the cell. For every buffer gas used, once the data for 0.8 torr CF_3I plus 0.8 torr buffer gas had been fit, the data with higher partial pressure of buffer gas was simulated simply by keeping the contributions to τ_v and τ_b inversely proportional to the buffer gas partial pressure. A typical fit for Xe as the buffer gas is shown in Fig. 11. Because the effects of Kr and Xe were the same to within experimental error (see Fig. 6) no effort was made to fit the Kr data separately. Table 1 contains a summary of the rate constants obtained for the speed and anisotropy relaxations and the ratio τ_v/τ_b for all buffer gases used.

The difference between the noble gases and the polyatomic molecules as buffer gas appears in the ratio τ_v and τ_b. Although N_2's atomic mass is only slightly larger than Ne's, the ratio of τ_v to τ_b is much greater, whereas the ratio for CO_2 is the smallest measured. This suggests that the hard sphere collision theory no longer applies for these polyatomic molecules.

Earlier measurements of He used as a buffer gas with C_3F_7I estimated relaxation times for He to be approximately 10 times as long as for C_3F_7I [2]. If we use this estimate along with the relaxation times

Buffer Gas	$k_v * 10^{11}$ cm^{3}mol^{-1}s^{-1}	$k_b * 10^{11}$ cm^{3}mol^{-1}s^{-1}	τ_v/τ_b
He	—	—	—
Ne	1.1	3.0	2.7
Kr	5.2	11	2.1
Xe	5.2	11	2.1
N_2	2.1	11	4.0
CO_2	3.7	6.5	1.7

Table 1: Rate constants obtained from the rate equation model

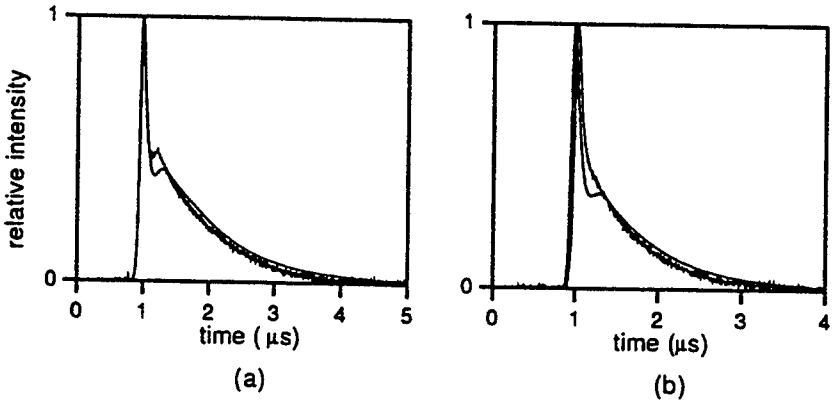


Figure 12: Experimental and theoretical pulse shapes for 0.8 torr CF_3I plus (a) 0.8 torr He, (b) 2.4 torr He. τ_v and τ_b were not changed from (a) to (b). Only $\Delta\nu_D$ and $\Delta\nu_h$ were changed.

measured for CF_3I, and assume the relaxation times for C_3F_7I and CF_3I to be comparable,the change in the total relaxation times is too small to account for the measured change in the pulse shape as He was added to the gain cell. In fact, we were able to fit the measured pulse shapes by changing the homogeneous line width and final inhomogeneous linewidth only, without changing τ_v or τ_b (see Fig. 12). Thus, we concluded that the affect of He on the relaxation times was too small to be measured at our working pressures.

5 Conclusions

The pulsed output of an iodine photodissociation laser operating at a low pressure of the active medium has distinct features which are not explained in a simple rate equation model. These features include a long delay between dissociation of the CF_3I and onset of lasing, a long pulse duration as compared to the

cavity life time, and a distinct second peak in the iodine pulse. Furthermore, these features are dependent on the amount of buffer gas that is introduced into the cell, although the number of active iodine atoms is kept constant. As more buffer gas is used, the build up time and pulse duration decreases, and the second peak becomes a shoulder and finally disappears. We were able to model all of these phenomena using a comprehensive rate equation model that included the relaxation of the speed and the angular anisotropy in the velocity distribution and relaxation between the $^2P_{1/2}$ hyperfine levels. This model neglected effects due to diffusion and the existence of all but one (the 3-4) of six transitions in atomic iodine. Although the rate equation model was able to fit the build up times, pulse duration, and decay in the tail of the pulse strikingly well for all buffer gases, reproducing the exact shape of the second peak proved to be difficult. There are several possible explanations for this.

From the beginning, the separability of the velocity distribution was only an approximation, and one of the effects that tends to reduce the separability over time is collisions; the more massive the collision partner, the greater the effect [13]. Because rapid changes in $I(\theta, \phi, t)$ are being affected in the area of the second peak in the pulse, we expect the degree of separability of $F(\vec{v}, t)$ plays an important role in the dynamics of the second peak.

Furthermore, during the modeling of the data, it was found that the homogeneous linewidth was a critical factor in determining the shape of tail of the pulse, because it determines the width of the hole that is burned in the Doppler profile, and thus, how that hole is filled. In the model, it was assumed that the homogeneous linewidth was a constant. However, this is not the case. In the early time after dissociation, before the iodine atoms have had a chance to undergo many collisions, the homogeneous linewidth is poorly defined. And in later times, as the gas comes into thermal equilibrium and diffusion begins to play a more important role, the homogeneous linewidth must certainly be time dependent. Further experiments are required to determine the dynamics of the homogeneous linewidth.

Although we could not obtain relaxation times for pure CF_3I, since there was always a significant amount of CF_3I plus dissociation products in the active volume of the laser, we were able to obtain relaxation times for a number of buffer gases. In addition, we were able to successfully explain in detail the dynamics of the laser pulses as a function of buffer gas pressure. Our measurements confirmed that the heavier buffer gases were more effective in relaxing the velocity than in relaxing the angular anisotropy. In addition, the polyatomic molecules showed different ratios for τ_v and τ_b than the monatomic Noble gases.

References

[1] B. Shizgal and R. Blackmore. *Chemical Physics*, 77:417, 1983.

[2] J. I. Cline, C. A. Taatjes, and S. R. Leone. *J. Chem. Phys.*, 93:6543, 1990.

[3] R. Schmiedl, H. Dugon, W. Meier, and K. H. Welge. *Z. Phys. A*, 304:137, 1982.

[4] J. Nicholson, W. Rudolph, J. McIver, R. Tate, and G. Hager. *Applied Physics Letters*, 67:64, 1995.

[5] G.Brederlow, E.Fill, and K.J.Witte. *The high-power iodine laser.* Springer, 1983.

[6] H.J.Baker and T.A.King. *J.Phys.D:Appl.Phys.*, 8:L31, 1975.

[7] Heon Hwang, Min Hee Lee, and Sang Soo Lee. *IEEE J. of Quantum Electron.*, QE-18:148, 1982.

[8] V. A. Alekseev, T. L. Andreeva, V. N. Volkov, and E. A. Yukov. *Soviet Physics JETP*, 36:238, 1973.

[9] M.A.Kelly, J.K.McIver, R.F.Shea, and G.D.Hager. *IEEE J. Quantum Electron.*, QE-27:263, 1991.

[10] W. Thieme and E. Fill. *Optics Communications*, 36:361, 1981.

[11] R. N. Zare. *Molecular Photocemistry*, 4:1, 1972.

[12] G. N. A. Van Veen, T. Baller, A. E. De Vries, and M. Shapiro. *Chemical Physics*, 93:277, 1985.

[13] C. A. Taatjes, J. I. Cline, and S. R. Leone. *J. Chem. Phys.*, 93:6554, 1990.

[14] Peter Felder. *Chemical Physics*, 143:141, 1990.

[15] H. J. Baker and T. A. King. *J. Phys. D: Appl. Phys.*, 9:2433, 1976.

AN INTRODUCTION TO PULSED POWER TECHNIQUES FOR X-RAY LASERS

Bryan R. Olsen
Graduate Student
Department of Physics and Astronomy

Brigham Young University
Provo, Utah 84606

Final Report for:
Graduate Student Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington DC

and

Phillips Laboratory

July 1995

AN INTRODUCTION TO PULSED POWER TECHNIQUES FOR X-RAY LASERS

Bryan R. Olsen
Graduate Student
Department of Physics and Astronomy
Brigham Young University

Abstract

Recent demonstrations of x-ray wavelength lasers in capillary discharge devices indicate the important role that pulsed power technology will play in the development of "table top" x-ray lasers. This paper reports on the Marx generator that is being prepared for x-ray laser studies at Phillips Laboratory. Some research groups have also theorized that fast pulses, at energies low enough that a Marx generator is not necessary, could be useful in creating the necessary plasma conditions for x-ray lasing. In light of this theory, details on a small pulse forming network are also contained in this report.

AN INTRODUCTION TO PULSED POWER TECHNIQUES FOR X-RAY LASERS

Bryan R. Olsen

Introduction

Lasers operating in the soft x-ray region of the spectrum have been researched for many years in hopes of producing better sources for plasma diagnosis and biological imaging. The majority of the operational x-ray lasers use exploding foil generated plasmas pumped by very large optical lasers such as the Nova laser at Lawrence Livermore National Laboratory[1]. The cost of operation of such large facilities, however, makes the practical application of exploding foil x-ray lasers unfeasible. Recently, however, Jorge Rocca and his coworkers at Colorado State University have demonstrated an x-ray laser based on a plasma produced in a small capillary by a fast electrical discharge. Rocca's device uses a 700-kV Marx generator to pulse charge a 3-nF liquid dielectric capacitor which is then rapidly discharged through the capillary[2]. Researchers led by Larry Knight at Brigham Young University have suggested that the plasma conditions necessary for lasing could be established in a capillary using a slow discharge to ionize and heat the plasma followed by a very fast (~1 ns) low energy pulse to create a population inversion. This paper reports on the Marx generator that is being prepared at Phillips Laboratory to duplicate Rocca's work. Also included is the design of a smaller pulser built at Phillips lab that could be modified to produce the short pulses required by the BYU laser scheme.

Cronos Marx Generator

Researchers in the high energy plasma physics division at Phillips Lab have recently begun an x-ray laser program with the initial goal of duplicating Rocca's work. In order to reduce the time and expense required to build an operational x-ray laser device, a decision was made to refurbish the existing Cronos Marx generator and use it to power the capillary discharge. The Cronos consists of eleven stages, each with a capacitance of 275 nF. Fig. 1 is a schematic of the Cronos Marx. This Marx generator is significantly larger than the one used by Rocca in his experiments so a water resistor has been included in the design of the initial discharge device (see Fig. 2) to absorb the excess energy supplied by the Marx. Computer simulations of the system's performance (Fig. 3) predict that the discharge should closely match the parameters reported by Rocca so there is a high degree of confidence in being able to duplicate Rocca's results once the device has been completed. Currently, the refurbishment of the Marx--cleaning the switches, replacing a damaged capacitor and

installing a trigger circuit--has been completed and progress is being made on the manufacture of the water capacitor and other components of the discharge device.

Fast Pulser

For a second project unrelated to the x-ray laser program, a compact pulse forming network (PFN) was recently constructed which is designed to deliver a 150 ns wide square wave to a 25 ohm load. Fig. 4 is a schematic of the PFN which was easily constructed using capacitors salvaged from unused equipment and hand wound coils. Tests of the PFN, conducted at charge voltages of 15-20 kV, demonstrated pulses very similar in shape to the simulation shown in Fig. 5. Ultimately, the PFN will be immersed in oil which will allow charge voltages to be increased to 100 kV resulting in a 50 kV output. Although this particular device is not intended to be used in the x-ray laser program, it should be possible to build a similar device that would fulfill the requirements for the fast pulse discharge required by the group at BYU for their x-ray laser proposal.

Conclusion

Current capillary discharge experiments show great promise for making x-ray lasers economically feasible in which case pulse power systems will take over the role formerly played by huge optical lasers in pumping the x-ray system. In the immediate future, Marx generators will most likely power the majority of the capillary discharge x-ray laser devices. There are, however, experiments in progress which, if successful, could yield even more compact and less expensive devices than those achievable with a Marx generator.

Acknowledgments

I would like to acknowledge the invaluable assistance of Jack D. Graham and the other technicians and personnel of Maxwell Laboratories, Albuquerque Division as well as the Air Force personnel of the High Energy Plasma Division of Phillips lab who made this work possible.

References

1. MacGowan, B.J., et al. Phys. Fluids B 4 (7), July 1992.

2. Rocca, J.J., et al. Phys. Rev. E 47 (2), Feb. 1993.

Fig. 1 – Schematic of the Cronos Marx generator

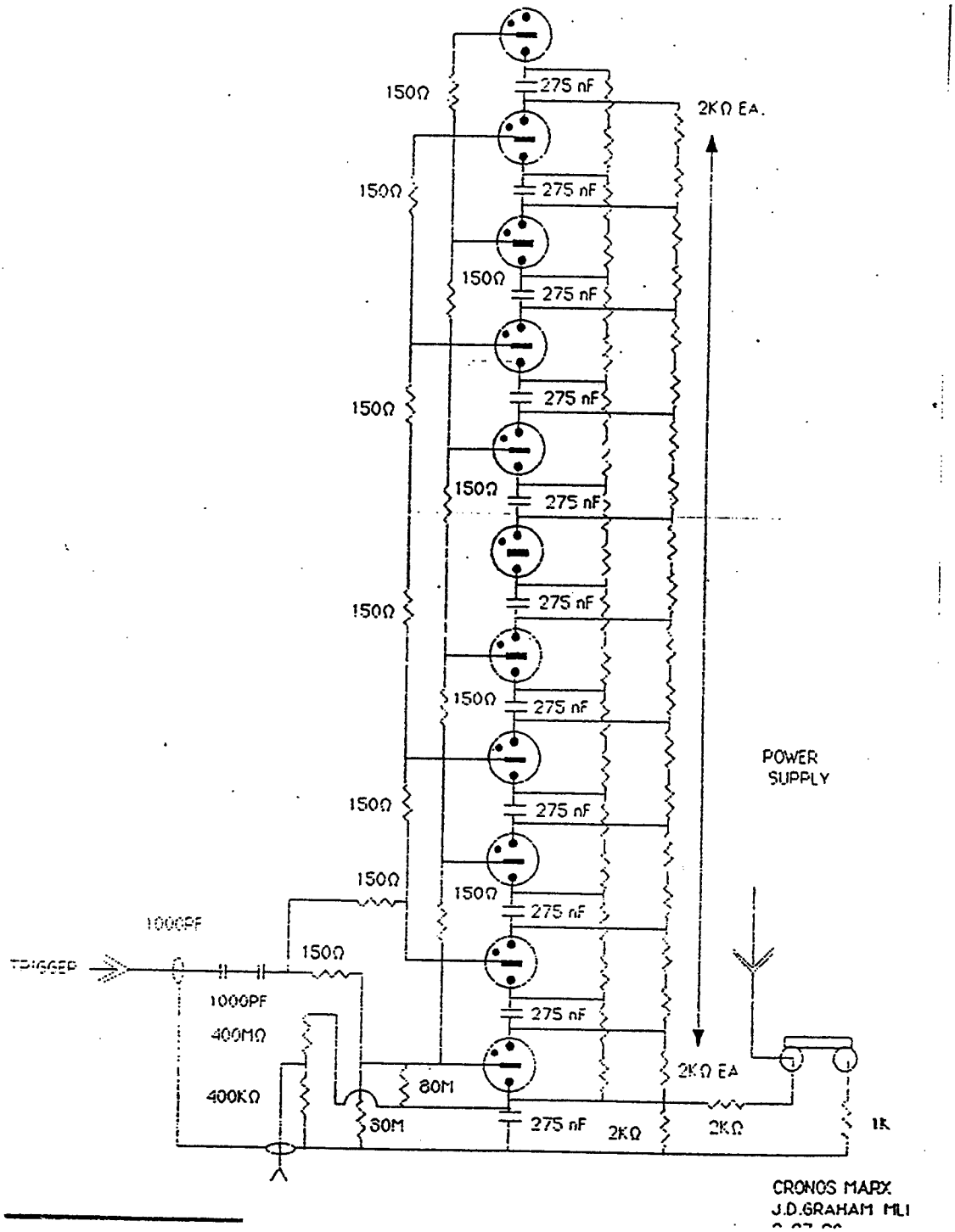


Fig. 2 – Capillary discharge device

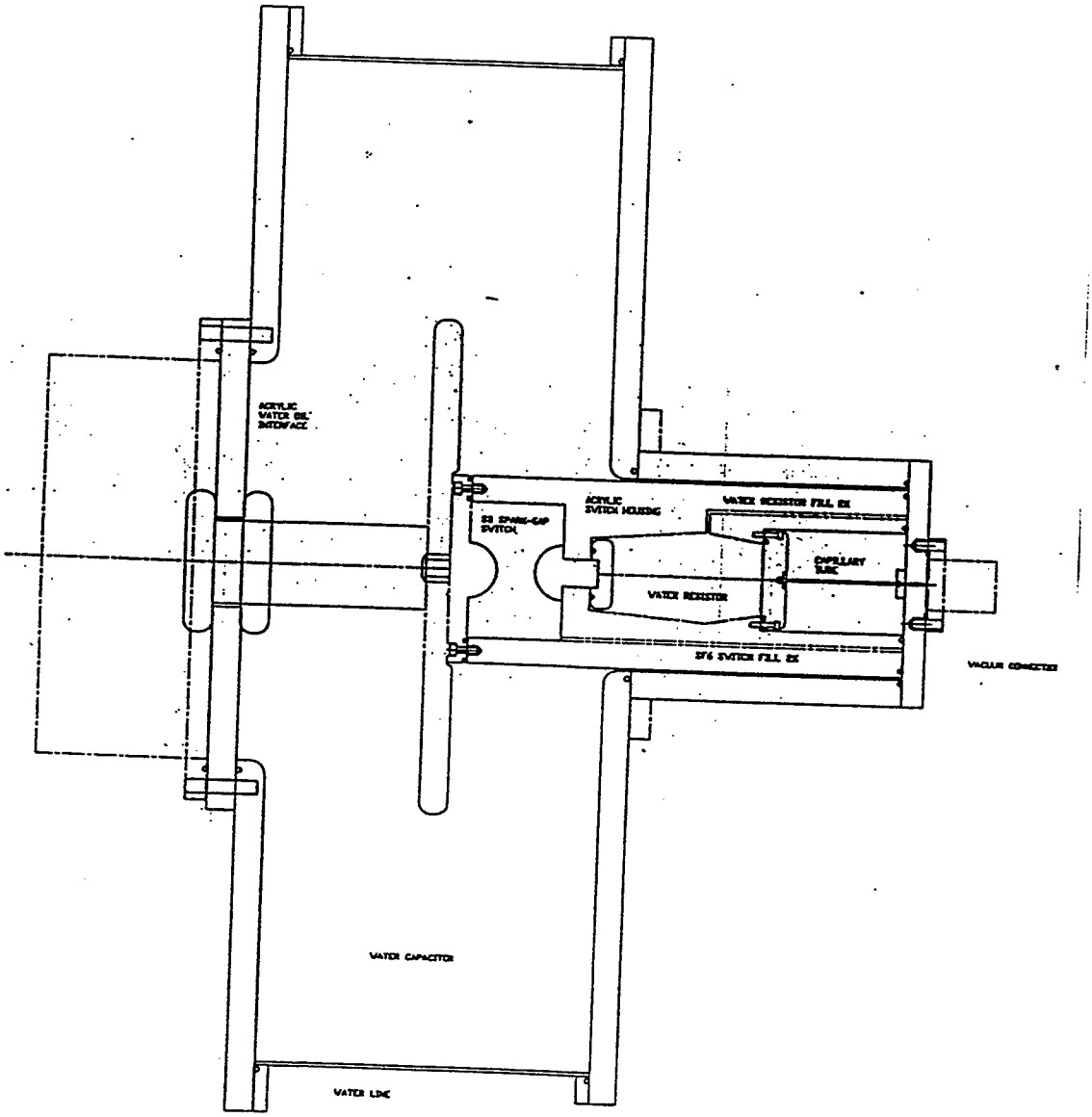


Fig. 3 – Computer simulations of the Marx discharge

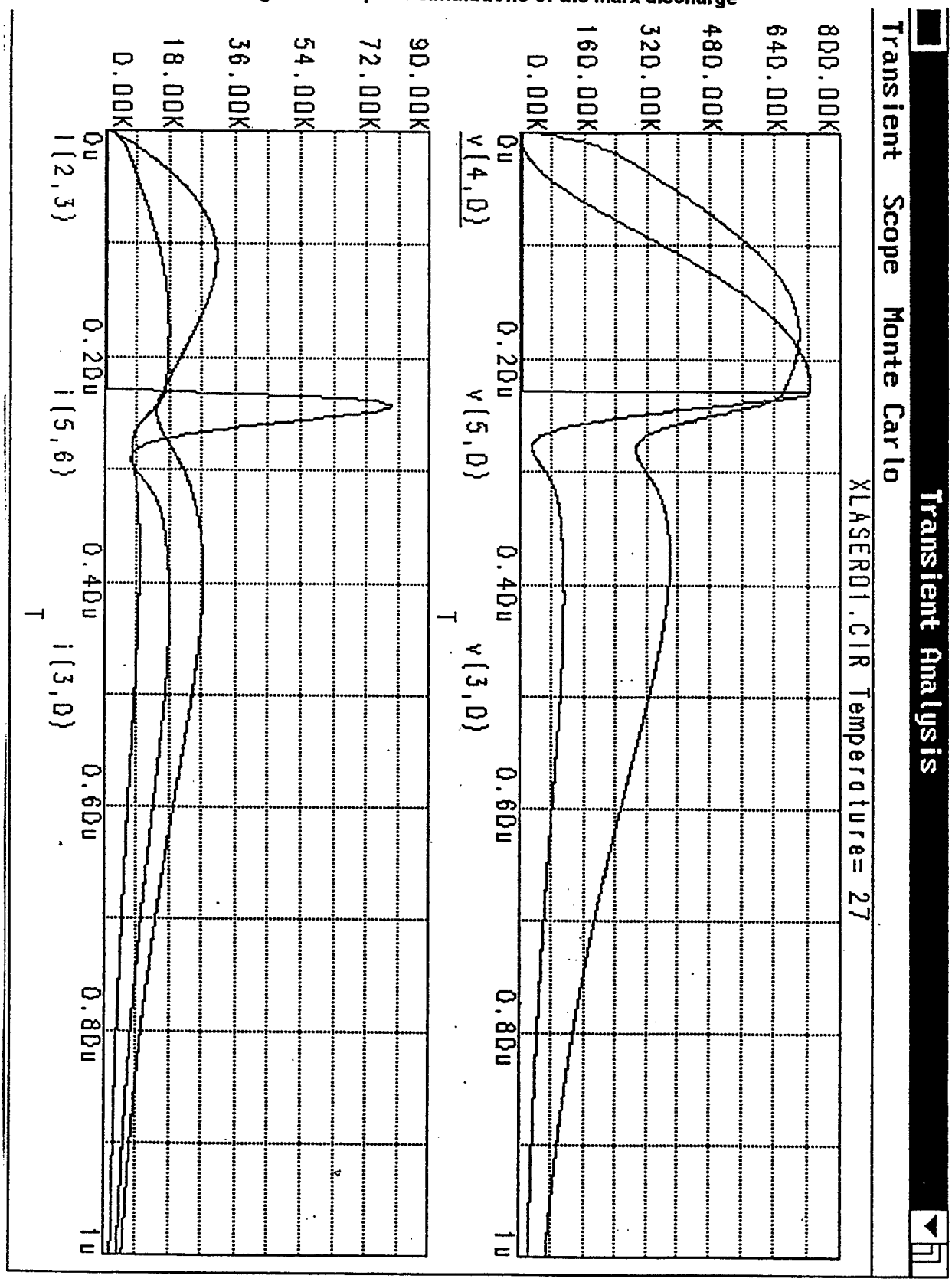
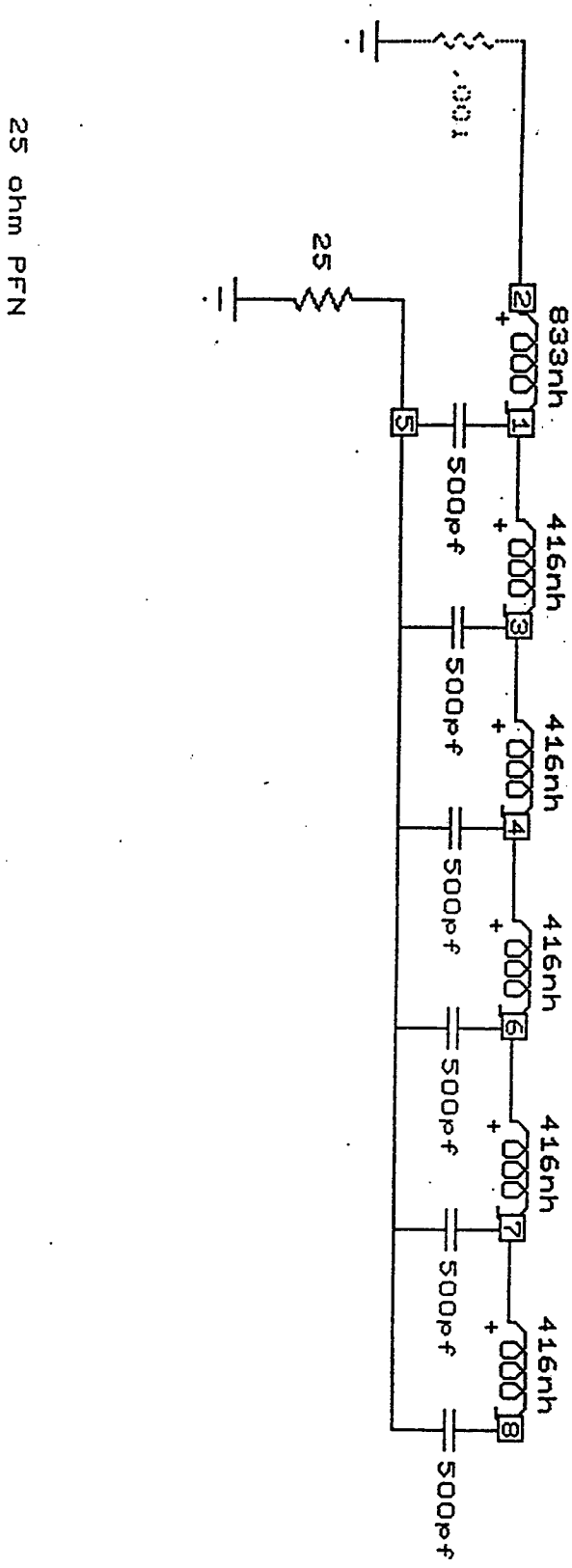


Fig. 4 – PFN Schematic



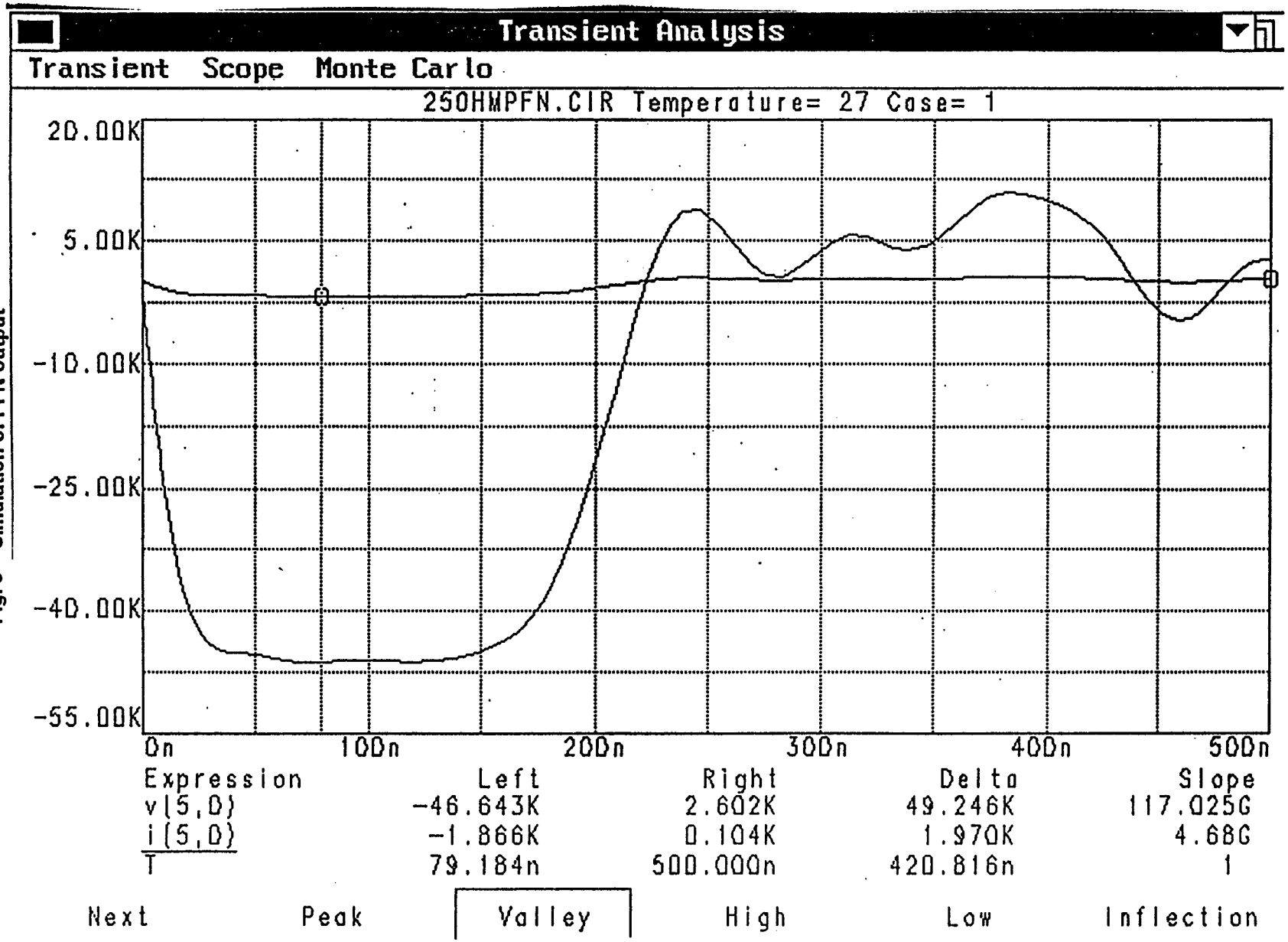


Fig. 5 – Simulation of PFN output

UNIDIRECTIONAL RING LASER WITH MULTIPLE QUANTUM WELL MEDIA

Paul Pulaski
Graduate Student
Jean-Claude Diels
Professor
Department of Physics and Astronomy
The University of New Mexico
Albuquerque, New Mexico 87131

Report for:
Faculty Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Phillips Laboratory, Albuquerque

September 1995

UNIDIRECTIONAL RING LASER WITH MULTIPLE QUANTUM WELL GAIN MEDIA

Paul Pulaski
Graduate Student
Jean-Claude Diels
Professor
Department of Physics and Astronomy
The University of New Mexico
Albuquerque, New Mexico 87131

Abstract

We have designed a multiple quantum well structure with asymmetric reflection properties. This structure will be the gain medium inside a ring laser forcing the lasing action to be unidirectional.

1 Introduction

The research we are working on is to desgn an optically non-reciprocal multiple quatum well (MQW) laser. The MQW sample is being grown at the University of New Mexico's Crystal Growth Facility using Metal Organic Chemical Vapor Deposition (MOCVD). The same structure will also be grown at Sandia National Labs using Molecular Beam Epitaxy (MBE).

The function of the MQW is to exhibit a zero reflection in one direction and a non-zero refelction in the other direction. We have shown that this device can be modeled using two thin (thickness $\ll$ wavelength) layers - a dielectric and a gain. By choosing the appropriate values for the reflection coefficient of the dielectric, the gain, and the separation between the layers, a light beam incident from one side of the device experiences a constructive interference in reflection, while a beam incident from the other side experiences a desructive interference in reflection, giving rise to its non-reciprocity. The transmission is the same in both directions. It is no longer non-reciprocal when the gain goes to zero. Computer calculations which model the effect of this device in an external ring cavity show that the steady state intensity in one direction is non-zero while in the other direction the intensity drops almost to zero. We have extended this idea to a practical device by designing an anti-symmetric MQW structure using the standard matrix formalism which calculates the total reflection and transmission as a function of index, thickness of the individual layers, gain and wavelength. By varying these parameters we are able to design a device which exhibits zero reflection in one direction and non-zero reflection in the other direction. The MQW uses GaAs for the gain layers (quantum wells) and AlGaAs for the spacers. The structure will be placed in an external ring cavity pumped with a Ti-Sapphire laser.

One application of this unidirectional laser is to construct a laser gyroscope. The idea is to have two of these lasers placed one on top of the other so they are decoupled from each other but have the same optical path length. One can grow

a MQW structure on the hypotenuse of a prism made of some semiconductor material that is transparent to the laser light. This can be optically contacted to the hypotenuse of another prism making a small cube. Any cross section perpendicular to the MQW makes an effective "cavity" with the light being internally reflected off the sides of the cube. The position of these cavities and the direction of light propagation depends on where the MQW is being pumped. If it is pumped at two different heights then two cavities can be made where the light travels in counterpropagating directions. Since the cavities are decoupled there will be no backscattered radiation from one cavity into the other, which is a major problem with the conventional laser gyros.

2 The Unidirectional Element

A single quantum well has a thickness d much smaller than the wavelength. Let us consider an inverted single quantum well. An electric field of amplitude $\tilde{\mathcal{E}}_i$ incident from the left, will, after transmission through the layer, become a transmitted field $\tilde{\mathcal{E}}_t = \tilde{\mathcal{E}}_i/(1-\alpha)$. The reflected field will be $\tilde{\mathcal{E}}_r = \alpha\tilde{\mathcal{E}}_i/(1-\alpha)$. The single amplifying quantum well can also be represented by a layer of purely imaginary index of refraction $n_g = -i\kappa$.

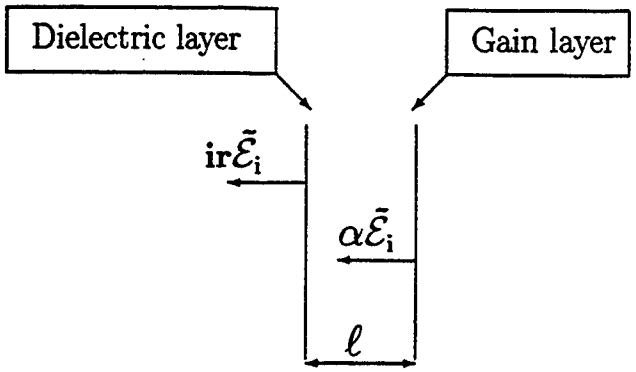
Let us consider next a thin dielectric layer, and the combined reflection from both interfaces. It can easily be shown by solving the wave equation for a monochromatic plane wave incident on the structure that, in the limit of zero thickness (thickness negligible compared to the wavelength), for an incident field E_i, the reflected field is $E_r = irE_i/(1-ir)$, where r is a real quantity, proportional to the discontinuity in index Δn. This result is consistent with the previous one: for a gain medium, $\Delta n = -i\kappa$, and the reflected field should be of the form $E_r = \alpha E_i/(1-\alpha)$.

The pairing of a thin gain layer with a dielectric reflecting layer leads to a structure that is non reciprocal. Let us consider, from left to right, a thin gain

layer (transmission $E_t = E_i/(1-\alpha)$; reflection $E_r = \alpha E_i/(1-\alpha)$) followed at a distance $\ell = 3\lambda/8$ by a dielectric layer (reflection $E_r = irE_i/(1-ir)$; transmission $E_t = E_i/(1-ir)$). The round trip between the two layers corresponds to a phase factor $\exp\{-2ik\ell\} = i$. Let us neglect for simplicity multiple reflections. The combined reflection from both interfaces, for a beam incident from the left, is: $E_r = (\alpha - r)E_i$, which is zero if $\alpha = r$. For a beam incident from the right: $E_r = i(\alpha + r)E_i$, which is finite if $\alpha = r$. (Fig. 1) The transmission is the same for both directions.

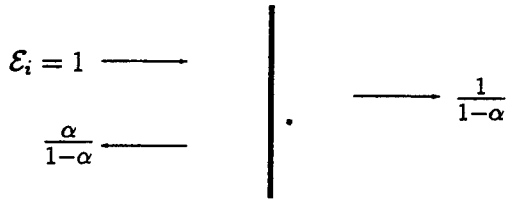
It is important for gyro related applications that this finite reflectivity has a phase factor of $\pi/2$ with respect to the transmission from left to right. Other cases are possible. For instance, if the spacing between the layers is $\ell = \lambda/8$, then it is the reflection from the right that can be made equal to zero. But the non-zero reflection from the left is then in phase (or π out of phase) with the transmission from the right, which is an undesirable situation for laser gyro applications.

At first glance one would think that this device does not conserve energy because the transmission is the same in both directions while the reflectivity is different. So the question is where does the energy go when the reflection is zero for the one direction? To answer this we need to look at what happens at the gain layer. The gain experienced at the gain layer is equal to some constant times the local field squared. It turns out that the local field at the gain layer is not the same depending on which direction the incident field is coming from. The local field at the gain layer can be calculated by summing the partial fields at the position of the gain layer. Then one subtracts the gain at the layer when coming from the left and the right. This value is equal to the non-zero reflection since the transmission is the same. Therefore this structure does conserve energy.(Fig. 2)



$r = \alpha$ and	FROM LEFT TO RIGHT $(\mathbf{i\,r} + \alpha\mathbf{e}^{-2\mathbf{ik}\ell})\tilde{\mathcal{E}}_1$	FROM RIGHT TO LEFT $(\alpha + \mathbf{i\,re}^{-2\mathbf{ik}\ell})\tilde{\mathcal{E}}_2$
$2k\ell = \frac{\pi}{2}$	$\mathbf{0}$	$\mathbf{2}\alpha\tilde{\mathcal{E}}_2$
$2k\ell = \frac{3\pi}{2}$	$\mathbf{2i}\alpha\tilde{\mathcal{E}}_1$	$\mathbf{0}$

Figure 1:Non-reciprocal field reflection from a combination of a thin dielectric and a gain layer.



$$GAIN = g|\mathcal{E}\alpha|^2$$

$$GAIN = R + T - 1$$

$$1 + GAIN_L = R_L + T_L$$
$$1 + GAIN_R = R_R + T_R$$

$$R_R = 0 \Longrightarrow GAIN_L - GAIN_R = R_L$$

$$g(|\mathcal{E}_\alpha|_L^2 - |\mathcal{E}_\alpha|_R^2) = R_L$$

Figure 2:Energy conservation

3 Real MQW Structure

The standard matrix formalism for multilayered structures is used to calculate the reflection and transmission as a function of wavelength. The total matrix for the MQW structure is constructed by mutiplying a series of interface and propagation matrices for each layer. The interface matrix is a function of the field reflection and transmission coefficients for the boundaries between the individual layers. The propagation matrix is a function of the phase factor which depends on the thickness and the complex index of refraction. The reflection and transmission coefficients for the whole structure can be calculated from the total matrix. It is important to note that the transmission is always the same in both directions. The reflection, however can be different if gain is included in some of the layers. The gain is represented by a negative imaginary index and for this particular structure the gain is in the GaAs quantum wells.

Since the MQW structure will be place in a ring cavity, we designed an anti-reflection (AR) coating stack on both sides of the quantum wells. The first layer is a half wavelength thick followed by a series of quarter wavelength layers. Between the quantum wells and the AR stack are layers which are critical for insuring unidirectionality. These layers are asymmetric with respect to the position of the quantum wells. The resonant wavelength for this structure, which is determined by the thickness of the quantum wells (10 nm), is 858 nm. In between the quantum wells are spacers such that the distance between two adjacent quantum wells is a half wavelength. Since the thickness of the quantum wells and the AR stack are fixed the adjustable parameters are the thickness, composition and the number of layers in between. Once the compostion and number of layers are determined it is still a formidable task to adjust the relative thickness of these layers so that we obtain a zero reflection in one direction and a $\pi/2$ phase shift of the non zero reflection with respect to the transmission at the resonant wavelength (Table I).

MQW structure

Table I

material	thickness (nm)	index (r)	index (i)	comment
air		1.	0.000	boundary
$Al_{0.25}Ga_{0.75}As$	126.30	3.3967	0.000	stop etch layer
$AlAs$	71.90	2.9835	0.000	
$Al_{0.25}Ga_{0.75}As$	63.15	3.3967	0.000	
$AlAs$	71.90	2.9835	0.000	AR
$Al_{0.25}Ga_{0.75}As$	63.15	3.3967	0.000	
$AlAs$	71.90	2.9835	0.000	AR
$Al_{0.25}Ga_{0.75}As$	63.15	3.3967	0.000	
$AlAs$	71.90	2.9835	0.000	AR
$Al_{0.25}Ga_{0.75}As$	35.00	3.3967	0.000	reflector
$AlAs$	116.0	2.9835	0.000	reflector
$Al_{0.25}Ga_{0.75}As$	94.00	3.3967	0.000	spacer $r - \alpha$
$AlAs$	36.90	2.9835	0.000	(2 layers)
$Al_{0.25}Ga_{0.75}As$	94.00	3.3967	0.000	spacer $r - \alpha$
$AlAs$	36.90	2.9835	0.000	(2 layers)
$Al_{0.25}Ga_{0.75}As$	118.38	3.3967	0.000	spacer $r - \alpha$
$GaAs$	10.00	3.6470	-0.014	Quantum well
$Ga_{0.80}Al_{0.20}As$	115.00	3.4130	0.000	pump layer
$GaAs$	10.00	3.6470	-0.014	Quantum well
$Ga_{0.80}Al_{0.20}As$	115.00	3.4130	0.000	pump layer
...	...	...	...	19 pairs total between ——
$GaAs$	10.00	3.6470	-0.014	Quantum well
$Ga_{0.80}Al_{0.20}As$	115.00	3.4130	0.000	pump layer
$GaAs$	10.00	3.6470	-0.014	Quantum well
$Ga_{0.80}Al_{0.20}As$	115.00	3.4130	0.000	pump layer
$GaAs$	10.00	3.6470	-0.014	last Quantum well
$Al_{0.25}Ga_{0.75}As$	118.38	3.3967	0.000	spacer
$AlAs$	36.90	2.9835	0.000	(1 layer)
$Al_{0.25}Ga_{0.75}As$	94.00	3.3967	0.000	spacer
$AlAs$	36.90	2.9835	0.000	(1 layer)
$Al_{0.25}Ga_{0.75}As$	124.00	3.3967	0.000	reflector
$AlAs$	71.90	2.9835	0.000	reflector
$Al_{0.25}Ga_{0.75}As$	63.15	3.3967	0.000	AR
$AlAs$	71.90	2.9835	0.000	AR
$Al_{0.25}Ga_{0.75}As$	63.15	3.3967	0.000	AR
$AlAs$	71.90	2.9835	0.000	AR
$Al_{0.25}Ga_{0.75}As$	63.15	3.3967	0.000	AR
$AlAs$	71.90	2.9835	0.000	AR
$Al_{0.25}Ga_{0.75}As$	1389.27	3.3967	0.000	etch layer
$AlAs$	10.00	2.9835	0.000	1st etch
$GaAs$		3.6470	0.000	substrate

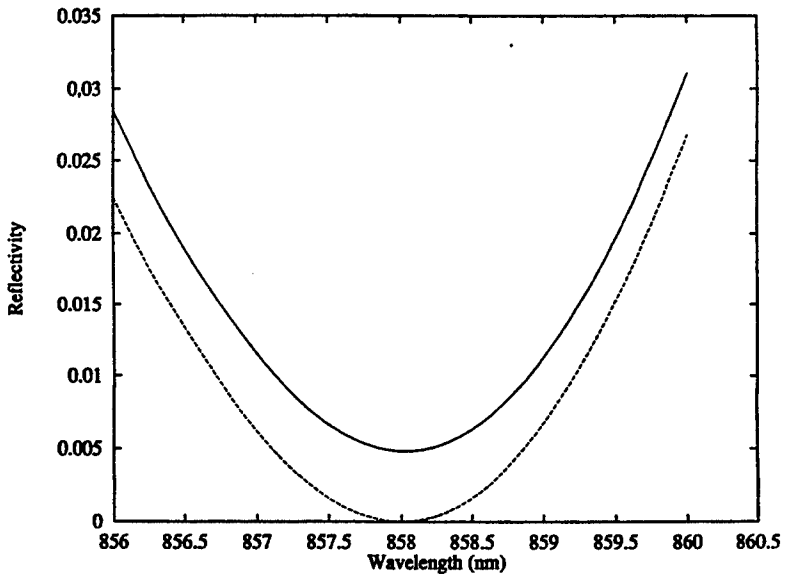


Figure 3:Reflectivity as a function of wavelength for both directions.

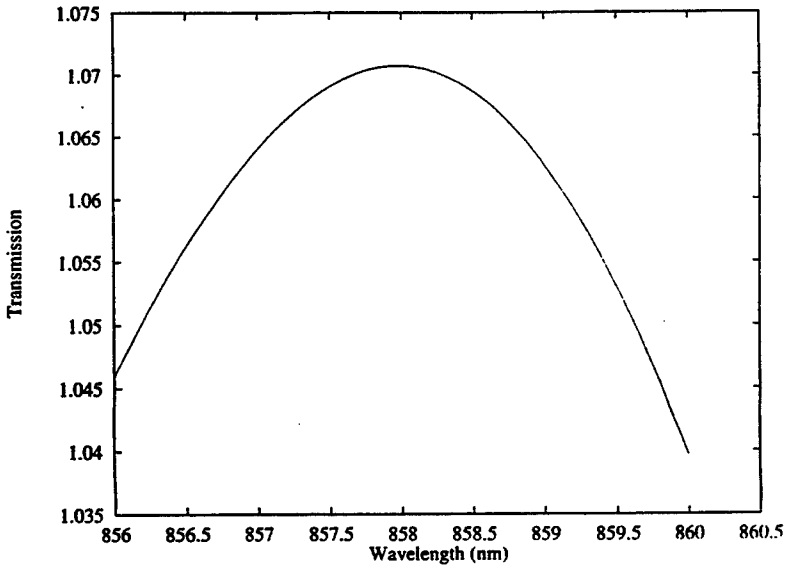


Figure 4:Transmission as a function of wavelength.

4 Etching

The MQW is grown on top of a thick GaAs substrate. In order to prepare the sample we need to etch a hole through this substrate so that it can be placed inside a ring cavity. We are using a wet chemical etchant to etch away the substrate. Since the substrate is fairly thick the first step is to polish the sample down then do the wet etch.

A small piece of the sample (about a 1 cm square) is cleaved from the 2 inch wafer. This piece is then waxed onto a small steel block with four microscope cover slides surrounding it. The thickness of the slides is about 100 microns. A smooth glass plate is used as the polishing surface. An alumina micropolish is mixed with a little bit of distilled water until mixture is a milky liquid that flows easily. Starting with a 1 micron grit the sample is polished using a figure eight motion until the sample is almost the thickness of the cover slides. A 0.3 micron grit is then used to finish the polishing.

Next the polished piece is waxed to a regular microscope slide leaving a very small hole (1 mm diameter) for the chemical etch. We use a fast etch to start and a selective slow etch as we get close to the MQW structure. The fast etchant is made with phosphoric acid/hydrogen peroxide/water with ratios 1:2:2. The main problem with this etchant is that the etch rate decreases rapidly which makes it difficult to calculate the time needed to etch down close to the MQW. If the sample stays in the solution too long, it will etch right through the sample, in which case you need to start over with a new hole. The intial etch rate is about 4.5 microns/minute and decreases to about 3 microns/minute in 30 minutes. These rates change depending on how fresh the hydrogen peroxide is when the mixture is prepared. A new bottle yields the fastest etch rate. The depth of the etched hole is measured with a mechanical stylus profilometer. Every 5 to 10

minutes I would check the depth to keep track of how close the etched surface was to the structure. A stop etch layer (10 nm AlAs) was grown between the stucture and the substrate. This was provided so we could use a selective slow etch once we got close to structure.

The formula for the slow etch consists of citric acid/hydrogen peroxide in a ratio of 10:1. The etch rate for this mixture is about 0.2 microns/minute for GaAs and .002 microns/minute for AlAs which gives a selectivity of 100:1. Compared to other etchants, this one gives better selectivity and smoother etched surfaces.

5 Results

We had recieved a sample grown here at UNM with the MOCVD machine a few months ago. There were a few problems with that sample. First, after the sample was etched it started to bow. There are two ways to deal with this. First we can make the etched hole smaller which will reduce the strain on the etched surface. Also, since the MQW structure is only about 4 microns thick, we decided to increase this thickness by adding a few microns to one of the end layers such that is does not change the non-reciprocal property of the device. Increasing the thickness will also help to dissipate the heat generated when this sample is pumped with a focused laser beam. Heat sinking may be a major concern since we will have to be careful not to melt the MQW. The etched sample will be mounted on a sapphire window that has a hole drilled in it to allow the light to pass through. We may also place this on TE cooler to help heat sink the device even more.

One of the major problems with MOCVD is that it is not nearly as precise as the MBE process. It is difficult to control the thickness of the individual layers with MOCVD, whereas the MBE process can monitor the thickness at each layer. We did a series of computer calculations for the reflection of the

sample as the thickness is changed. We were told that all of the AlGaAs layers would change by about the same amount when growing this sample with MBE. So we changed the thickness of the AlGaAs layers with respect to the GaAs layers. We found that the general behavior of the reflection in both directions does not change that much when the thicknesses were changed between plus and minus 1%, however the minimum reflection which occurs at the resonant wavelength shifts does shift by as much as 10 nm.

Once the sample is prepared, it will be placed inside a ring cavity and pumped with a Ti-Sapphire laser (Fig. 5). The pump beam will be on axis with ring cavity beam path and focused onto the sample with short focal length diode lenses. One of the mirrors will be an output coupler so that we can detect the counterpropagating beams from the ring cavity laser.

6 The Laser Cube

One application of this unidirectional structure is to construct a laser gyroscope with two oppositely directed unidirectional lasers, one on top of the other. This will eliminate the main nemesis of the laser gyroscope known as the lock-in effect which is caused by backscattered radiation off the cavity optics, coupling from one mode into the counterpropagating mode. Since the two lasers are separated spatially, there will be no chance for this backscattered radiation from one laser to affect the other laser. The bi-level laser is a cube which consists of two GaAs 90^0 turning prisms. A nonreciprocal gain layer structure is grown on the hypotenuse face of one prism ($InGaAs$). The two prisms are thereafter optically contacted to form a cube. Depending of the location of the pump spot (pump radiation around 900 nm), the lasing will be clock-wise or counterclock-wise in the cross-section of the cube. With two pump spots at different height in the prisms, lasing can occur in opposite direction in adjacent planes. Since there is no coupling between the cavities, the beat frequency between the two lasing

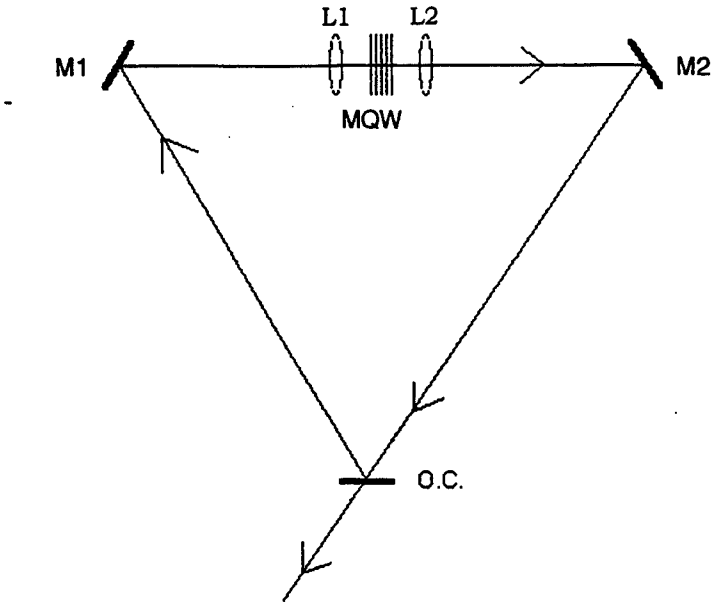


Figure 5:Ring laser cavity configuration

planes should exhibit a perfect gyroscopic response.

7 Conclusion

We are able to model a unidirectional device with a gain and dielectric layer which exhibits a zero reflection in on direction and a non-zero reflection in the other direction. This can be extended to a real MQW structure which is designed by using the standard matrix formalism for determining the reflection and transmission of a multilayered structure. Once we obtain the sample, it will have to be prepared by polishing and etching away the substrate before it is placed inside a ring resonator. The samples will be grown using MOCVD here at UNM and MBE at Sandia National Labs here in Albuquerque. It is our guess that the MBE process will yield a higher quality sample.

Associate did not participate in program.

Turbulent Atmosphere Modeling and Atmospheric Short Exposure OTF Evaluation

Kevin Scales
Graduate Student
Department of Physics

University of New Mexico
800 Yale NE
Albuquerque, NM 87131

Final Report for:
Graduate Student Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Laboratory

July 1995

Turbulent Atmosphere Modeling and Atmospheric Short Exposure OTF Evaluation

Kevin Scales
Graduate Student
Department of Physics
University of New Mexico

Abstract

Modeling of propagation through a turbulent atmosphere was studied by evaluating a solution for the Helmholtz equation. Refractive index variations were modeled in a three dimensional manner unlike the common phase screen approach. An evaluation of the short exposure atmospheric OTF is evaluated numerically and compared to Fried's approximation.

Kevin Scales

Introduction

Two projects were worked on during the summer of 1995 at Phillips Lab. The first involved the creation of a wavefront generator using known properties of atmospheric turbulence. By generating a statistically appropriate model of refractive index variations, it then becomes possible to model the propagation of optical waves through such a medium.

Also relating to the optical transmission, the exact frame-averaged optical transfer function for a turbulent medium was computed and compared to the commonly used approximation. The assumption that Gaussian statistics are valid for atmospheric modeling was used throughout.

Atmospheric Index Modeling — Background

The index of refraction for the atmosphere is written as the sum of a non-varying constant plus a zero mean fluctuating value [1]:

$$n(\mathbf{r}) = n_0(\mathbf{r}) + n_1(\mathbf{r})$$

The spatial autocorrelation function for a homogeneous n_1 is defined as:

$$\Gamma_n(\mathbf{r}) = E[n_1(\mathbf{r}_1)n_1(\mathbf{r}_1 - \mathbf{r})]$$

The three dimensional fourier transform of $\Gamma_n(\mathbf{r})$ is the power spectral density:

$$\Phi_n(\kappa) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \int \Gamma_n(\mathbf{r}) e^{j\kappa \cdot r} d^3\mathbf{r}$$

We have defined $\kappa = (\kappa_X, \kappa_Y, \kappa_Z)$ as a vector spatial frequency. According to the often used Kolmogorov spectrum,

$$\Phi_n(\kappa) = 0.033 C_n^2 \kappa^{-11/3}$$

It is known [2] that values of $n_1(\kappa)$ and $n_1(\kappa')$ are uncorrelated. For this reason, it is acceptable to model the atmosphere in κ space by assigning random values, properly scaled, to each cell in κ space. This method allows one to calculate values for a three dimensional array of n_1 for a volume of atmosphere.

Propagation through the Atmosphere

There are two possible ways to model atmospheric propagation. One approach is done in Goodman [1], using a Green's function integral approach. The other is to solve the Helmholtz equation directly.

Integral equation

This approach in Goodman assumes that just as we can write the refractive index as the sum of a constant and a perturbative term, we can also write our field, $\mathbf{U}$ as the sum of a non-perturbed and a corrective term: $\mathbf{U} = \mathbf{U}_0 + \mathbf{U}_1$. The solution after travelling through a volume V is straightforwardly found to be:

$$\mathbf{U}_1(\mathbf{r}) = \frac{1}{4\pi} \int\int\int_V \frac{e^{jk_0|r-r'|}}{|\mathbf{r}-\mathbf{r}'|}[2k_0^2 n_1(\mathbf{r})\mathbf{U}_0(\mathbf{r}')]d^3\mathbf{r}'$$

Before one can use this, the question arises of whether it is safe to assume that higher order terms in $\mathbf{U}$ can in fact be safely neglected. Following a similar method but keeping our $\mathbf{U}_2$ terms, it can be shown that higher order terms contain significant phase differences. Since the calculations involved in keeping higher order terms quickly become lengthy, we shall use a solution to the Helmholtz equation instead.

Helmholtz equation solution

The Helmholtz equation is known as $(\nabla^2 + k^2)\mathbf{U}(\mathbf{r}) = 0$. Given that $k^2 = \frac{\omega^2 n^2}{c^2}$, $n^2 = {n_0}^2 + 2n_0n_1 + {n_1}^2$, and $\nabla_T{}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the transverse Laplacian, we can write the Helmholtz equation as:

$$(\nabla_T{}^2 + \frac{\partial^2}{\partial z^2})U(\rho,z) + \frac{\omega^2}{c^2}({n_0}^2 + 2n_0n_1 + {n_1}^2)U(\rho,z) = 0$$

Our next step is to write the propagating wave as the product of a plane wave and a slowly varying envelope. This can be expressed as:

$$U(\rho,z) = e^{i\frac{\omega}{c}n_0 z}u(\rho,z)$$

This can be easily plugged into the Helmholtz equation above. After cancelations and rearangement of terms, we are left with:

$$\frac{\partial}{\partial z}u(\rho,z)+\frac{1}{2ik_0}{\nabla_T}^2u(\rho,z)+\frac{k_0n_1}{in_0}u(\rho,z)=0$$

The alert reader will notice that this equation is of the same form as the Schrödinger equation. By changing t to z and choosing appropriate numbers for V, we can transform it mathematically to the Schrödinger equation. Then we are able to use known solutions to solve it. [3,4,5]

By setting $H=-{\nabla_T}^2+(\frac{-1}{2}n_1)$, we get:

$$u(\rho,z+\Delta z)=e^{-iHz}u(\rho,z)$$

The next step is to use this equation for as many layers of atmospheric n_1 values as we need. Considering that inner scale sizes are on the order of a few milimeters, we find that we require at least one hundred thousand layers for each kilometer traversed.

Some explanation is in order concerning the numerical approximations made. The transverse Laplacian requires the taking of second derivatives, a process that requires three data points, one at the location of interest and one on either side of it in each direction. This means that at each succeeding layer, the size of the two dimentional array of u values is reduced in size by one data point on each side.

Rather then lose all data, we can make some assumptions about our values of u on the edges after each propagation. A possible assumption is that the value of each "lost" data point takes on the value of the point directly across from it on the array, performing a wraparound effect. This would provide a hopefully accurate approximation about the behavior of the wave on the edges.

Another possibility is to assign values of u on the edges equal to the arithmetic mean of values over the area of the wave at a given layer.

Preliminary tests show that in both cases, for an arbitrary set of n_1 values, the results of wave propagation were virtually identical. This relates to the fact that values of n_1 have a very small variance, and thus cause very little wave perturbation at any given altitude.

The Optical Transfer Function

It is known that any aperture passes only spatial frequencies lower then a certain limit, found to be $\nu = \frac{d}{\lambda}$. How well it passes frequencies in that range is measured by the optical transfer function.

Basic Theory

In general, the optical transfer function for frequencies ν_U and ν_V is given as:

$$\mathbf{H}(\nu_U, \nu_V) = \frac{\int_{-\infty}^{\infty} \int \mathbf{P}(x,y)\mathbf{P}^*(x - \lambda\nu_U, y - \lambda\nu_V)\mathbf{t}_s(x,y)\mathbf{t}_s^*(x - \lambda\nu_U, y - \lambda\nu_V)dxdy}{\int_{-\infty}^{\infty} \int |\mathbf{P}(x,y)|^2 |\mathbf{t}_s(x,y)|^2 dxdy}$$

For propagation through a turbulent medium that follows the Kolmogorov spectrum, it can be shown (see Goodman [1]) that after removing tilt effects, the numerator of the OTF is written as:

$$\int_{-\infty}^{\infty} \int \mathbf{P}(x_1,y_2)\mathbf{P}(x_2,y_2) \times exp[(\chi_1 + \chi_2) + j(S_1 - a_X x_1 - a_Y y_1) - j(S_2 - a_X x_2 - a_Y y_2)]$$

where we have defined $(x_1, y_1) = (x, y)$ and $(x_2, y_2) = (x - \lambda\nu_U, y - \lambda\nu_V)$. Also, S_1 and S_2 are the phases at the two points of interest. a_X and a_y are defined to minimize the squared error:

$$\Delta = \int_{-\infty}^{\infty} \int P(x,y)[S(x,y) - (a_X x + a_Y y)]^2 dxdy$$

$$a_X = \frac{64}{\pi D_0^4} \int_{-\infty}^{\infty} \int xP(x,y)S(x,y)dxdy$$

$$a_Y = \frac{64}{\pi D_0^4} \int_{-\infty}^{\infty} \int yP(x,y)S(x,y)dxdy$$

To get a useful optical transfer function, we need to average this expression over many different frames of atmospheric turbulence. We need now to make two assumptions.

(1) We assume that $S(x,y)$, $a_X x$, and $a_Y y$ are uncorrelated, Gaussian random variables.

(2) We assume that $(S_1 - a_X x_1 - a_Y y_1) - (S_2 - a_X x_2 - a_Y y_2)$ is independent of $(\chi_1 + \chi_2)$

We can now reduce the problem to the evaluation of:

$$E\{exp[(\chi_1 + \chi_2) + j(S_1 - a_X x_1 - a_Y y_1) - j(S_2 - a_X x_2 - a_Y y_2)]\}$$

$$= exp\{-\frac{1}{2}D_\chi(\lambda\nu_U, \lambda\nu_V) - \frac{1}{2}\langle[(S_1 - a_X x_1 - a_Y y_1) - (S_2 - a_X x_2 - a_Y y_2)]^2\rangle\}$$

Evaluation of Atmospheric OTF

Now we are ready to compare the exact OTF under the assumptions above to the approximation by Fried [6] commonly used. The last averaged term above can be expanded out as:

$$\begin{aligned}&[(S_1 - a_X x_1 - a_Y y_1) - (S_2 - a_X x_2 - a_Y y_2)]^2 \\ &= \{(S_1 - S_2)^2 - (a_X \lambda\nu_U + a_Y \lambda\nu_V)^2 - 2[(S_1 - a_X x_1 - a_Y y_1) \\ &\quad -(S_2 - a_X x_2 - a_Y y_2)] \times [a_X \lambda\nu_U + a_Y \lambda\nu_V]\}\end{aligned}$$

If we now use assumption 1 from above, we can assume that the last term goes to zero. This is the assumption used in the literature which we shall now ignore.

The OTF depends on values of aperture diameter, which has been set at one meter throughout, and on r_0, the atmospheric coherence diameter. The expression for r_0 is related to C_n^2 by:

$$r_0 = 0.185[\frac{\lambda^2}{\int_0^z C_n^2(z)dz}]^{3/5}$$

For this derivation, C_n^2 has been assumed constant, and the height of the atmosphere has been taken to be 1 kilometer. Figures 1–4 illustrate the exact versus the approximate

calculation. The figures show that the two expressions remain close in value over many ranges of ν and D/r_0.

Some words are needed about what we are seeing. In the expression:

$$\mathbf{H}(\nu_U, \nu_V) = \mathbf{H}_0(\nu_U, \nu_V)\mathbf{H}_L(\nu_U, \nu_V)$$

which relates the average OTF of an imaging system to the product of the OTF of the atmosphere and the OTF of the system alone, we are seeing only the OTF of the atmosphere. In a real imaging system, the OTF will continue decreasing for all values of ν.

The question still remains of why the OTF goes back up for higher spatial frequencies. The answer is found in the removal of tilt that was performed early on. Tilt is simply the linear term that describes the linear shift of an image over an aperture. There are still higher order aberration terms present that are stronger for the middle ranges of spatial frequencies than the tilt is.

We see then that as tilt is removed, the remaining aberration terms tend to cause the OTF to drop accordingly. Since tilt accounts for roughly ninety percent of the wavefront aberration, its removal from center frequencies allows the OTF to return to a large value for high spatial frequencies.

Conclusions

In the field of atmospheric modeling, we still see that much work remains to be done. For numerical computations, a full three dimensional fourier transform to generate refractive index variations having the most accurate spatial correlation requires more memory than is available on most modern computers.

Consider that the smallest turbulent eddies are on the order of millimeters and the largest are on the order of tens of meters. For any reasonable propagation length, the number of data points becomes very high.

Thus, we find that approximations are required. At the time of this writing, such calculations have not been completed by the researchers involved in that portion of the study. It is hoped that a model can be developed using a model more accurate than the commonly used phase screen aproach in which only five to ten layers are evaluated.

The results of this work will be of aid to long baseline optical interferometry. Because of the importance of phase information, a more accurate model may be of aid in predicting the performance of future interferometers.

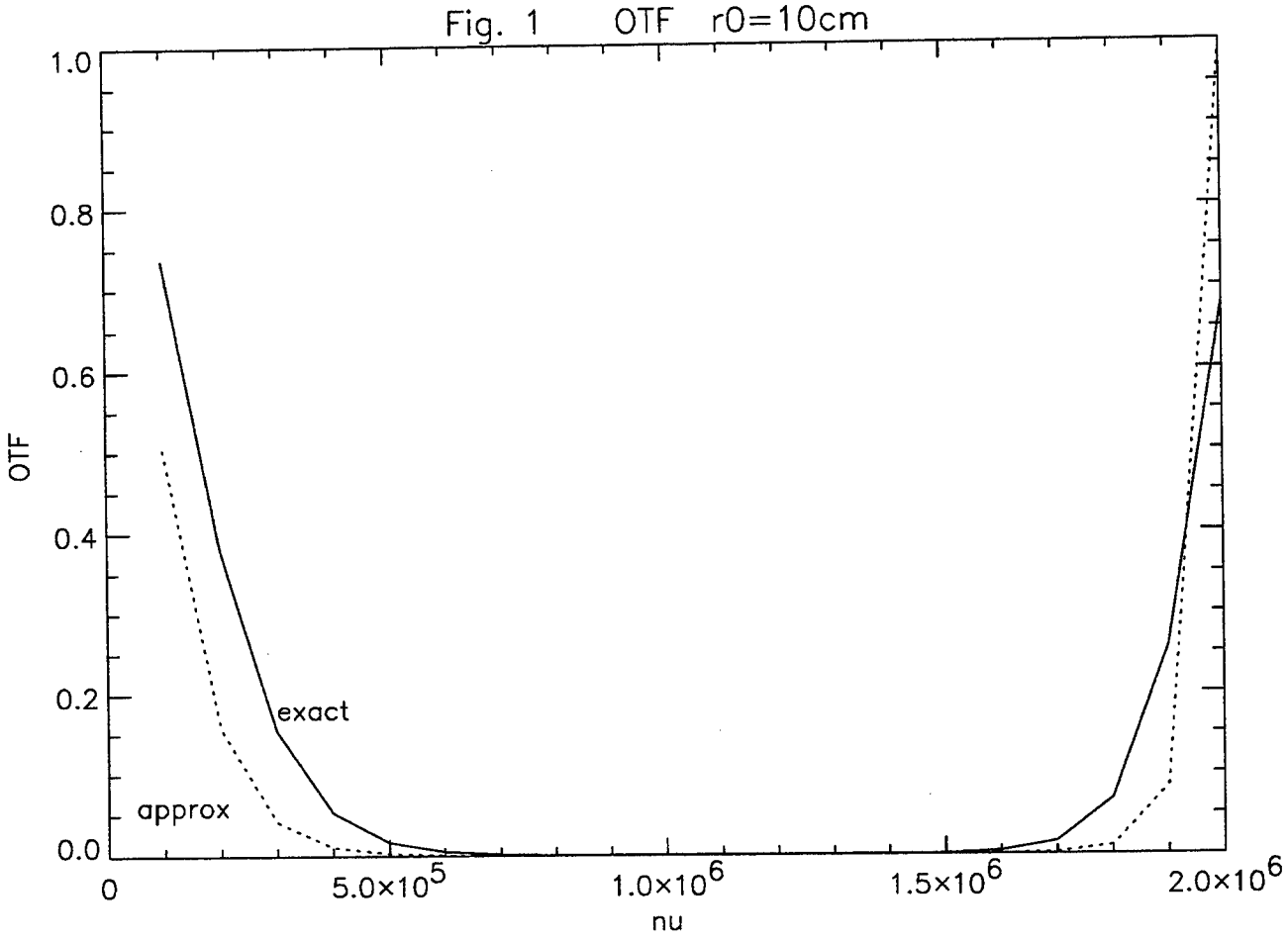


Fig. 1 OTF r0=10cm
1.0
0.8
0.6
0.4
0.2
0.0
OTF
exact
approx
0
5.0x10^5
1.0x10^6
1.5x10^6
2.0x10^6
nu

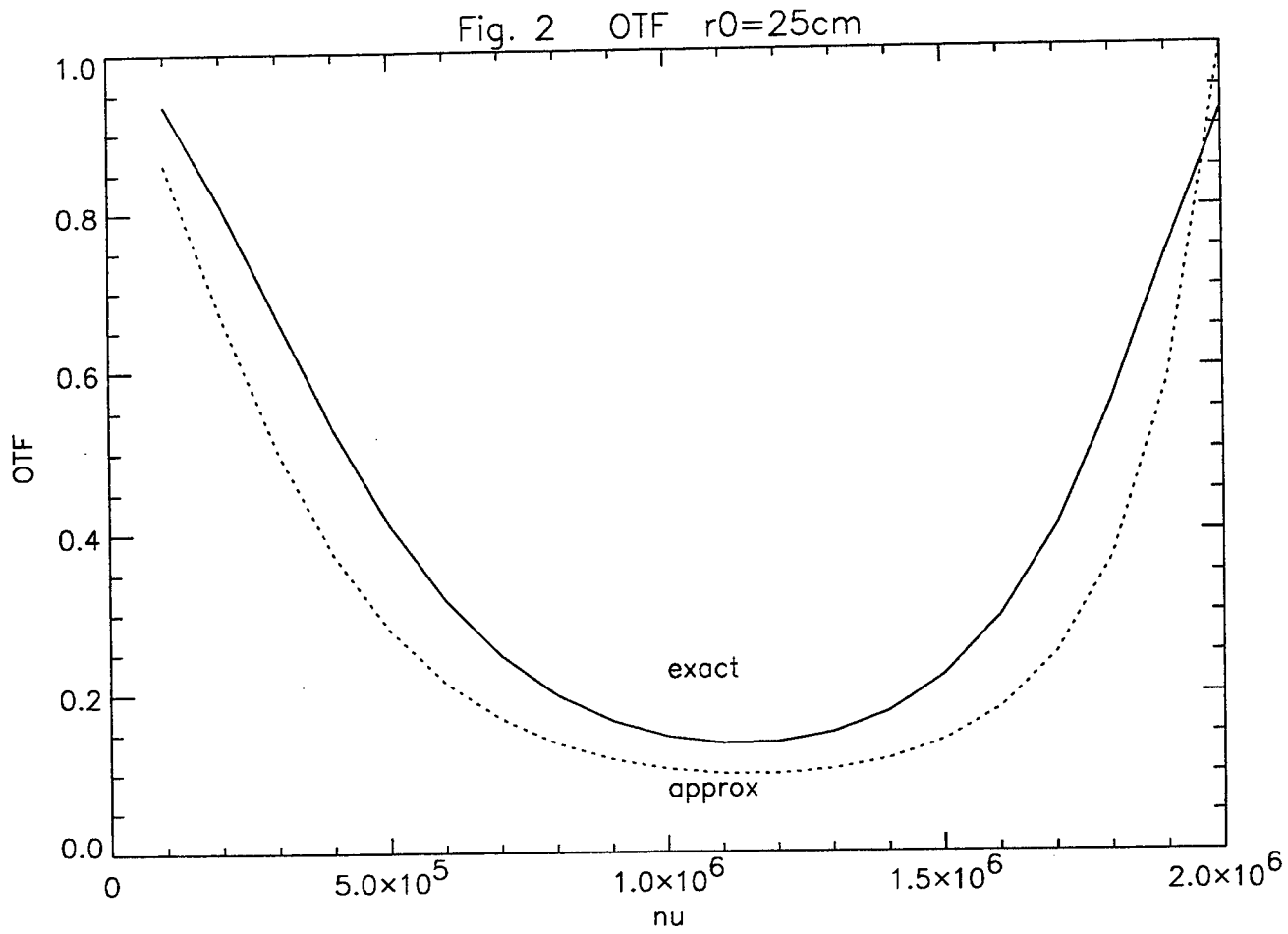


Fig. 2 OTF r0=25cm
1.0
0.8
0.6
0.4
0.2
0.0
OTF
exact
approx
0
5.0×10^5
1.0×10^6
1.5×10^6
2.0×10^6
nu

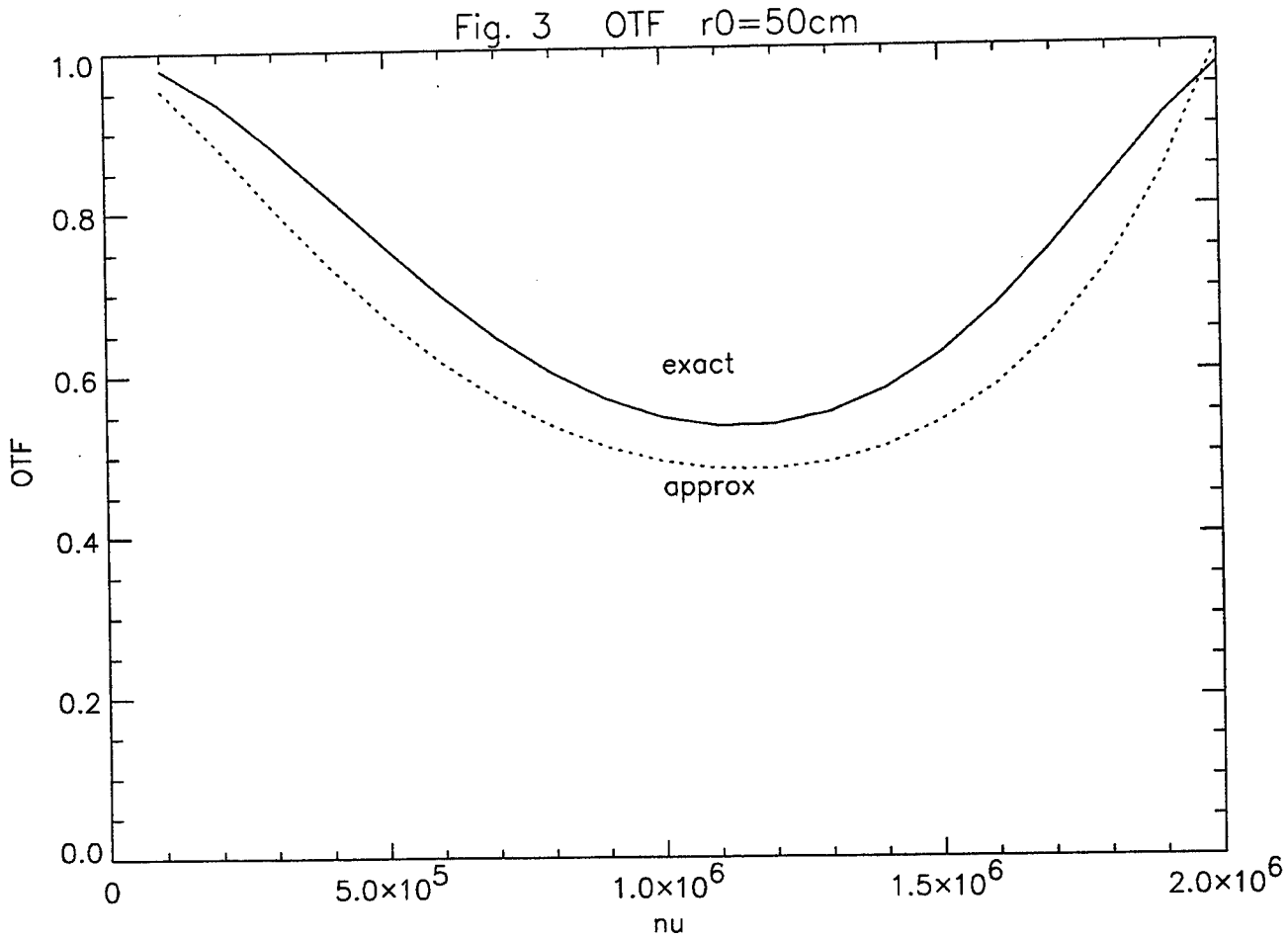


Fig. 3 OTF r0=50cm
1.0
0.8
0.6
0.4
0.2
0.0
OTF
exact
approx
0
5.0×10^5
1.0×10^6
1.5×10^6
2.0×10^6
nu

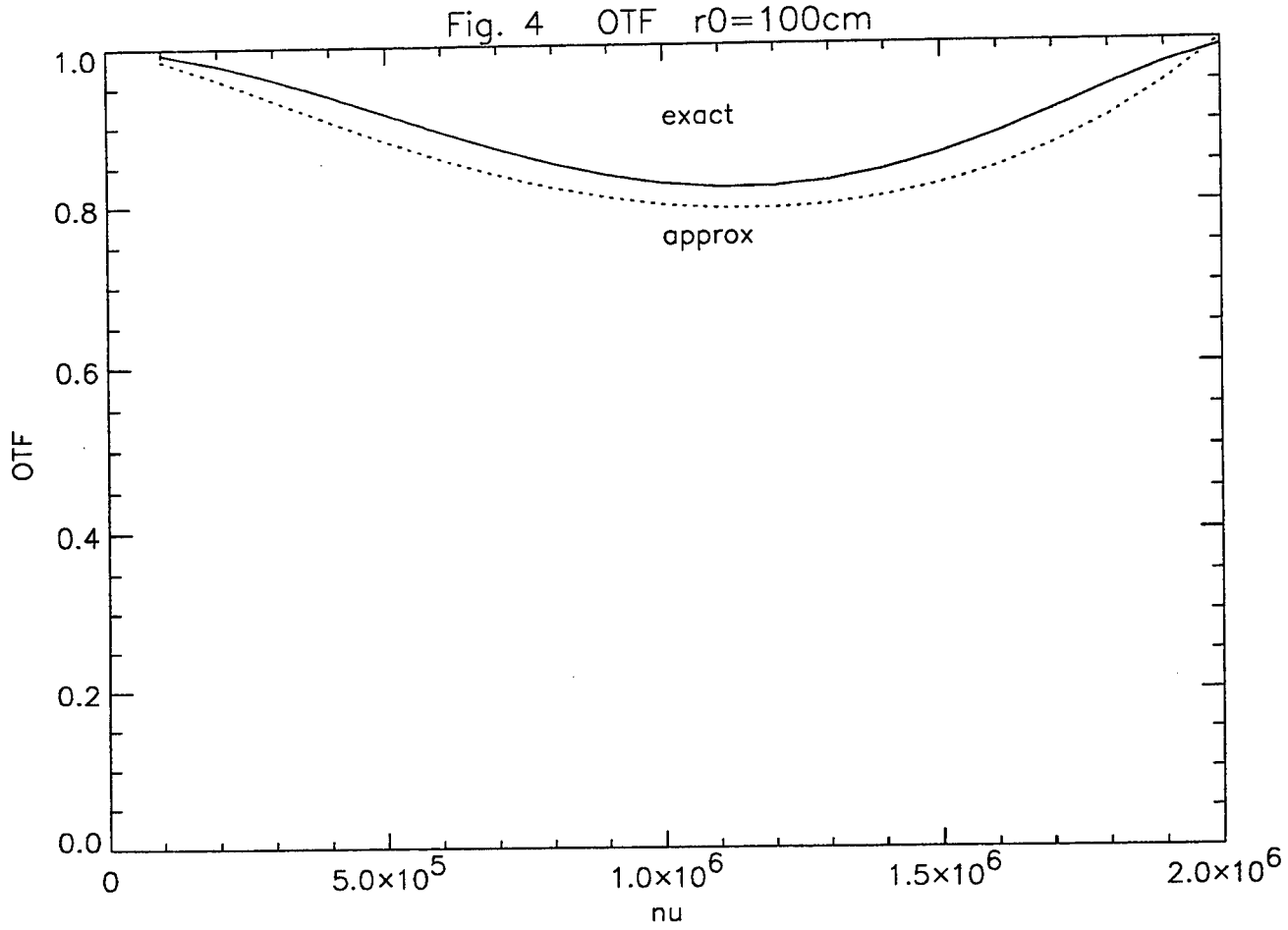


Fig. 4 OTF r0=100cm
1.0
0.8
0.6
0.4
0.2
0.0
OTF
exact
approx
0
5.0x10^5
1.0x10^6
1.5x10^6
2.0x10^6
nu

Bibliography

[1] J. Goodman, *Statistical Optics*, John Wiley & Sons, New York, 1985.

[2] J.W. Strohbehn, "Optical Propagation Through the Turbulent Atmosphere," *Progress in Optics*, Vol. IX (E. Wolf, editor), North Holland Publishing Amsterdam (1971).

[3] A. Goldberg, H.M. Schey, J.L. Schwartz, *Am. J. Phys.*, **35**, 177 (1967).

[4] I. Galbraith, Y.S. Ching, E. Abraham, *Am. J. Phys.*, **52**, 60 (1984).

[5] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes*, Cambridge University Press, New York, 1986.

[6] D.L. Fried, *J. Opt. Soc. Am.*, **56** 1372 (1966).

Associate did not participate in program.

SHIFTING OF BIFURCATION POINTS IN A DRIVEN Nd FIBER LASER BY SUBHARMONIC PERTURBATIONS

David W. Sukow
Graduate Research Assistant
Department of Physics

Duke University
Box 90305
Durham, NC 27708-0305

Final Report for:
Graduate Student Research Program
Phillips Laboratory, Kirtland AFB

Sponsored by:
Air Force Office of Scientific Research
Bolling AFB, Washington DC
and
Phillips Laboratory

August 1995

SHIFTING OF BIFURCATION POINTS IN A DRIVEN Nd FIBER LASER BY SUBHARMONIC PERTURBATIONS

David W. Sukow
Graduate Research Assistant
Department of Physics
Duke University

Abstract

We experimentally investigate the nonlinear dynamical behavior of a neodymium-doped fiber laser with a modulation applied to the pump. A perturbation is then added to the modulation. This perturbation is resonant with one of the subharmonics of the system. We observe shifting of the bifurcation point of the orbit having the same frequency as the applied perturbation. The amount and the direction of the shift is shown to depend on the strength of the perturbation and the relative phase between the perturbation and the master driver.

SHIFTING OF BIFURCATION POINTS IN A DRIVEN Nd FIBER LASER BY SUBHARMONIC PERTURBATIONS

David W. Sukow

Introduction

The study of nonlinear dynamics has given the scientific community many tools for understanding systems which display complex behaviors. Because the presence of irregularity in many applications is detrimental to performance, much effort has been devoted to learning how to manipulate the instabilities which characterize such systems. These efforts have generally involved the application of some sort of perturbation to a system through an accessible parameter. These perturbations can be considered in two classes: methods based on feedback, and open-loop methods. By application of these techniques, it has been demonstrated that the dynamics of a system can be greatly altered; instabilities can be suppressed or induced. In addition to the usefulness of stabilizing systems, this sort of study yields very basic information about the nature of the instabilities themselves.

We have chosen to pursue this type of experiment on a fiber laser, a system which displays a rich variety of nonlinear dynamical behavior. The driven Nd fiber laser has been studied experimentally and theoretically [1-3], and shows period doubling bifurcations, quasiperiodicity, chaos, coexisting attractors, and hysteresis. It also displays antiphase dynamics and has been successfully modeled as two coupled class-B lasers [3]. Similar investigations have been made on an erbium-doped fiber laser [4,5]. The *undriven* Nd fiber system is also known to spontaneously destabilize its steady state and undergo a period doubling route to chaos. The unstable periodic orbits and steady states of this autonomous system have been studied using feedback control techniques [6].

Experimental system

A schematic diagram of our experimental setup is shown in Figure 1. The active medium of the fiber laser system is a silica fiber doped with neodymium ions (manufactured by

Fibercore). The fiber is 3m long and has a core diameter of 3.5 μm. Its lasing wavelength is λ=1088 nm. It is strongly absorbing at suitable pump wavelengths (8.5 dB/m at λ=810 nm).

The laser cavity is formed by placing mirrors at the ends of the fiber. The input coupling mirror has a reflectivity greater than 99.5% at the lasing wavelength, but has transmission at pump wavelengths greater than 95%. The fiber is coupled to this mirror using an index-matching liquid to decrease losses. The output coupling mirror is 95% reflective at the laser wavelength, and the fiber is simply butt-coupled to this mirror.

A semiconductor laser (SDL-5412-H1, 100 mW) pumps the fiber. Its wavelength is temperature-tuned to 811 nm by using the built-in TE cooler to hold the laser at 13.5 C. The pump light is collimated using an AR-coated lens with a numerical aperture of 0.4. Because the pump's beam is strongly elliptical, it is poorly mode-matched to the fiber and would couple very inefficiently into the fiber if not corrected somehow. We therefore shape the beam with a set of anamorphic prisms (Melles Griot 06 GPA 001, 3x expansion). These optics are placed approximately 6 cm from the collimating lens. Then, at a distance of 20 cm from the prisms, we place a biconvex lens with a focal length of 500 mm (Newport KBX082). This lens reduces the beam diameter that is incident on the final input coupling lens, located 33 cm downstream. This final lens (Newport FL10B) has a focal length of 12 mm and a numerical aperture of 0.25. The pump light is focused by this lens through the input coupling mirror and into the fiber core.

Our experiments involve a *driven* laser system. To drive the fiber, we sinusoidally modulate the semiconductor laser pump current I_d. This has the effect of modulating the pumping field applied to the fiber. We use a function generator (HP 3325A) to drive the modulation input of the current supply (ILX Lightwave LDC-3722). The current varies proportionally with the voltage applied at the input, with a ratio of 20 mA/V. So if the modulation amplitude is A_1 = 1.5 V, the current will swing + and - 30 mA from the DC value. The drive current of the diode laser relates to its power output as follows. The threshold value of the current is about 25 mA, at which point the output power is 0.8 mW. The power then

increases approximately linearly with the current until the diode is emitting 108 mW at I_d=120 mA.

Characterization of the fiber laser dynamics

As stated in the introduction, the driven fiber laser is a system which is known to display a wide variety of nonlinear dynamical behaviors. To characterize our laser, we have generated various diagrams based on time series measurements of its total output intensity. The laser output was detected using a New Focus model 1811 photodetector, and this signal was sent to a digitizer. From the raw digitized time series, we can produce bifurcation diagrams, return maps, power spectra, and reconstructed attractors.

With the correct parameters, the fiber laser will display period doubling behavior. Furthermore, there are several experimentally accessible quantities which can be used as the bifurcation parameter: modulation frequency f_1, modulation amplitude A_1, and the DC pumping level. Figure 2 shows representative time series of the fiber laser's output intensity as it progresses through the period doubling cascade as a function of A_1. Wave (a) shows a limit cycle (period-1). The small peaks are relaxation oscillations and are not counted when determining the periodicity of the wave. For this wave, A_1= 0.57 V. Waves (b), (c), and (d) show progressively higher periodicities and have values for A_1 of 1.65, 1.97, and 2.15 V respectively. For all four waveforms, the diode laser pump current was I_d = 85 mA and the modulation frequency f_1 = 11.95 kHz. The waveforms were sampled by the digitizer at a frequency of 5 MHz, corresponding to about 420 points per drive cycle.

Another way to represent this sort of data is to construct a bifurcation diagram. Figure 3 is such a diagram, demonstrating how the system period doubles into chaos as a function of the modulation frequency f_1. Each dot represents a measured peak value when the system is driven at the frequency shown on the x-axis. To generate this diagram, f_1 was swept up from 12 - 22 kHz in 50 Hz increments. The digitizer recorded 32k points at each frequency setting, sampling

every 300 ns. The peaks were then determined by simply locating the maxima above an arbitrary threshold value, set in order to reject the relaxation peaks. This entire procedure was automated. The other parameters for the laser for this diagram were I_d=65 mA and A_1=2.1 V.

There are a few important features in Figure 3. First, there is a definite period doubling route to chaos, visible up to period-4. Within the chaotic regime, there is also a window of periodicity. We found that these windows sometimes contained very high period orbits, with periods up to 38 times the drive period. Also, the chaotic region ends as the frequency becomes too high, and the system reverts to a limit cycle.

It is worth pointing out that one needs to be rigorous to claim that a seemingly irregular signal is in fact chaos. An FFT can reveal that what appears to be chaos may have a high periodicity. The best thing to do, perhaps, is to calculate the largest Lyapunov exponent if the time series in question is of sufficiently length. We have done such a calculation for a suspected chaotic time series and compared it with the results of the same calculation on a stable period-16 orbit. These time series were 256k in length. The stable orbit yielded an exponent of 7×10^{-5}, while the chaotic series produced an exponent seventy times larger. This makes a strong case that the system truly was exhibiting chaotic dynamics.

Experiment: Effects of a weak subharmonic perturbation

The experiment we performed on this nonlinear optical system was to examine the behavior of the system, and its bifurcation points in particular, when the system is subjected to a near-resonant perturbation. That is, consider that we normally drive the system through the pump field with a signal such as $A_0 + A_1\cos(2\pi f_1 t)$. Then, introducing a subharmonic perturbation means we add an additional term of $A_2\cos(2\pi f_2 t + \phi)$, where $f_2 = f_1 / n$ (integer n) and ϕ is just a relative phase difference between the driving term and the perturbation term.

It has been found that such a perturbation can shift the bifurcation point of the subharmonic being perturbed. That is, if the perturbation is at frequency $f_2 = f_1 / 2$, the point at

which the period-2 orbit becomes unstable may be shifted. Furthermore, the shift may be forward or backward. That is, the period-2 orbit may be stable over a region in which it was previously unstable, or it may become unstable in a previously stable region. There has been some general theoretical work on this type of situation [7-9], and experiments on a magnetostrictive ribbon have been carried out [10]. Similar experiments have been performed on a CO_2 laser with modulated losses [11], although in this case the thrust of the effort has been to create stabilized orbits in a regime which would otherwise be chaotic.

In our experiment, we generated the perturbation with a second function generator (HP 3314A). This generator is preprogrammed to generate an f_1/n wave, where f_1 is externally supplied. For these experiments, we always chose n=2, so that the perturbation had the same period as a period-2 orbit. The phase relationship between the two signals is also adjustable, allowing us access to ϕ. These two waves are then summed with an op amp, and a line driver delivers the resulting combination to the modulation input of the current supply.

We have collected data which highlights the effects of the perturbation strength A_2 and the phase difference ϕ. First, we set the parameters of the fiber laser such that it would undergo a period-doubling bifurcation from a period-2 into a period-4 regime. We characterized this unperturbed reference state by generating a modulation amplitude bifurcation diagram. This is shown in Figure 4. This diagram shows slightly more complicated behavior than Figure 3. Notice, however, the nice transition into a period-4 state at $A_1 = 1.94$ V. The parameters of this setup are $I_d = 80$ mA and $f_1 = 12.7$ kHz. This diagram was made using the same procedure as was used for Figure 3, with the exception that A_1 was the bifurcation parameter rather than f_1.

To begin our experiment, we set $A_1 = 2$ V, placing the system just into the period-4 state. To understand the effects of the subharmonic perturbation, we needed to consider both the effects of the phase ϕ and the strength A_2 of the perturbation. To measure these effects, we made a series of bifurcation diagrams, using ϕ as the bifurcation parameter. Each successive diagram

had a greater value for A_2. That is, we set the ratio A_2/A_1, made a bifurcation diagram as ϕ was swept from 0-180 degrees, increased A_2/A_1, and repeated the process.

The results are shown in Figure 5. Diagram (a) has no perturbation (A_2=0), and so is a reference showing the original period-4 state. The successive diagrams (b) through (f) are in order of increasing A_1/A_2 in 8% increments, to a maximum of 40%. Notice the wide variety of behavior induced by the perturbation. These effects include stabilization of a period-2 state, leaving the system in a period-4 state, or destabilizing the original state into period-8 or chaotic dynamics. It is also striking that all these behaviors can be found for a single setting of A_2/A_1, depending on the particular value of ϕ. Notice also that as the perturbation becomes quite strong, the system is more likely to be locked onto a period-2 orbit.

Another way of approaching this experiment is to consider how the bifurcation point shifts as a function of A_1 as the perturbation is applied. For this question, we set A_2 at a fixed value of 0.2 V and made bifurcations diagrams by sweeping A_1. This was done for several values of ϕ, and the results are shown in Figure 6. The bifurcation in the unperturbed system occurred at A_1 = 1.94 V, as shown in diagram (a). The vertical line locates the bifurcation point. With the perturbation at different phases, however, the bifurcation point clearly shifts. In diagram (b), with ϕ = 20 degrees, the bifurcation point shifts down to A_1= 1.7. Diagram (c) has ϕ = 125 degrees and the bifurcation at A_1= 1.8 V, and in (d), with ϕ = 170 degrees, the point shifts in the other direction and occurs at A_1= 2.05 V. It is worth noting that there is some hysteresis present here, depending on if A_1 is swept up or down. For consistency, these plots were made by starting at the upper value of A_1 and incrementally decreasing.

A more compact way to represent this type of data is to generate a two-dimensional bifurcation diagram. For this type of diagram, we determine the periodicity of the laser dynamics on a grid of two parameters, A_1 and ϕ. As before , we located a regime where the laser would display a period doubling bifurcation into a period-4 orbit. The parameters were I_d = 78 mA and f_1 = 12.5 kHz. In this case, the bifurcation of interest occurred at A_1 = 1.92 V. We then fixed A_2

at a value of 0.27 V. To take the data, the acquisition program would set a value for ϕ, sweep through values of A_1 incrementally, then increment ϕ, `and so on. At each point (A_1,ϕ) the program would examine the peaks of the time series and determine the periodicity of the wave. The period was determined by calculating which orbital period gave the minimum average deviation for corresponding peaks.

The results are shown in Figure 7. On this two-dimensional grid, we have drawn contours to distinguish the borders between regions of different periodicity. While there is some complex behavior going on here, the effect of shifting the bifurcation point is quite apparent. There is a large island of period-2 behavior in the upper left half of the figure. As A_1 increases along the horizontal axis, the system eventually bifurcates into a period-4 regime. However, the particular value of A_1 at which the bifurcation occurs depends fairly smoothly on the phase. And again, the shift goes on both directions (recall that the value of A_1 for the unperturbed case is 1.92 V).

Conclusion

We have investigated the nonlinear dynamical behavior in a driven Nd-doped fiber laser system. This system displays period doubling behavior as a route to chaos. We have found that the application of a small subharmonic perturbation, with a period twice that of the drive, has dramatic effects on the system dynamics. Particularly, the point at which the period-2 orbit becomes unstable is shifted. This shift can be positive or negative, depending on the amplitude of the perturbation and the relative phase between the perturbation and the master driving signal.

REFERENCES

[1] M.W. Phillips, H. Gong, A.I Ferguson, and D.C. Hanna, Optics Comm. **68**, 215 (1987).

[2] D. Derozier, S. Bielawski, and P. Glorieux, Optics Comm. **83**, 97 (1991).

[3] S. Bielawski, D. Derozier, and P. Glorieux, Phys. Rev. A **46**, 2811 (1992).

[4] E. Lacot, F. Stoeckel, and M Chenevier, Phys. Rev. A. **49**, 3997 (1994).

[5] F. Sanchez, M. LeFlohic, G.M Stephen, P. LeBoudee, and P.L. Francois, IEEE J. of Quan. Elec. **31**, 481 (1995)

[6] S. Bielawski, D. Derozier, and P. Glorieux, Phys. Rev. A **47**, 2492 (1993); S. Bielawski, M. Bouazaoui, D. Derozier, and P. Glorieux, Phys. Rev. A **47**, 3276 (1993).

[7] P. Bryant and K. Wiesenfeld, Phys. Rev. A **33**, 2525 (1986).

[8] G. Cicogna and L. Fronzoni, Phys. Rev. E **47**, 4585 (1993).

[9] Sanju and V.S. Varma, Phys. Rev. E **48**, (1993).

[10] S.T. Vohra, F. Bucholtz, K.P. Koo, and D.M. Dagenais, Phys. Rev. Lett. **66**, 212 (1991); S.T. Vohra, L. Fabiny, and K.Weisenfeld, Phys. Rev. Lett. **72**, 1333 (1994); S.T. Vohra, L. Fabiny, and F. Bucholtz, Phys. Rev. Lett. **75**, 65 (1995).

[11] R. Meucci, W. Gadomski, M.Ciofini, and F.T. Arecchi, Phys. Rev. E **49**, 2528 (1994); M. Ciofini, R. Meucci, and F.T. Arecchi, Phys. Rev. E **52**, 94 (1995).

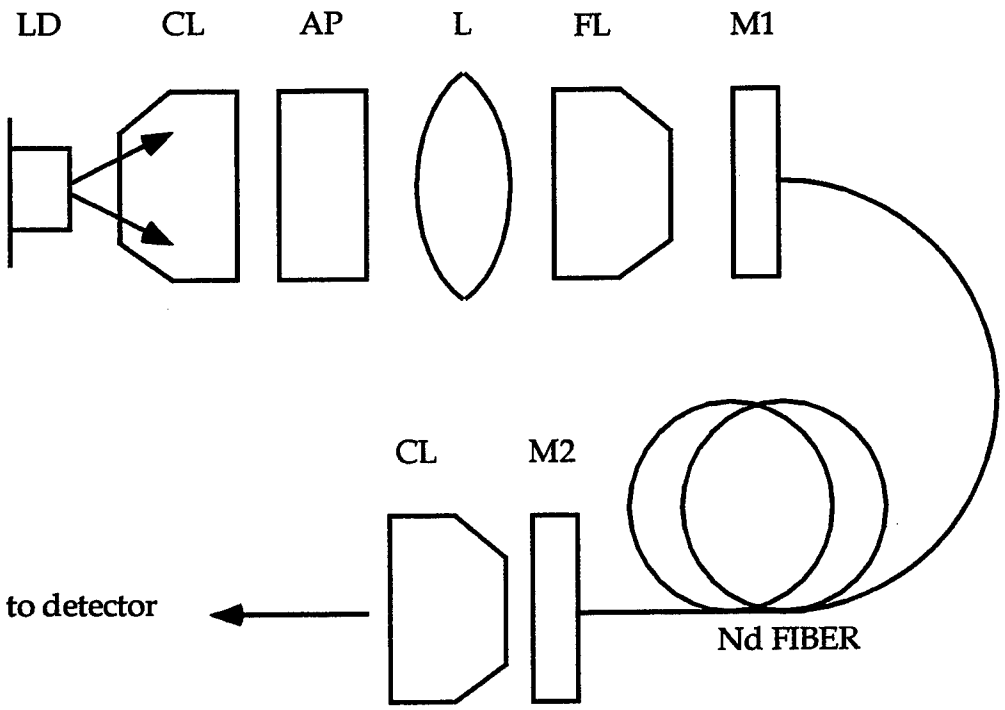


Figure 1: Experimental setup for diode-pumped Nd fiber laser. Notation: LD=laser diode, CL=collimating lens, AP = anamorphic prisms, L=lens, FL=focusing lens, M1= input coupling mirror, M2=output coupling mirror.

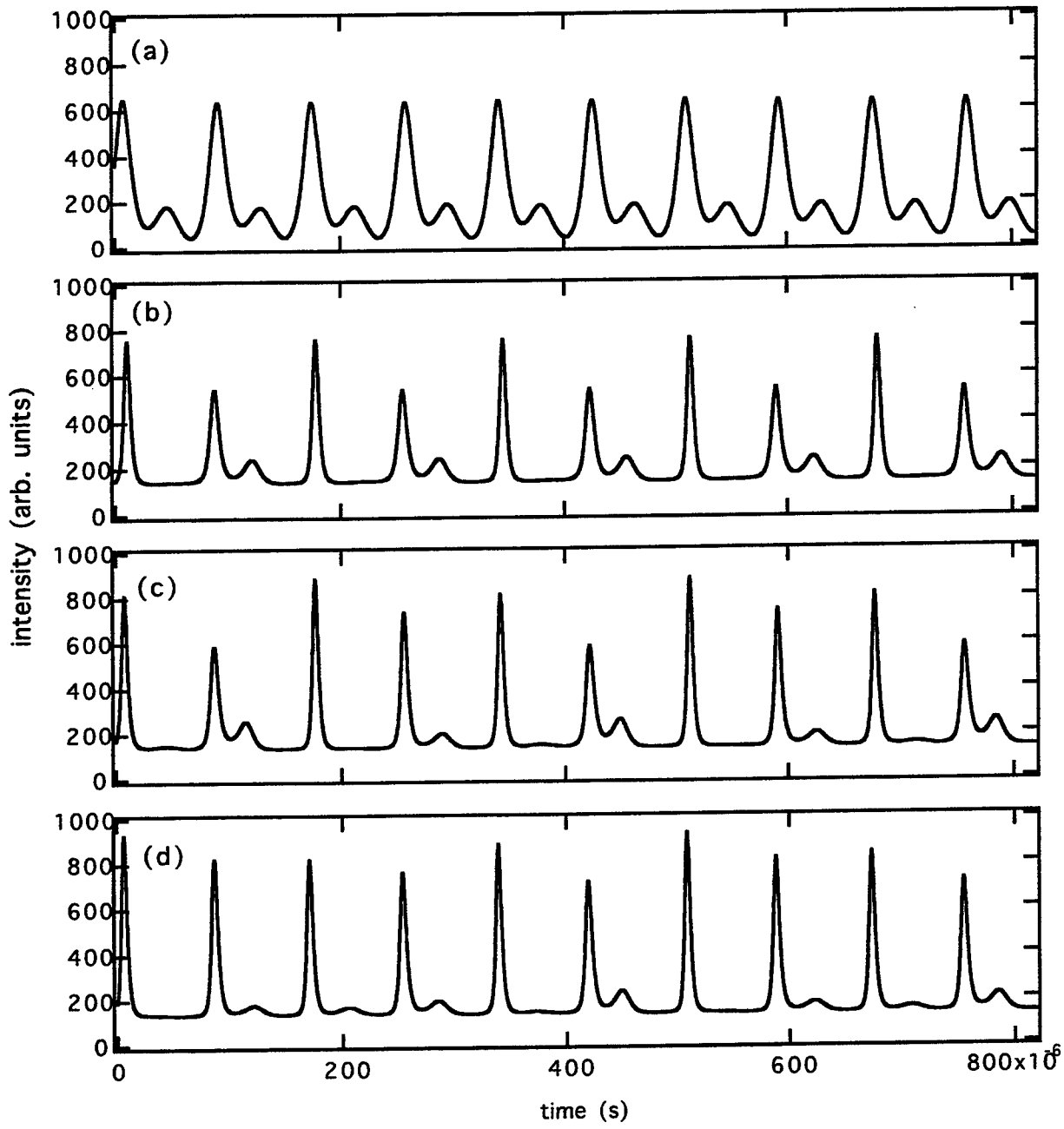


Figure 2: Representative waveforms of fiber laser output

(a) limit cycle with relaxation oscillation
(b) period-2
(c) period-4
(d) chaos

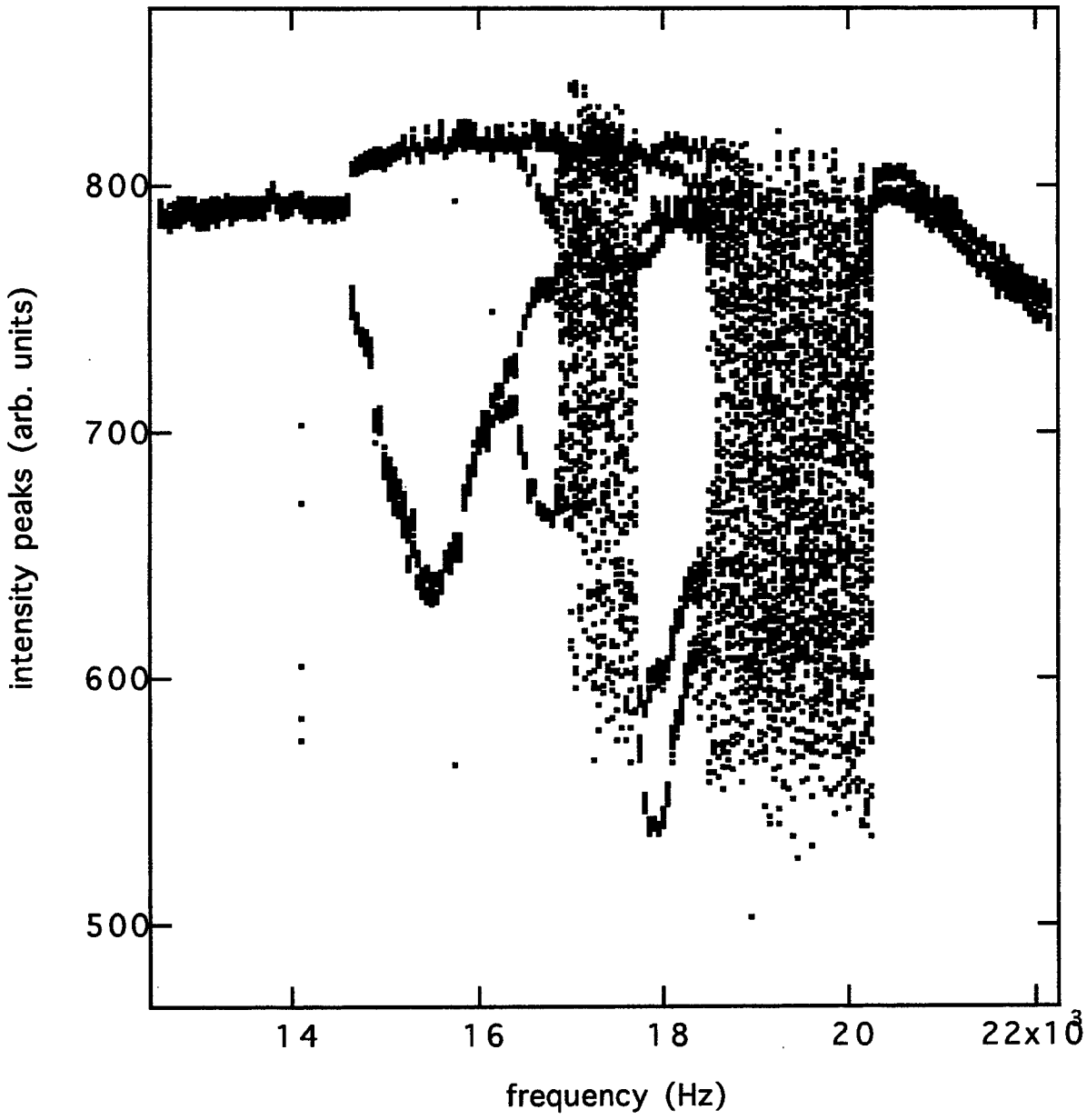


Figure 3: Frequency bifurcation diagram

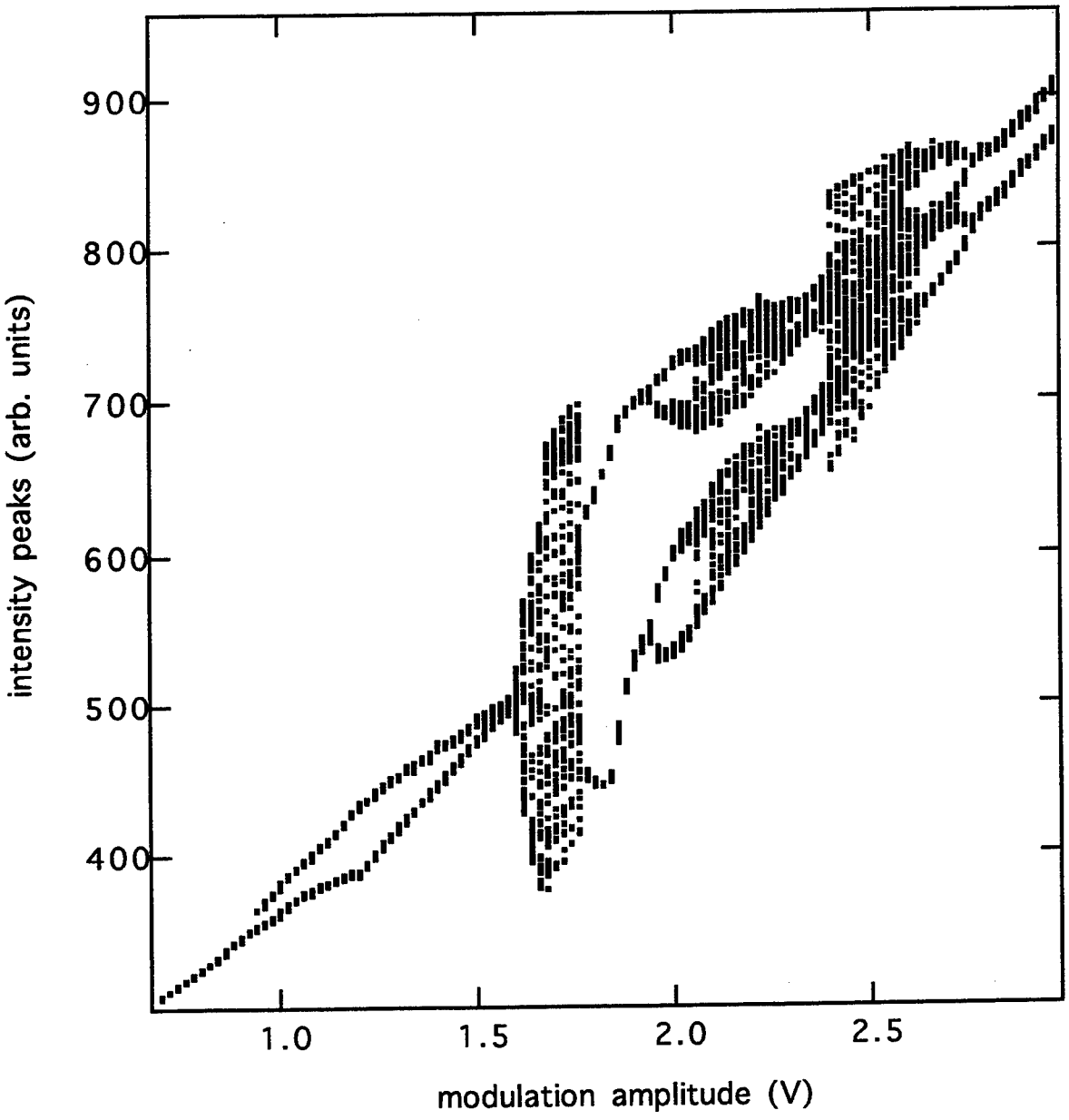


Figure 4: Amplitude bifurcation diagram

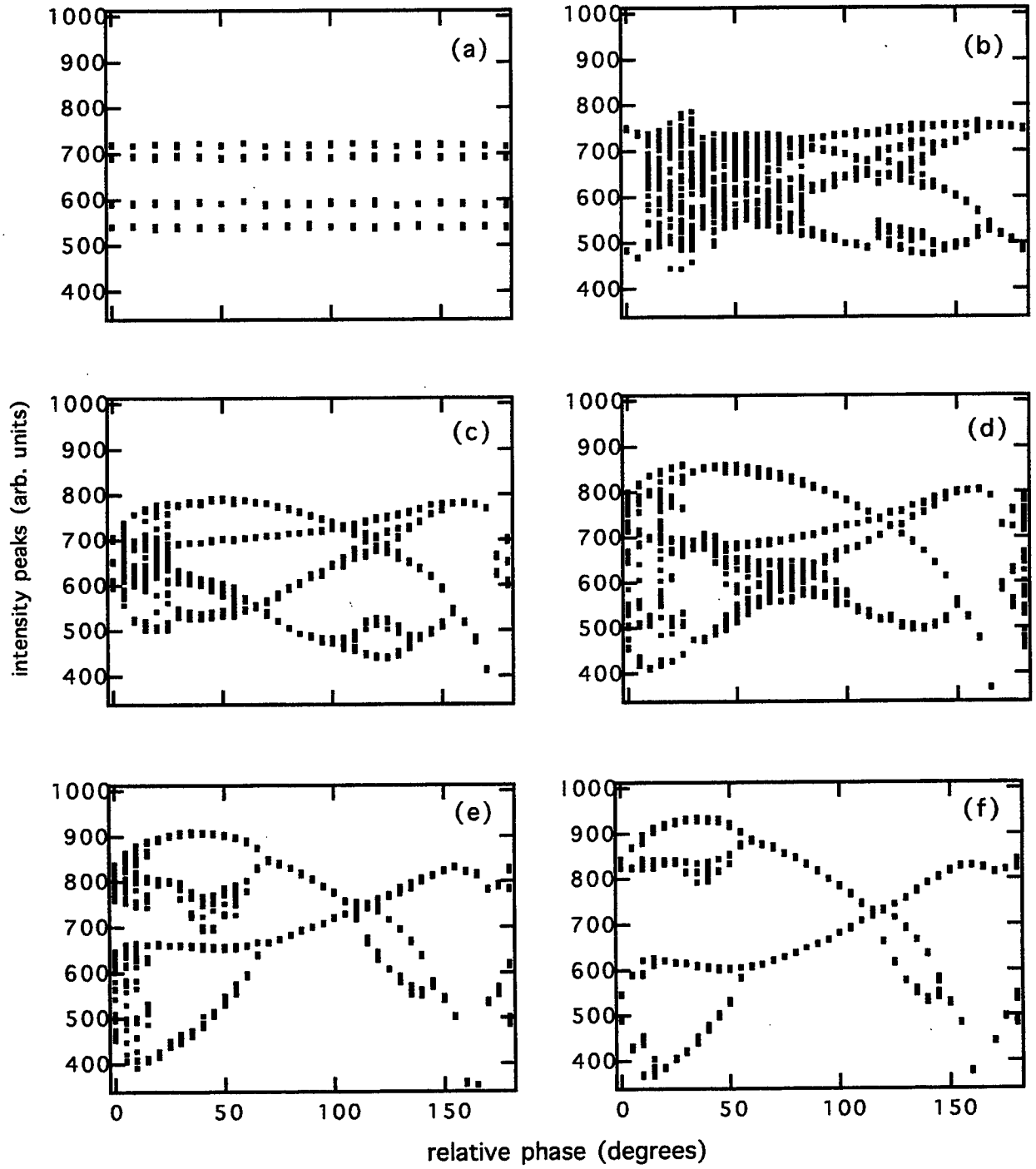


Figure 5: Phase bifurcation diagrams. The diagrams (a) - (f) are in order of increasing perturbation amplitude.

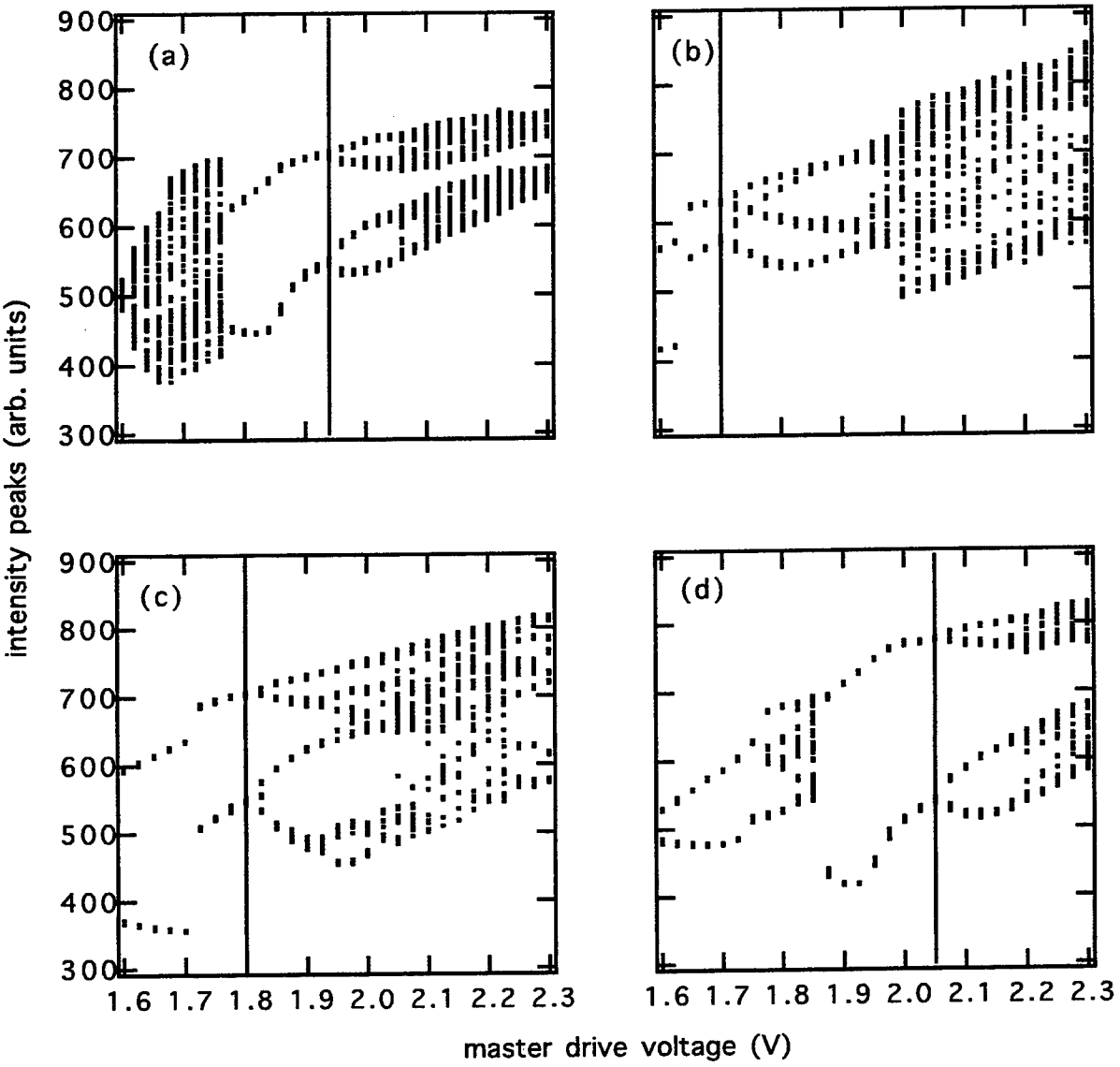


Figure 6: Amplitude bifurcation diagrams with subharmonic perturbation,demonstrating shift of bifurcation point. (a) is the unperturbed reference state, while (b), (c), and (d) have relative phase shifts of 20, 125, and 170 degrees, respectively.

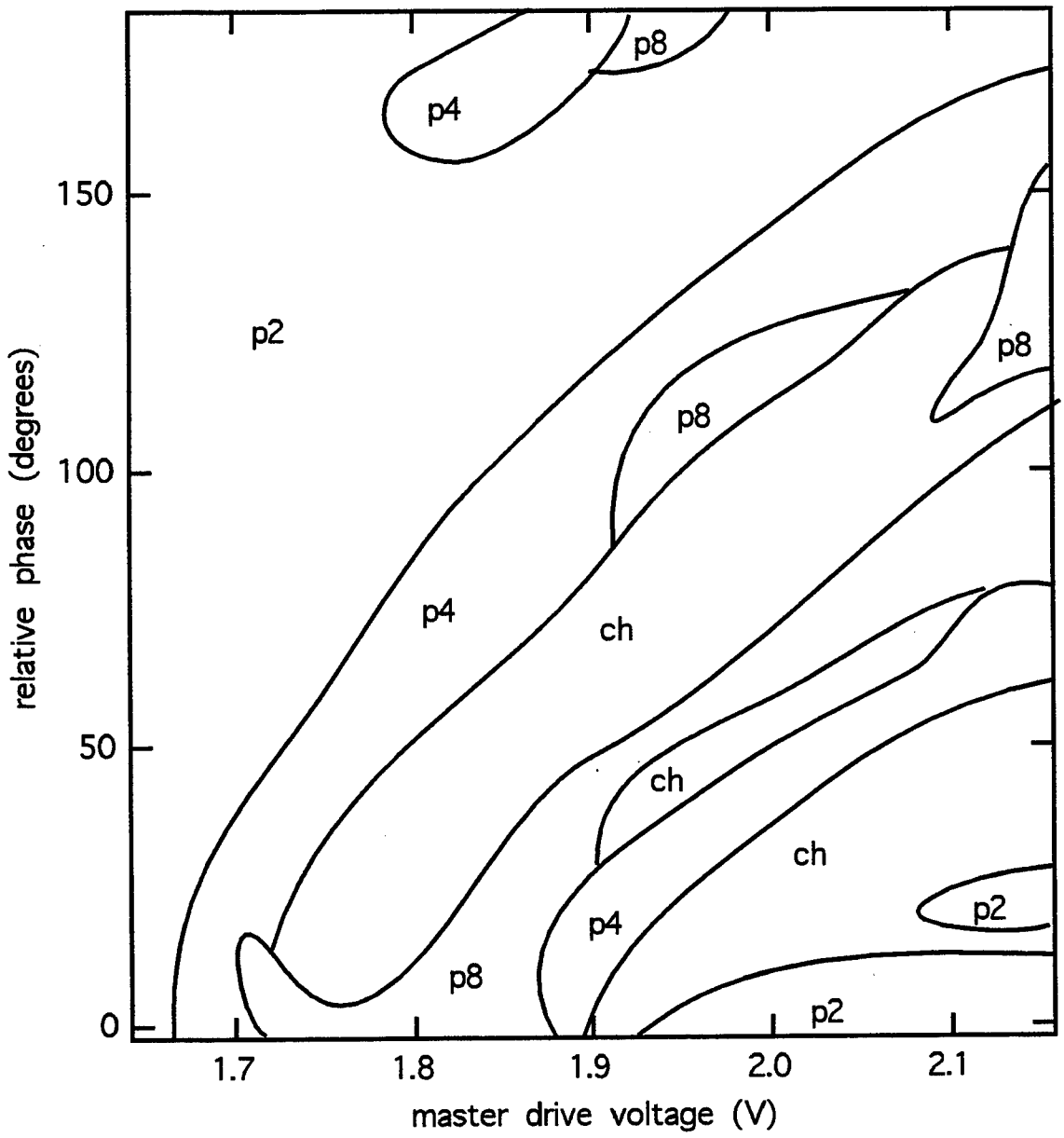


Figure 7: Two-dimensional bifurcation diagram

Monte Carlo Simulations of the MCD Spectra of Cryogenic NA/AR Matrices

Heidi A. Terrill-Stolper
Dept of Physics

New Mexico State University
Box 30001, Dept. 3D
Las Cruces, NM 88003

Final Report for:
Graduate Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Laboratory

September 1995

MONTE CARLO SIMULATIONS OF THE MCD SPECTRA OF CRYOGENIC NA/AR MATRICES

Heidi A. Terrill-Stolper
Graduate Student
Department of Physics
New Mexico State University
and
Chemical Physics Laboratory
Eastern New Mexico University

Abstract

Monte Carlo (MC) simulations of the magnetic circular dichroism (MCD) spectra of the pre-HEDM material consisting of sodium (Na) atoms matrix isolated in cryogenic argon (Ar) lattices revealed never before seen parameter dependent structures in the spectra. The dependance of the structures on the Na ^{2}P spin-orbit splitting, applied magnetic field, and "magnetic" temperature are charted. The moment analysis method is tentatively confirmed through evaluation of simulated absorption and MCD spectra.

MONTE CARLO SIMULATIONS OF THE MCD SPECTRA OF CRYOGENIC Na/Ar MATRICES

Heidi A. Terrill-Stolper

Introduction

It has long been known that the addition of small amounts of metal powder to solid rocket fuels increases their specific impulse, making the fuel more efficient. The problem with this method of doping the fuels, is that the energy gained in the oxidation of the metal is reduced by the binding energy of the metal clusters. If the metals can be trapped in the rocket fuels as individual atoms, then this problem can be overcome, producing even larger increases in fuel efficiency.

The goal of the current research is to study the nature and energetics of the high energy density matter (HEDM) solid molecular hydrogen (H_2) doped with atoms of light metals such as lithium (Li), boron (B), and sodium (Na). Since hydrogen is very difficult to model due to quantum effects, the rare gases, xenon (Xe), krypton (Kr), argon (Ar), and neon (Ne) are used as pre-HEDM matrix hosts in which the trapping site dynamics can be studied and characterized before attempting the more difficult solid hydrogen host. The rare gases make reasonable models for H_2 since both species are closed shell and the essentially free rotations of the H_2 make it effectively spherical.

The primary focus of the summer research has been Monte Carlo (MC) simulations of the magnetic circular dichroism (MCD) spectra of Na atoms matrix isolated in cryogenic rare gas solids, in this case Ar. The theory of these MC-MCD simulations was recently prepared by Kenney [1] as a straight forward extension of the MC simulation theory of Boatz and Fajardo [2] for the absorption spectra. Na and Ar were chosen primarily because they are easiest to work with from a theoretical stand point. Both atoms are heavy enough to be treated quasi-classically, thereby avoiding the complication of quantum effects, but are small enough not to require relativistic corrections as do Xe and Kr. Furthermore, there are good pair potentials for Na–Ar, but not the other metal/rare gases combinations.[3]

Background [4]

The $^2S \rightarrow {}^2P$ electronic absorption spectra of alkali metals (M) matrix isolated in cryogenic rare gases are characterized by a three peaked absorption structure called a "triplet." The three peaks arise due to a dynamic splitting of the alkali metal (M) electronic P states by the vibrating rare gas (Rg) matrix atoms. So far, three different triplets have been identified in the M/Rg cryogenic matrices. The first is called the "red triplet", a triplet absorption structure centered around the isolated metal atom transition energy. The "red" trapping site is most likely a large multi-atom hole in the host matrix. The second triplet structure is the "blue triplet", which is blue shifted slightly with respect to the red triplet. It is believed that the trapping site(s) that produce(s) this feature are slightly tighter and therefore higher energy than that/those of the red triplet site. The third triplet is called the "violet triplet" and is blue shifted with respect to the blue triplet structure. This feature is accessible only through laser ablation

of the M guest atoms and, thus far, has been observed only in the Na/Ar system. The trapping site producing the violet triplet is therefore interpreted to be tighter still than the blue triplet site.

Magnetic circular dichroism (MCD) is the difference between the absorbance of left and right circularly polarized light of a system placed in an axial magnetic field. MCD is highly sensitive to site dynamics, making it an excellent diagnostic tool for the study of the M/Rg systems. Since MCD measures a difference in absorbance, small variations in site structure are expected to alter the shape and characteristics of the MCD spectra as we move between trapping sites and vary the type of impurity or host atoms.

Description of Monte Carlo Program and Simulation Parameters [5]

One of the most widely used simulation techniques for the study of condensed phase systems is the Metropolis Monte Carlo (MC) method. In a MC simulation program, the initial positions of the particles of interest, in this case the matrix atoms, are chosen using an experimental radial distribution function or an "ideal" lattice structure. A new position for each particle in the simulation is then randomly generated, and the new configuration is tested to see if it is more energetically favorable than the previous configuration. If it is, then the new configuration is accepted, if it is not, then the new configuration is accepted with a Boltzmann probability. After each particle in the simulation has been moved (or not moved as the case may be), the absorption and MCD spectra of that configuration are recorded. Then the process is repeated. Most MC simulations run for several tens of thousands of configurations, with the first 20-25% of the configurations ignored in the averaging to allow the system to equilibrate, and the energy, spectra and other properties averaged over the subsequent configurations. The basic idea behind this modeling method is that a single system, observed at many thousands of configurations, with its properties averaged over these configurations, will have basically the same spectra as many thousands of systems, each in a different configuration, observed simultaneously.

Since a real lattice contains millions of atoms, it is not feasible to model the "real" system. Therefore, the number of atoms in the "simulation" box is chosen as one of the adjustable parameters of the simulation, with periodic boundary conditions applied. Hence the simulation approximates the bulk behavior of the solid, with the accuracy of the approximation increasing with the number of atoms included. Based on previous simulations of the absorption spectra of Na atoms in Ar [2], most of the present simulations are run using a solid argon "box" containing 108 atoms, some number of which have been removed and replaced with the sodium impurity atom. At the beginning of each simulation, the program is given a specific lattice structure where each of the argon atoms is placed at the "ideal" location of a face-centered cubic (fcc) lattice. The sodium is placed at the center of the hole created by removal of one or more nearest-neighbor argon atom(s), then the system is allowed to equilibrate by completing a specified number of MC moves before averaging begins.

Another limit to the precision of the program output is the number of configurations over which the simulation is averaged. In most cases, 80,000 configurations are generated, with the energy and spectral properties averaged over 60,000 configurations. Generating 80,000 configurations and completing the spectral evaluation of each

configuration takes approximately four hours of CPU time on a high-performance workstation. For higher simulation temperatures, regions of high interaction, or small magnetic fields, the signal to noise ratio of the spectra using only 80,000 configurations is often to small for meaningful assessment. Simulations over larger numbers of configurations, however take more time, e.g. a simulation over 200,000 configurations takes approximately 11 hours to complete. Therefore, simulations over large numbers of configurations are run only when the results of a lesser number of configurations show they are needed, or when a greater level of confidence in the result is needed.

The MC program is designed to allow user definition of the parameters governing the perturbations examined in the simulation. The parameters of most interest in the context of this research include choice of size of the trapping site hole, simulation temperature, spin-orbit splitting, magnetic field (for simulation of the MCD), and number of configurations.

One of the user-selected features built in to the MC-MCD program is called the "magnetic temperature." The magnetic temperature allows the Boltzmann distribution that determines the relative populations of the magnetically split ^2S ground state to be determined independently from the simulation temperature. This option is important since the temperature of the simulation necessary to get the correct vibronic structure is higher than the corresponding temperature in the real system. Therefore the simulation temperature can be chosen to get the correct vibronic structure, and the magnetic temperature selected to obtain realistic Boltzmann distributions.

In a real matrix, the local orientation of the matrix trapping sites with respect to the magnetic field (B) are not known. In the simulation however, the magnetic field is oriented along one of the principal axes (in this case the z-axis) of an arbitrary reference frame. This is done to simplify the application of the selection rules for MCD spectra. In an attempt to correct the simulation for this arbitrary reference frame, a simulation effect called "random rotations" was built into the program; selecting this feature effects a rotation of the rare gas perturbation matrix at each configuration through randomly chosen Eulerian angles.

Simulation Temperature vs. Real Temperature

Calibration of Simulation Temperature with Real Temperature

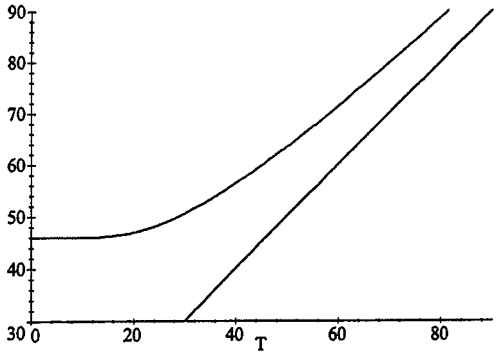
One of the effects of the semi-classical approximation necessary in the simulations is that the simulation underestimates the vibronic motion of the molecules since it cannot account for the zero point vibrations of the lattice. Therefore the temperature of the simulation that governs the thermal fluctuations of the atoms must be higher than the "real" temperature which it is modeling. The simulation temperature verses "real" temperature question was addressed in two ways.

Bergsma et al. [6] discussed the thermal harmonic correction, and provided an equation that relates the simulation temperature, T_s, to the "real" temperature, T, in molecular dynamics simulations.

$$T_s = \frac{\hbar\omega}{2k_B}\left[\tanh\left(\frac{\hbar\omega}{2k_BT}\right)\right]^{-1} \tag{1}$$

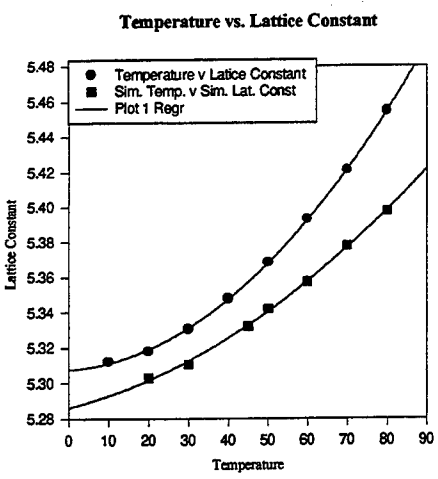
The Debye energy, $k_B\Theta_D$, was used to estimate the energy, $\hbar\omega$, due to the nuclear vibrations at zero Kelvin (K).

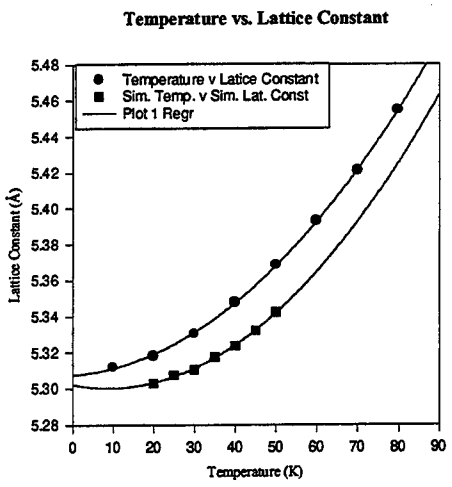
Taking $\Theta_D = 92K$, this provided a prediction that the required simulation temperature to mimic real temperatures below 10 K would be 46 K. Using this as a starting point, several simulations were run with a simulation temperature of 45 K, with various magnetic temperatures.



Real Temperature vs. Simulation Temperature:

A second method was suggested [7] in which the "real" temperature could be estimated by comparing the solid argon simulation lattice constant with the experimentally determined lattice constants. Simulations were run on solid argon, and the radial distribution function (rdf) was examined. The lattice constant was determined by integrating, using a program written for this purpose, over the first peak in the rdf to find the most probable separation between nearest neighbor atoms. The distance between first nearest neighbors, d, is related to the lattice constant, a, by the relation: $a = \sqrt{2}d$. Based on this analysis a simulation temperature of 31 K is necessary to model the zero point motion of the real system.

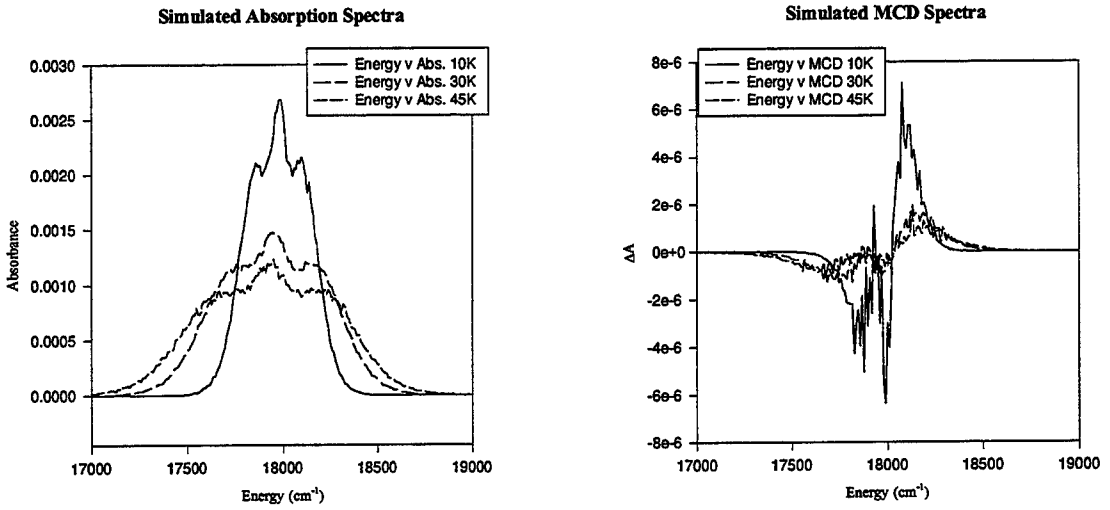




The simulation temperature required to model a given real temperature should approach the real temperature in the limit of large temperatures. The simulation temperature verses lattice constant exhibits the expected behavior at simulation temperatures less than 50 K, however the simulation temperature and real temperature curves diverge for higher temperatures. Increasing the number of atoms included in the simulation did not have an appreciable affect on this result. This suggests that modeling of systems requiring simulation temperatures in excess of 50 K is not possible.

Simulations at the "Real" Temperature

Since most of our MC-MCD simulations are carried out at a simulation temperature of 10 K, which has been shown in the preceding section to be less than the temperature required to mimic the zero point motion of the lattice, an important question to be answered is how much of an effect does the simulation temperature have on the results of the simulation. To answer this question a small number of simulation sets have been run at the simulation temperatures that have been determined to give the correct vibronic structure in the above analyses. For the most part, the primary effect of increasing the simulation temperature was a broadening the spectra and shifting of the absorption energy barycenter to slightly lower energies.



Comparison of the temperature effects in the absorption and MCD spectra for the case of $SO = 17cm^{-1}$, $B = 1.0\,T$, magnetic temperature, $T_m = 10\,K$, at simulation temperatures of 10, 30, and 45 K

Moment Analysis Method

One of the most commonly used methods of analyzing the MCD spectra of matrix isolated species is the moment analysis method. It was first developed for use in analyzing the MCD and absorption spectra of F-centers.

It purportedly can be used to extract the spin-orbit splitting, g-factor, and cubic and non-cubic mode parameters from spectra.

Moment analysis is based on the calculation of band moment integrals over the spectral features, the ratios of which are related to physical quantities. The n^{th} absorption and MCD moment integrals are [8]

$$A_n = \int \left(\frac{A}{\mathcal{E}}\right)(\mathcal{E}-\overline{\mathcal{E}})^n \, d\mathcal{E} \tag{2}$$

$$M_n = \int \left(\frac{\Delta A}{\mathcal{E}}\right)(\mathcal{E}-\overline{\mathcal{E}})^n \, d\mathcal{E} \tag{3}$$

where the subscript, n, indicates the order of the moment, and $\overline{\mathcal{E}}$ is the absorption barycenter, given by

$$\overline{\mathcal{E}} = \frac{\int A d\mathcal{E}}{\int \left(\frac{A}{\mathcal{E}}\right) d\mathcal{E}} \tag{4}$$

There are four equations relating the zeroth and second absorption moments, and the first, second, and third MCD moments to the physical parameters. They are [8]

$$\frac{M_1}{A_0} = 0.934\left(g - \frac{\zeta}{2kT}\right) \tag{5}$$

$$\frac{M_3}{M_1} = \frac{3}{4}\zeta^2 + 3\Delta_C^2 + \frac{3}{2}\Delta_{NC}^2 \tag{6}$$

$$\frac{A_2}{A_0} = \frac{1}{2}\zeta^2 + \Delta_C^2 + \Delta_{NC}^2 \tag{7}$$

$$\frac{M_2}{A_0} = 0.934\frac{\zeta}{4kT} \tag{8}$$

where $\zeta \equiv$ the spin-orbit parameter ($\zeta = \frac{2}{3}\xi$, the spin-orbit splitting of the state), $\Delta_C \equiv$ the cubic mode parameter, $\Delta_{NC} \equiv$ the non-cubic mode parameter, g is the Landé g-factor, and k is the Boltzmann constant.

In the context of this research, the moment analysis method has been applied to the absorption and MCD spectra of metals trapped in cryogenic rare gas solids [8,9,10,11]. Those of primary interest in HEDM research are the alkali metals, Li and Na. As yet there has been very little experimental data on these metals in the lighter rare gases. Absorption and MCD spectra of Li in Ar, Kr and Xe [9] and Na in Xe [9,10] have been obtained by the Schatz experimental group at University of Virginia (UVa). The Li/Ar and Li/Kr spectra were clearly a mixture of the red and blue trapping sites. Li/Ar absorption and MCD spectra have also been obtained in the Physical Chemistry Laboratory (PCL), at Eastern New Mexico University (ENMU). These spectra are primarily of "blue" site character, but side bands in the spectra interfere with unambiguous analysis of these data.[12] There is a distinct difference between the Li/Ar MCD spectra from the UVa lab as from the PCL. The PCL Li/Ar MCD spectra have one down peak and two up peaks with increasing energy, while the Li/Ar MCD spectra from UVa have two down peaks and one up peak with increasing energy.

For the Li/Rg and Na/Xe experimental spectra [8,9,11] that have been explored using the moment analysis, the values derived from the moment analysis equations for the ^2P spin-orbit splitting were large and negative. The magnitude of the SO splittings obtained from moment analysis for Li/Ar and Li/Kr are substantially reduced as

Correction to Final Report of Heidi A. Terrill-Stolper

In reviewing my final report for the AFOSR/RDL Graduate summer research program, I found an error in one of the equations. On page 24-8 of the report, equation 8 currently reads

$$\frac{M_2}{A_0} = 0.934\frac{\varsigma}{4kT}$$

This equation should read

$$\frac{M_2}{A_0} = 0.934\frac{\varsigma^2}{4kT}$$

Thank you.

compared with that of Li/Xe.[9] In the Li/Ar spectra the calculated SO splitting is slightly more than 10% of the Li/Xe value and the SO splitting in Li/Kr is approximately half of the moment analysis derived value for Li/Xe. These large negative values (several orders of magnitude larger and of opposite sign than the free atom values) seem unphysical. The explanation given for them was that there was substantial interaction between the metal atom valence electron and the rare gas. Although such delocalization of the electron seems reasonable for the F-center case, where an electron is trapped in a negative ion vacancy in an ionic crystal, it seems less likely for the case of a metal atom in a rare gas. Furthermore, simulations on the effects of varying the SO splitting of the impurity atom have shown that the same gross spectral features could be obtained using either large negative spin-orbit splitting or small positive spin-orbit splitting. This suggests that the moment analysis method may be flawed.

To test this hypothesis, a computer program was written to perform the moment integrals on the simulated spectra produced in the MC-MCD program. Since the SO coupling, trapping site symmetry, magnetic field and temperature of the simulation can be chosen explicitly, if the method did not produce the expected spin orbit splitting, the analysis method would be precisely falsified. The results of this experiment were surprising. *The moment analysis consistently reproduced the input spin-orbit splitting to within reasonable variations.* As yet, the only times in which the moment analysis has not reproduced the spin-orbit splitting to within reasonable limits is when the input magnetic field is large. In the cases examined, as the magnetic field is increased, the moment analysis increasingly underestimates the SO splitting. However, this is not a major concern since most experiments are done at magnetic fields much smaller than those producing the observed deviations.

The results of the moment analysis applied to the MCD simulation data lends support to the method. Therefore it seems that the physics behind the method is much more general than we had previously suspected; it is now necessary to study and fully understand the moment analysis method and the physics behind it. Furthermore, as this consistency in the method was not expected, a test of the moment analysis program was devised: the moment analysis program was modified to accept experimental data as well as simulation data, then a Li/Xe experimental spectrum from the published literature [11] was digitized and evaluated using this program. The results were consistent with the published figures. This test therefore had the dual accomplishment of confirming the accuracy of our moment analysis program, and, in the light of the results of the moment analysis on the simulated spectra, the accuracy of the values for the SO splitting published for the Li/Xe and Na/Xe systems. Inasmuch as our reasons for rejecting the moment analysis method were based in the apparent unphysicality of the Li/Xe and Na/Xe spin-orbit splitting derived from the moment analysis, there is also the added challenge of understanding the physics occurring in the M/Xe systems and how it differs from that of the M/Ar systems. One observable difference between the two types of systems are the differences in the experimentally observed absorption spectra. The M/Xe systems have absorption spectra that consist of a singlet peak slightly separated from a doublet.[8,9,11] The M/Ar absorption spectra are much more symmetric.[8,12,13] These revelations suggest that there are novel effects occurring in the M/Xe systems that are not occurring in the M/Ar systems. It has been suggested [14] that the Li is resting on one side of the vacancy/trapping site, so that it sees an asymmetric potential which would cause a separation of the Li P

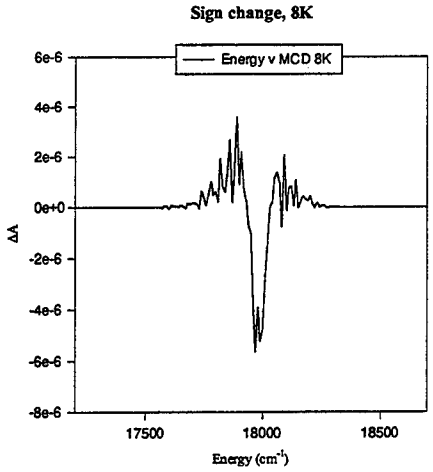
state such that p_x and p_y will be in one irreducible representation, with p_z in another. Thus the asymmetry would produce a splitting of the P state into two levels, one degenerate, the degeneracy of which would then be lifted through vibronic interactions. This is consistent with the Na/Ar MC-MCD simulations of the asymmetric trapping sites which produce triplet absorption spectra that consists of singlet and doublet peaks. A model proposing such an asymmetry in the interaction can be maintained even in the tight single substitutional sites, where the size of the M atom is comparable to that of the Xe it is replacing if it is assumed that the alkali metal and xenon interact in "quasi-covalent" manner where they share electrons. Such a large interaction between the M and the Xe could account for the large negative SO parameter derived from the moment analysis. The idea of bonding between the M and Xe is highly speculative and needs to be explored more fully. Presently there are no reliable M/Xe pair potentials with which to model this system, however it might be possible to obtain pseudo-potentials for the Na/Xe diatomic so that simulations of the system can be performed. Exploring this hypothesis experimentally is extremely challenging, since it would require depositing the M atoms on the surface of the Rg substrate; obtaining concentrations of isolated M atoms sufficiently high that a signal can be observed is doubtful.

While the moment analysis has been extremely successful in extracting the spin-orbit splitting of the simulations, and hence the g-factor, the results for the square of the cubic mode parameter, and to a less obvious extent the square of the non-cubic mode parameter, have been highly erratic. Dramatic changes in the value of the square of the cubic mode parameter occur with simple changes in the choices of magnetic temperatures evaluated in a single set of simulations. In some cases this value has been negative, which is not physically possible since it is a squared parameter. This does not necessarily mean that the moment analysis has failed in this area, however, since the cubic and non-cubic modes are very sensitive to the vibronic structure of the lattice. In the case of a simulation, this means that the accuracy of the extracted cubic and non-cubic modes will depend highly on the relationship between the input simulation temperature and the "real" temperature which it is modeling. For the most part the simulations are run at a simulation temperature of 10 K, which underestimates the zero point vibration of the Rg lattice. Determination of the stability of the moment analysis with respect to the cubic and non-cubic mode parameters will have to wait until there are improved interatomic potentials (currently being generated [15]) and there is time to run accurate simulation temperature/real temperature simulations.

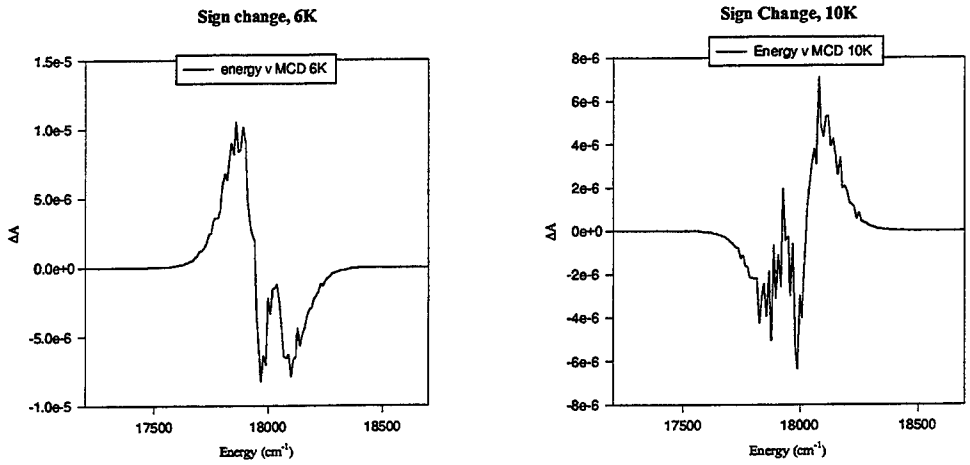
Low Temperature Studies

The magnetic temperature feature of the MC simulation has been used to explore the behavior of the MCD spectra with temperature. The limiting behavior of the MCD spectra for the magnetic temperature at zero (0.01 K) and at infinity (10^6 K), using simulation temperature of 10 K, SO = 17 cm^{-1}, B = 1.0 T in a single substitutional site, was studied. It was found that as the magnetic temperature was increased or decreased, the MCD spectra became very antisymmetric about the absorption barycenter, with the low temperature side having one peak up and two peaks down and the high temperature side having two peaks down and one peak up, with increasing energy. This sign change in the MCD spectra is very interesting, and warranted further study. It was found that the MCD

spectra changed sign at relatively low temperatures. This sign change does not occur at a single temperature, but over a range of temperatures; the "temperature" of the sign change is, therefore, a convenient temperature near the center of the range.



The low temperature sign change at 8K for the Na/Ar system using SO = 17 cm(-1) and B = 1.0 T



Spectra on either side of the low temperature sign change in the Na/Ar system at SO = 17 cm^{-1} and B = 1.0 T.

Spin-Orbit Splitting Effects

A major determining factor of the temperature at which this sign change occurs seems to be the SO splitting of the system. SO splittings of 0.0 cm^{-1}, 0.5 cm^{-1}, 10 cm^{-1}, 17 cm^{-1}, -17 cm^{-1}, and 34 cm^{-1} have been

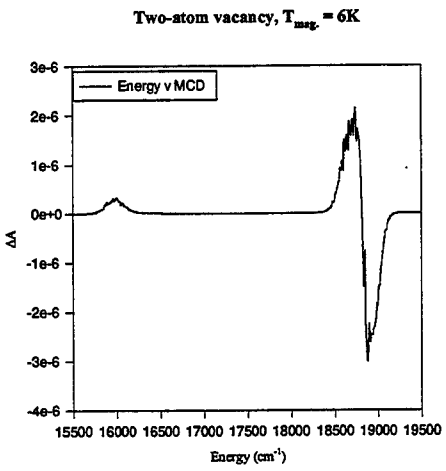
explored. For a fixed value of B, a simple empirical relation relating the ratio of the SO splittings to the ratio of the temperatures at which the sign change occurs was observed:

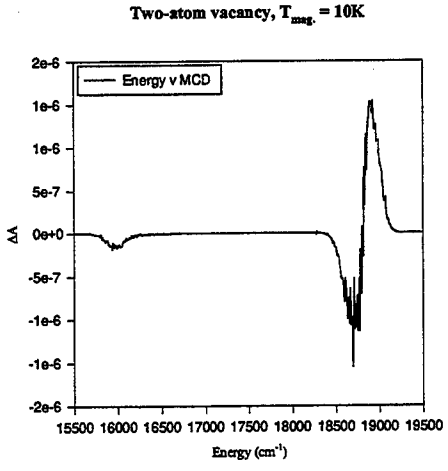
$$\frac{\Delta}{17} = \frac{T_m}{8.3K} \tag{9}$$

Here $\Delta \equiv$ the spin-orbit splitting (in cm^{-1}), $T_m \equiv$ the sign change temperature for that spin-orbit splitting. The 8.3 in the equation is the temperature at which the sign change occurs for SO splitting of 17 cm^{-1}. This is by no means an exact relationship, rather it provides a guide line for looking for the sign change at SO splittings other than 17 cm^{-1}. Futhermore, this relation does not hold when the magnetic field splitting is larger than the SO splitting. The splitting due to the magnetic field, $\Delta E_B = g_l \Delta m_j \mu_B B$, with $g_l \equiv$ the Landé g-factor for the electronic energy level, $\Delta m_j \equiv$ the change in the m_j quantum number between states (in this case it is 1), $\mu_B \equiv$ the Bohr magneton, is approximately 1 cm^{-1} per tesla.

The sign change temperature has been explored for the SO = 17 cm^{-1} case in a variety of different situations. For the most part, there was no change in the (magnetic) temperature of the sign change with varying simulation temperatures and trapping site. There were some variation with magnetic field, this will be discussed in the following subsection.

The sign change for the two- and five-atom vacancy trapping sites, SO = 17 cm^{-1} B = 1 T, occurred around 8 K, as with the one-, four-, and six–atom vacancy trapping sites, however, there were some striking differences in the spectra. Some of these features, the separation of the spectra into one peak at low energy widely separated from a doublet, and the increased width of the spectral band, were already know from simulations of the absorption





spectra [2]. The most striking difference, however, was a complete surprise, and this is a change in the character of the MCD through the sign change. In all of the symmetric trapping sites, regardless of the SO splitting chosen or the magnetic field strength, the center peak of the MCD spectra remained negative. In the case of the two-

and five-atom vacancies, all peaks in the MCD changed sign as the magnetic temperature passed through the sign change temperature.

Simulations using more configurations are needed to tell if the simulated MCD spectra of the thirteen-atom vacancy, SO = 17 cm^{-1} B = 1 T, has the same features as the other symmetric sites at 6 and 10 K, however the sign change very clearly was centered around 8 K as with the other trapping sites having these SO and B values.

At SO = 10 cm^{-1}, B = 1.0 T, the sign change occurred between 4 and 6 K, very near 5 K. The relation 9 predicts the sign change for the 10 cm^{-1} case to occur around 4.9 K. This is supported by the plot of the MCD for this case. The exact temperature of the center of the sign change region has not yet been charted out; it is most likely slightly less than the 4.9 K predicted above, since the magnetic field also plays a role in the position of the sign change, with larger magnetic fields implying lower sign change temperatures (see next subsection).

At SO = 0.5 cm^{-1}, B = 1.0 T the relation 9 predicts the sign change to occur around 0.24 K. However, all spectra in the range of $0.1 \leq T_m \leq 45 K$ had the "high" temperature character, with no sign change observed. In the light of the magnetic field effects (see next subsection) on the sign change temperature, this result is not surprising. In this case the splitting of the states due to the magnetic field is twice that of the SO splitting. Therefore the magnetic field has completely overwhelmed any SO effects.

As with the SO = 0.5 cm^{-1} case, the cases of SO = -17 cm^{-1} and 0.0 cm^{-1} showed no sign change at any of the temperatures checked. To the extent that relation 9 is accurate, for SO = -17 cm^{-1} the sign change is predicted to occur at an (unphysical) negative magnetic temperature so no sign change would be observed. For the 0.0 cm^{-1} case, the sign change temperature is predicted to be 0 K, or negative due to the magnetic field effects (see next subsection). This spectra showed no variation at all in the MCD for temperatures from 2 to 10 K.

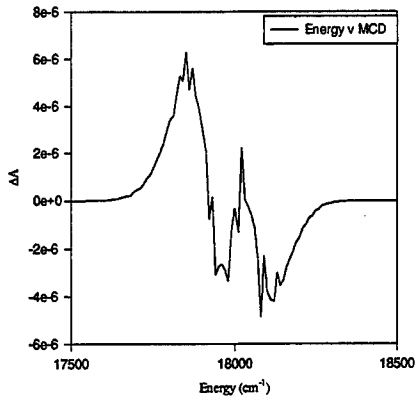
For SO = 34 cm^{-1}, B = 1 T, the temperature of the sign change had shifted to between 10 and 20 K, with its most likely position around 16 K (the exact location has not been confirmed).

Magnetic Field Effect Studies

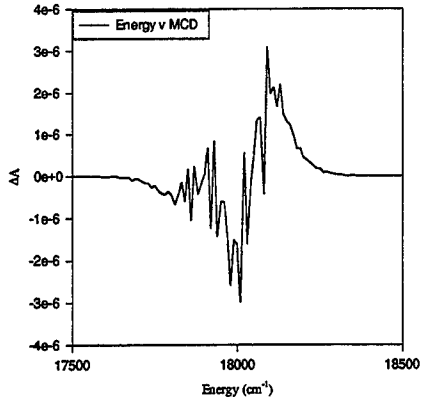
Although the sign change temperature seems to be affected most strongly by the SO coupling, the magnetic field also plays a role. If it is assumed that the dominant effect of the magnitude of the magnetic field is on the relative Boltzmann populations of the magnetically split ^2S ground state of the metal atom, then intuitively the sign change would be expected to shift to higher temperatures as the magnetic field is increased. Clearly this is not the case. A shift to lower sign change temperatures with increasing B can be clearly seen in the plots of the MCD spectra at 4 and 5 K for the SO = 10 cm^{-1} case. In these plots the sign change can be seen to shift from centered slightly below $T_m = 5 K$ at B = 1 T, to centered slightly above $T_m = 4 K$ at B = 5 T.

The case of SO = 17 cm^{-1} was explored at B = 1, 2, 5, and 10 T for the single substitutional site. As in the afore mentioned tests with SO = 10 cm^{-1}, the sign change temperature shifted from around 8.3 K (for B = 1 T) to lower temperatures as B was increased. This shift in the position of the sign change, however, was slower relative to the SO = 10 cm^{-1} case, with the shift evident at B = 5 T and more pronounced at B = 10 T. Judging by the

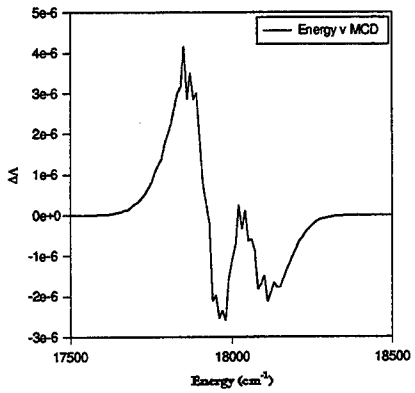
$T_{mag.}$ = 4K, B = 1T



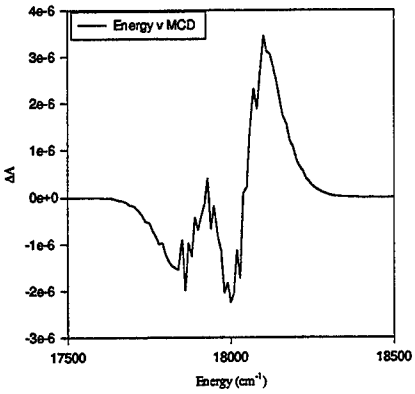
$T_{mag.}$ = 5K, B = 1T



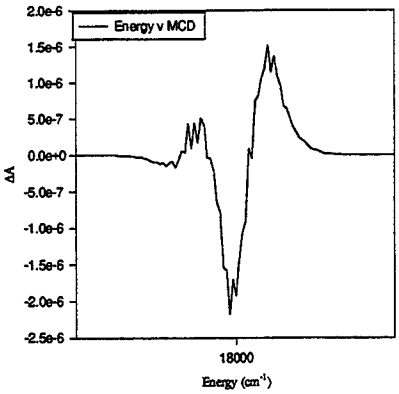
$T_{mag.}$ = 4K, B = 3T



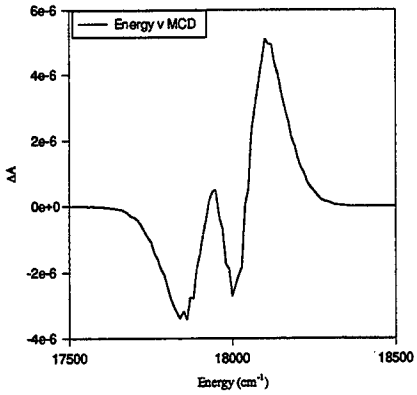
$T_{mag.}$ = 5K, B = 3T



$T_{mag.}$ = 4K, B = 5T



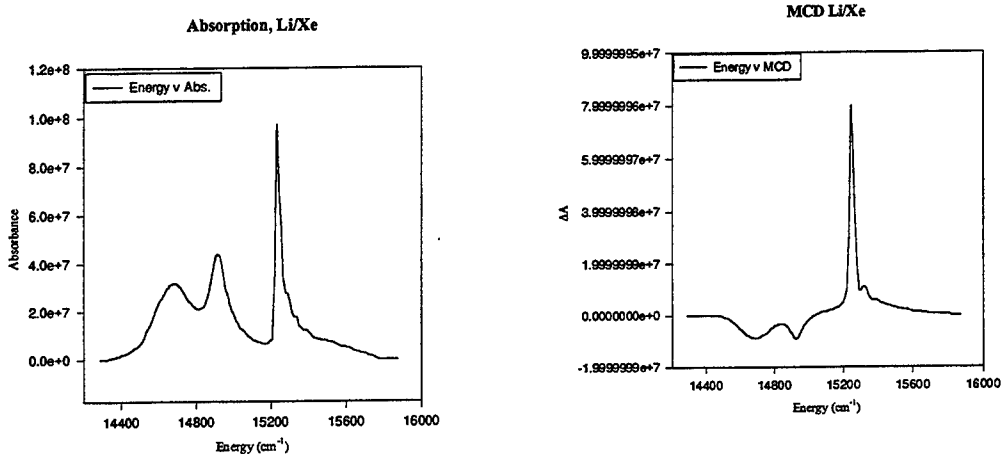
$T_{mag.}$ = 5K, B = 5T



appearance of the spectra at 6 and 8 K in the B = 10 T case at SO = $17cm^{-1}$, the sign change temperature for these conditions is around 7K. Although more study is needed to be sure, it appears that the rate at which the sign change shifts to lower temperatures as B increases is inversely proportional to the SO splitting.

Summary

The significance of this low temperature behavior is not yet clear. There is, to date, no experimental data indicating whether or not this is a real effect or an artifact of the simulation. The fact that no sign change is observed in the Li/Xe and Na/Xe does not negate the possibility that this is a real effect since the SO splitting, as calculated using the moment analysis method is negative, implying a negative sign change temperature. If the sign change proves to be a real feature of the low temperature behavior of the Na/Ar system, then our simulations have produced an extremely important piece of data, the governing physics of which is not yet understood. Further exploration of this effect is intended as a follow-on project, with the recognition that the temperature of the sign change is also related to the interaction potentials of the atoms involved. Although it has been amply demonstrated that the sign change temperature is not strongly affected by changes in trapping site, it might be instructive to study the effect of varying the strength of the rare gas matrix perturbation. Also if suitable interatomic potentials can be found, the generality of the low temperature sign change can be explored for systems other than Na/Ar. If the result is sufficiently general, it might also provide a way of estimating the SO splitting of a material before a moment analysis or other analysis is performed.

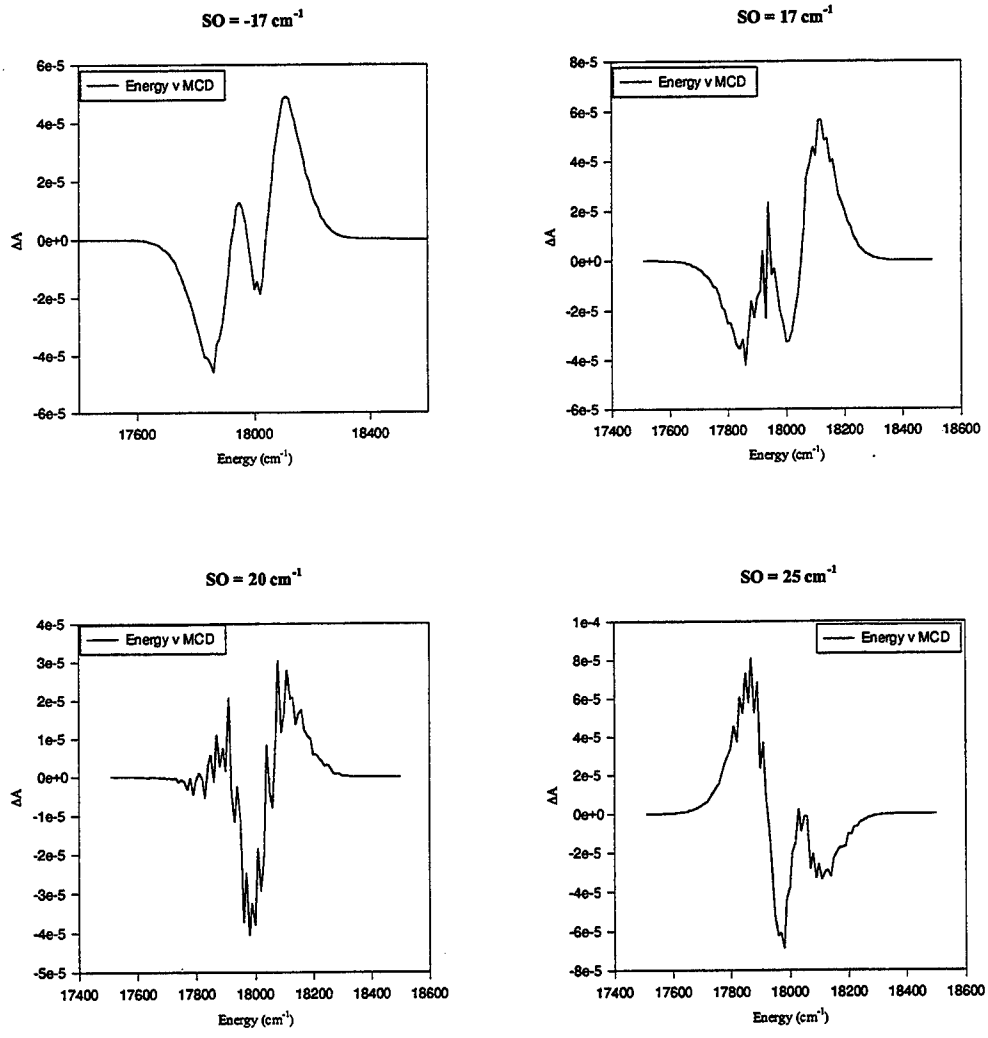


Experimental Li/Xe absorption and MCD Spectra [11]

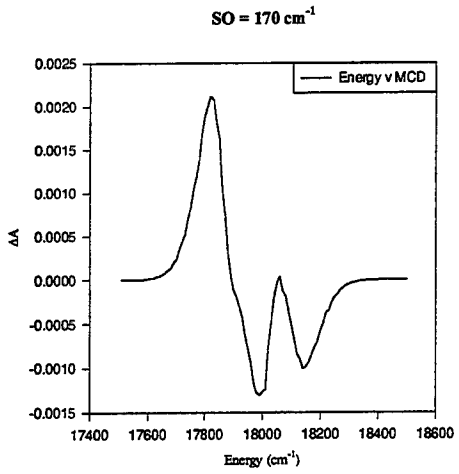
Spin-Orbit Coupling Studies

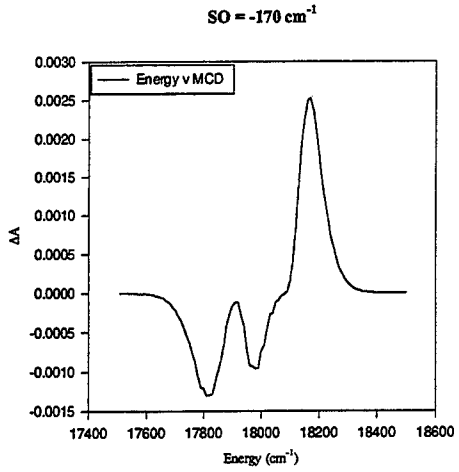
One of the reasons for doing the moment analysis studies is related to simulations run in spring of 1995, before

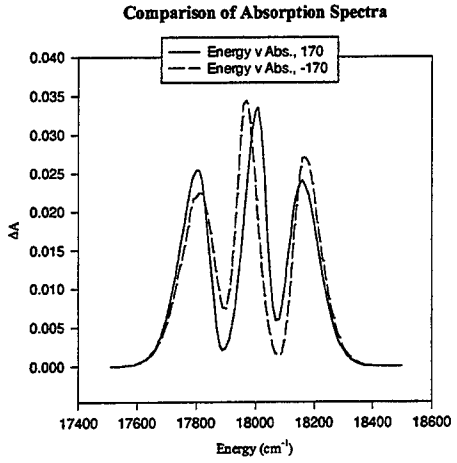
the start of the summer research. Since at that time we did not feel that the moment analysis was accurate, we decided to examine the effect of varying the spin-orbit splitting on the MCD spectra in our simulations. The M/Rg experimental data in the literature [8,9,11.] had MCD spectra characterized by two down "peaks" on the lower energy side and one sharp up peak on the high energy side. The SO splitting of this structure, as determined from the moment analysis was large and negative. It had been determined that setting the SO parameter in the MCD simulation to a large and negative value also produced a MCD spectra consisting of two down peaks and one up peak, going from high to low energy (see plot of the MCD spectra for SO = -170 cm^{-1}). The question was then, is this the only conditions under which the MCD spectra have this characteristic shape? Several simulations were run at a wide variety of SO splittings. It was found that the MCD spectra for *small positive* SO splittings also has two down peaks and one up peak with increasing energy. It was determined that the sign change occurs around SO



splitting = 20 cm^{-1}. This is consistent with relation 9 which predicts a sign change temperature of 9.8T for SO = 20 cm^{-1}.







The spectra for $SO = 17\ cm^{-1}$ and $SO = -17\ cm^{-1}$ have essentially the same character. The $SO = -17\ cm^{-1}$ MCD spectra closely resembles the $T_m = \infty$ spectra observed during the study of the limiting behavior of the MCD spectra, being very anti-symmetric about the energy barycenter. It is possible that the use of the random rotations feature, which was not yet operational when the SO = -17 cm^{-1} series was run, will remove some of this anti-symmetry. However, for the most part, the absorption and MCD spectra for $SO = 170\ cm^{-1}$ are a reversed version of the corresponding spectra for $SO = -170\ cm^{-1}$. The fact that the $SO = 17\ cm^{-1}$ and $SO = -17\ cm^{-1}$ MCD spectra are not mirror images is tentatively attributed to the small magnitude of the SO splitting as compared to the vibronic energy and the rare gas perturbation energy, as opposed to the $SO = 170\ cm^{-1}$ and $SO = -170\ cm^{-1}$

cases when the SO splitting is on the order of or larger than these other energies.

Future Work

There is still a great deal to be done in understanding all the nuances of the sign changes exhibited in the MCD spectra under various conditions. Many more simulations are required to completely map out the behavior of this intriguing feature of the Na/Ar MCD spectra.

One of the goals of the summer research that has not been realized is examination of the spectra at very small magnetic fields in the range from 0.1 to 0.01 T. This study was not completed due to the time that it takes to run a simulation for these parameters. For very small B, the perturbation due to the magnetic field is so small, that the statistical noise in the spectrum washes out the effect for simulations run over less than several hundred thousand configurations. It is anticipated that on the order of one million configurations will be needed for these simulations, so in the interests of time, it was decided to postpone the running of these simulations until after the summer research period.

Another project that was planned, but not completed is a test of the convergence of the simulated MCD spectra as a function of the number of configurations. This is to determine the minimum number of configurations needed for convergence.

One important project for the future is reconfiguring the MC program and lattice box to look selectively at the individual lattice vibrations believed to be the primary sources of the triplet splitting observed in the absorption spectra. It is hoped that this will make it possible to look at the effects of different coupling strengths between the Jahn-Teller active modes of the system. Past simulations of the MCD and absorption spectra of the M/Rg systems have been restricted to Jahn-Teller simulations of the case of equal coupling of the dominant vibronic modes acting on a predetermined site symmetry.[16]

Also planned for completion in the months following this summer research period are MC-MCD simulations of the multi-atom vacancy trapping sites using simulation temperature and magnetic temperature choices that more closely approximate the real thermal and vibronic conditions of the lattice. It is hoped that co-adding these spectra will give better predictions of the expected experimental spectra. "Real" temperature simulations are especially important in the asymmetric trapping sites such as the two-atom vacancy trapping site since it will determine if higher vibronic energy will allow these sites to deform to more symmetric geometries.

Conclusion

The MC-MCD simulations have revealed many interesting properties of the Na/Ar matrices. The details of this work has not yet been finalized, nor has the underlying physics been completely understood, however the framework has been laid. These MC-MCD simulations are especially important in that, contrary to the previously done simulations [16], they allow the choice of non-symmetric trapping sites and the allowed vibrations are not restricted to only the Jahn-Teller active vibrations.

Simulation based confirmation of the moment analysis method is an extremely important, never before accomplished theoretical milestone. This work will be extended by co-adding the simulated absorption and MCD spectra from several sites, generating more realistic spectra, to see if the moment analysis continues to extract the correct SO splitting.

Although the research completed this summer consisted of varying various MC simulation parameters, a goal of this research is the development of a parameter free model for the simulation of absorption and MCD spectra of atom-doped matrices. The purpose of varying the parameters is to get an idea of how each parameter affects the MCD spectra. Ultimately, it is hoped that the absorption and MCD spectra of a cryogenic M/Rg, and eventually a cryogenic M/H_2, matrix can be predicted with this MC-MCD simulation model before the experiment is done by putting the actual experimental parameters (Δ, B, T) into the simulation.

MODELING SPATIALLY EXTENDED SEMICONDUCTOR LASER DEVICES USING A SPLIT STEP FFT ALGORITHM

J. K. White

Arizona Center for Mathematical Sciences

University of Arizona

Tucson, AZ 85721

Final Report for:

Graduate Student Research Program

Phillips Laboratory

Sponsored by:

Air Force Office of Scientific Research

Bolling Air Force Base, DC

and

Phillips Laboratory

October 1995

MODELING SPATIALLY EXTENDED SEMICONDUCTOR LASER DEVICES USING A SPLIT STEP FFT ALGORITHM

J. K. White

Arizona Center for Mathematical Sciences

University of Arizona

ABSTRACT

Spatially extended semiconductor laser devices are modeled using a split-step FFT algorithm including nonlinear gain, phase, and dispersion dynamics. Nonlinear dispersion curves in a broad area Fabry-Perot laser are calculated and compared to experimental results. Preliminary results for master oscillator power amplifiers (MOPAs) are given.

INTRODUCTION:

Large aperture semiconductor laser amplifiers are extremely promising for applications combining pulsed and CW power levels $\sim 1-10$ W with high spatial coherence, as well as the usual advantages of semiconductor lasers. They are also useful vehicles for studying fundamental aspects of spatio-temporal dynamics, chaos, and pattern formation/recognition.

The basis for our model is the traveling wave laser rate equations [1]:

$$\begin{aligned}
\frac{n}{c}\frac{\partial F}{\partial t}+\frac{\partial F}{\partial z} &= \frac{i}{2k_0}\frac{\partial^2 F}{\partial x^2}+\frac{\Gamma}{2}[g(N)+i2k_0\Delta n(N)]F-\alpha F \\
\frac{n}{c}\frac{\partial B}{\partial t}-\frac{\partial B}{\partial z} &= \frac{i}{2k_0}\frac{\partial^2 B}{\partial x^2}+\frac{\Gamma}{2}[g(N)+i2k_0\Delta n(N)]B-\alpha B \qquad (1)\\
\frac{\partial N}{\partial t} &= D_f\frac{\partial^2 N}{\partial x^2}+\frac{J(x,z)}{ed}-\frac{1}{\tau}N-\frac{\epsilon_0 c g(N,\omega_0)}{2\hbar n\omega}(|F|^2+|B|^2)
\end{aligned}$$

where the symbols have their usual meanings. These equations are derived from Maxwell's equation where the usual slowly varying envelope approximation is made.

The gain and change in refractive index are functions of the carrier density N. They can either be calculated from the semiconductor Bloch equations in a quasi equilibrium regime [2, 3]

$$\left[i\frac{\partial}{\partial t}-(e_{e,k}+e_{h,k})\right]P_{\vec{k}}=(f_{e,\vec{k}}+f_{h,\vec{k}}-1)\left[d_{cv,\vec{k}}\frac{E_0(t)}{2\hbar}+\sum_{q\neq k}\frac{V_{|\vec{k}-\vec{q}|}}{\hbar}P_{\vec{q}}\right] \qquad (2)$$

or by using a more traditional linear gain model

$$g(N)=a(N-N_0) \qquad (3)$$

$$\Delta n(N)=bN=-\frac{aR}{2k_0}N \qquad (4)$$

where R is the linewidth enhancement factor.

A 100$\mathring{A}$ InGaAs quantum well is the usual material used in these simulations. Figure 1 shows a comparison of the theoretical nonlinear gain and refractive index to measured values. Small deviations in the nonlinear refractive index are a result of carrier contribution from the QW barriers and were not considered in this model.

The following scaling is used to non-dimensionalize the equations:

$$\begin{gathered}
\tilde{t} = \frac{ct}{nL}, \quad \tilde{z} = \frac{z}{L}, \quad \tilde{x} = \frac{x}{W} \\
\tilde{F} = \sqrt{\frac{aL\epsilon_0}{2\hbar\omega}}F, \quad \tilde{B} = \sqrt{\frac{aL\epsilon_0}{2\hbar\omega}}B, \quad \tilde{N} = \frac{N}{N_0} \\
\tilde{g} = \frac{g}{aN_0}, \quad \tilde{\Delta n} = \frac{\Delta n}{aN_0} \\
\tilde{D_p} = \frac{L}{2k_0W^2}, \quad \tilde{\alpha} = \alpha L, \quad \tilde{D_f} = \frac{nL}{cW^2}D_f \\
\tilde{J} = \frac{nL}{cN_0ed}J, \quad \tilde{\gamma} = \frac{nL}{c\tau}, \quad \tilde{\kappa} = \frac{LaN_0}{2}
\end{gathered} \tag{5}$$

where L is the device length, W is the device width, and the other symbols have their usual meaning. The scaled equations are (dropping tildes):

$$\begin{aligned}
\frac{\partial F}{\partial t} + \frac{\partial F}{\partial z} &= D_p\frac{\partial^2 F}{\partial x^2} - \alpha F + \kappa\Gamma[g(N) + \imath\Delta n(N)]F \\
\frac{\partial B}{\partial t} - \frac{\partial B}{\partial z} &= D_p\frac{\partial^2 B}{\partial x^2} - \alpha B + \kappa\Gamma[g(N) + \imath\Delta n(N)]B \\
\frac{\partial N}{\partial t} &= D_f\frac{\partial^2 N}{\partial x^2} + J(x,z) - \gamma N - \Gamma g(N)(|F|^2 + |B|^2)
\end{aligned} \tag{6}$$

ALGORITHM:

The equations (eqn. 6) are integrated on the characteristics using a split-step FFT solver with non-paraxial corrections. Computation is performed on a grid in x and z at time t. Following the propagation of the forward and backward field and the update of the carriers, the solution is then propagated forward in time along the characteristics (with appropriate gain dispersion added) to $t + \Delta t$ (see figure 2).

The non-paraxial operator splitting [4]

$$(F,B)(x,z \pm \Delta z,t) = \exp\left\{\frac{\imath \Delta z 2k_0 D_p}{2}\left[\frac{\nabla^2}{(\nabla^2 + k_0^2)^{1/2} + k_0}\right]\right\}\exp(\imath \Delta z \chi) \quad (7)$$
$$\exp\left\{\frac{\imath \Delta z 2k_0 D_p}{2}\left[\frac{\nabla^2}{(\nabla^2 + k_0^2)^{1/2} + k_0}\right]\right\}(F,B)(x,z,t)$$

$$\chi = -\alpha + \kappa\Gamma[g(N(x,z+\frac{\Delta z}{2},t+\frac{\Delta t}{2})) + \imath \Delta n(N(x,z+\frac{\Delta z}{2},t+\frac{\Delta t}{2}))] \quad (8)$$

which in the paraxial limit ($\nabla^2 << k^2$) gives the familiar operator splitting of

$$(F,B)(x,z \pm \Delta z,t) = \exp\left\{\frac{\imath \Delta z D_p}{2}\nabla^2\right\}\exp(\imath \Delta z \chi)\exp\left\{\frac{\imath \Delta z D_p}{2}\nabla^2\right\}(F,B)(x,z+\Delta z,t) \quad (9)$$

Since the field is propagated forward half a space step at the end of the previous time step and is then propagate forward another half step at the start of the next time step, it is possible to improve the computational efficiency by rearranging the splitting as

$$(F,B)(x,z \pm \Delta z,t) = \exp\left\{\imath \Delta z 2k_0 D_p\left[\frac{\nabla^2}{(\nabla^2 + k_0^2)^{1/2} + k_0}\right]\right\}\exp(\imath \Delta z \chi)(F,B)(x,z,t) \quad (10)$$

It is easiest to handle the linear operator by transforming into Fourier space

$$(F,B)(x,z\pm\Delta z,t)=\mathcal{F}^{-1}\exp\left\{\frac{\imath\Delta z 2k_0}{2}\left[\frac{-k^2}{(k_0^2-k^2)^{1/2}+k_0}\right]\right\}\mathcal{F}\exp(\imath\Delta z\chi)(F,B)(x,z,t) \quad (11$$

Before the nonlinear operator can be applied, N must be advanced in time. This is done using the first order implicit scheme (BTCS)

$$\begin{aligned}
N(x,z+\frac{\Delta z}{2},t+\frac{\Delta t}{2}) &= N(x,z+\frac{\Delta z}{2},t-\frac{\Delta t}{2}) \\
+\mu[N(x+\Delta x,z+\frac{\Delta z}{2},t+\frac{\Delta t}{2}) &- 2N(x,z+\frac{\Delta z}{2},t+\frac{\Delta t}{2})+N(x-\Delta x,z+\frac{\Delta z}{2},t+\frac{\Delta t}{2})] \\
+J(x)-\gamma N(x,z+\frac{\Delta z}{2},t-\frac{\Delta t}{2}) &- \Gamma g[N(x,z+\frac{\Delta z}{2},t-\frac{\Delta t}{2})] \qquad (12) \\
\times\Big[\frac{|F(x,z,t)|^2+|F(x,z+\frac{\Delta z}{2},t)|^2}{2} &+ \frac{|B(x,z,t)|^2+|B(x,z+\frac{\Delta z}{2},t)|^2}{2}\Big]
\end{aligned}$$

$$\mu=\frac{\Delta t}{\Delta x^2}$$

Finally some longitudinal diffusion is added to simulate the effect of the nonlinear gain dispersion by integrating the equation

$$\frac{\partial(F,B)}{\partial t}=D_g\frac{\partial^2(F,B)}{\partial z^2} \qquad (13)$$

using the first order explicit scheme (FTCS)

$$\begin{aligned}
(F,B)(x,t+\Delta t,z) &= (F,B)(x,t,z) \\
&+\mu[(F,B)(x,t,z+\Delta z)-2(F,B)(x,t,z)+(F,B)(x,t,z-\Delta z)] \quad (14)
\end{aligned}$$

$$\mu=\frac{\Delta t}{\Delta z^2}$$

Since $\Delta t=\Delta z$ on the characteristics, $\mu=1/\Delta z$.

BROAD AREA LASER:

Dave Bossert at the Phillips Laboratory has characterized the spatio-temporal dynamics of broad area devices by measuring the non-linear dispersion curves. These curves show the expected Fabry-Perot structure. The non-linear dynamics of the semiconductor material effects the intensity distribution along these curves. This is a clean, reproducible experiment which shows a connection between filamentation and temporal oscillations. It is a rigorous test of the time dependent spatially extend laser model.

Table 1 shows a listing of the device parameters used in the simulation. The nonlinear gain and refractive index are the same as in Figure 1. Nonlinear gain dispersion was chosen to give a representative number of dispersion curves but does not (at this stage) correspond to experiment.

parameter	*value*
λ	890 nm
W	100 μm
L	0.5 mm
d	100 $\AA$
n	3.5
a	$7.918\ 10^{-20}\ m^{-1}$
N_0	$1.2125\ 10^{24}\ m^{-3}$
D_f	$50\ 10^{-4}\ \frac{m^2}{s}$
τ	2 ns

Table 1. Table 1

Figure 4 shows comparison of the simulations with experiment for input currents of 130 mA and 400 mA. The mode spacing is approximately 90 MHz which agrees with both experiment and predicted mode spacing in a Fabry-Perot cavity.

Notice that for pumping near threshold the field intensity is concentrated in two discete transverse modes with an angular separation of approximately 2.5 degrees. At higher pumping the laser becomes chaotic and the dispersion curves are filled in with larger intensity concentrations centered about the nonlinear transverse modes.

MOPA:

Master Oscillator Power Amplifiers (MOPAs) hold the promise of high power, diffraction limited, compact coherent light sources. The aforementioned simulation was modified to run a semiconductor laser coupled to an amplifier. The goal of these simulations is to study coupling between the longitudinal modes of the master oscillator and the power amplifier. Such couplings are a result of low levels of feedback from the output facet and are suspected to be responsible for oscillations in the devices.

Working with Malcolm Wright of the Phillips Lab the first steps in simulating this device were taken. A stable $5\mu m$ laser was simulated and the output as coupled into a flared amplifier with no facet reflectivity. The resulting device relaxes to a steady state high power output with diffraction limited beam quality.

Further work on this device is in progress at the Arizona Center for Mathematical Sciences with the addition of week feedback and a DBR laser to replace the Fabry-Perot cavity.

---

1. H. Adachihara, O. Hess, R. Indik, J. V. Moloney, *J. Opt. Soc. Am. B*, **10**, 496, (1993).

2. H. Haug, S. W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors*, (World Scientific, Singapore 1990), Chap. 13, p. 241.

3. W. W. Chow, S. W. Koch, M. Sargent III, *Semiconductor Laser Physics*, (Springer Verlag, Berlin, 1994), Chap. 5, p. 158.

4. M. D. Feit, J. A. Fleck, *J. Opt. Soc. Am. B*, **5**, 633-640, 1988.

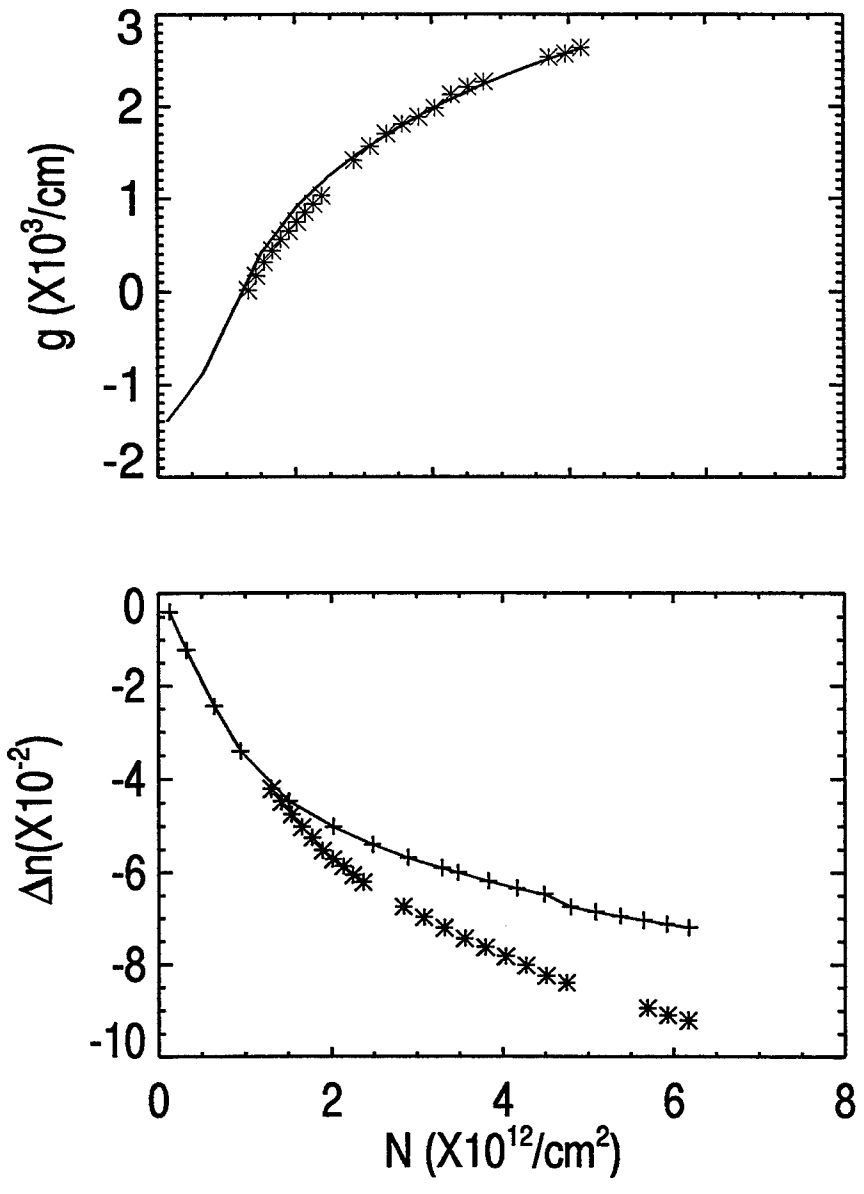
Fig. 1. Comparison of theoretical and measured nonlinear gain and refractive index.

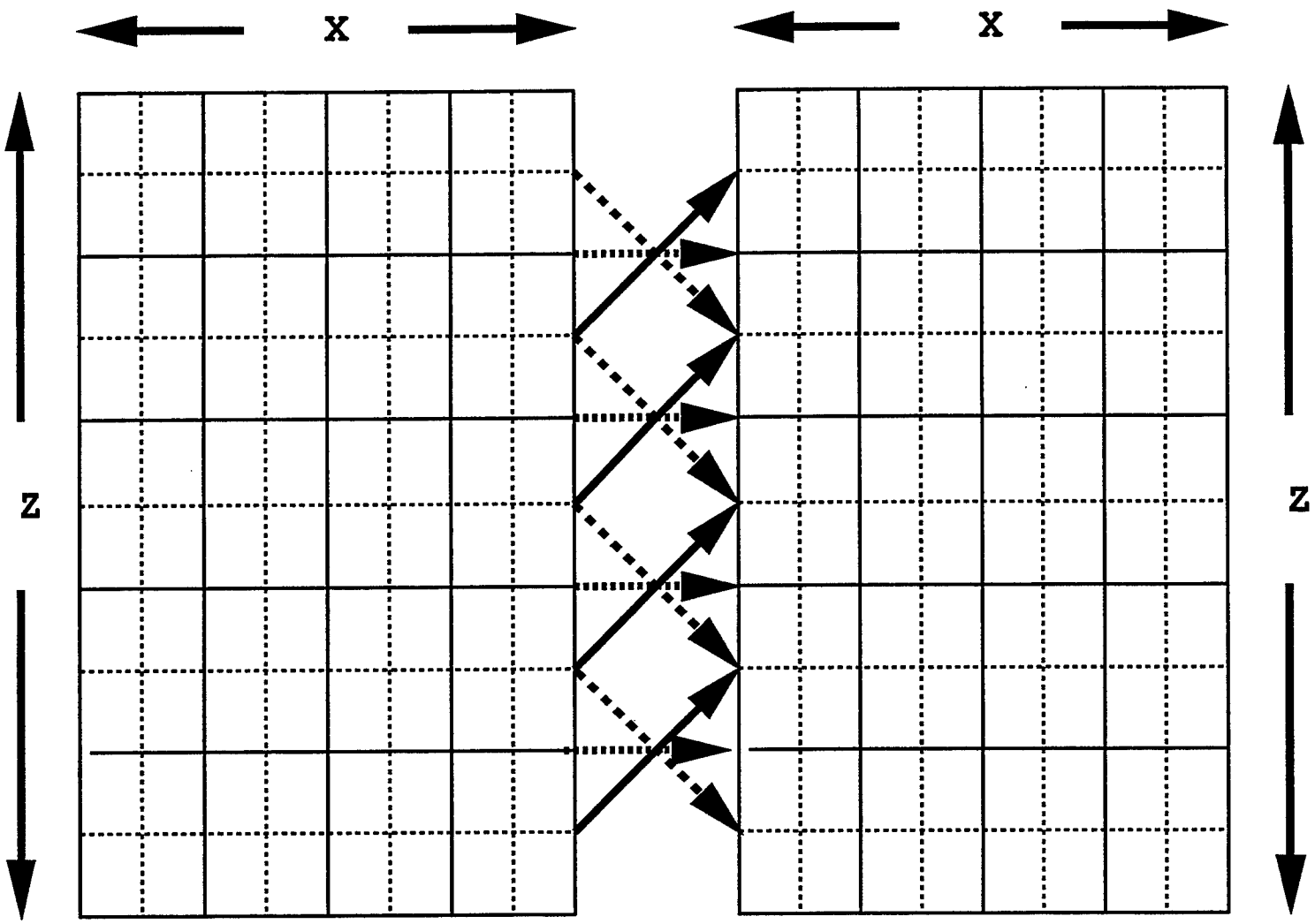
Fig. 2. Schematic of time evolution of F, B, and N.

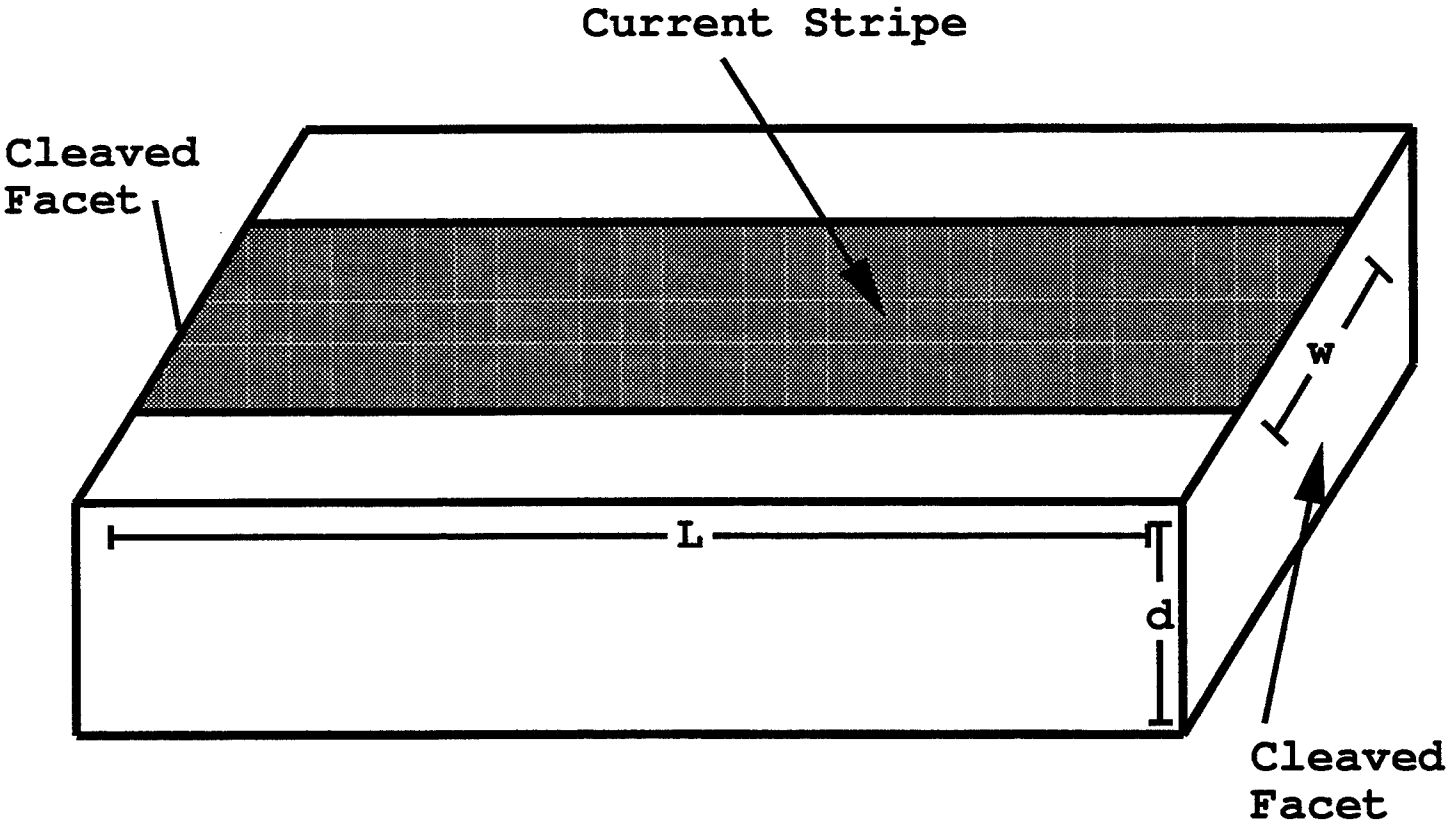
Fig. 3. Schematic of broad area laser.

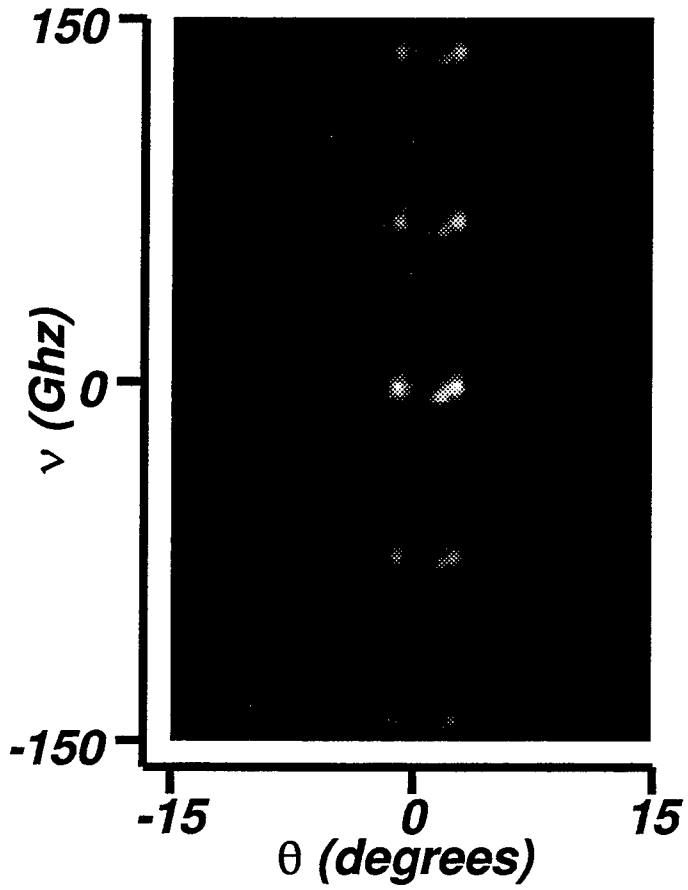
Fig. 4. Comparison of experimental and theoretical dispersion curves close to threshold (upper pair) and approximately four times threshold (lower pair). Experimental curves are on the left.

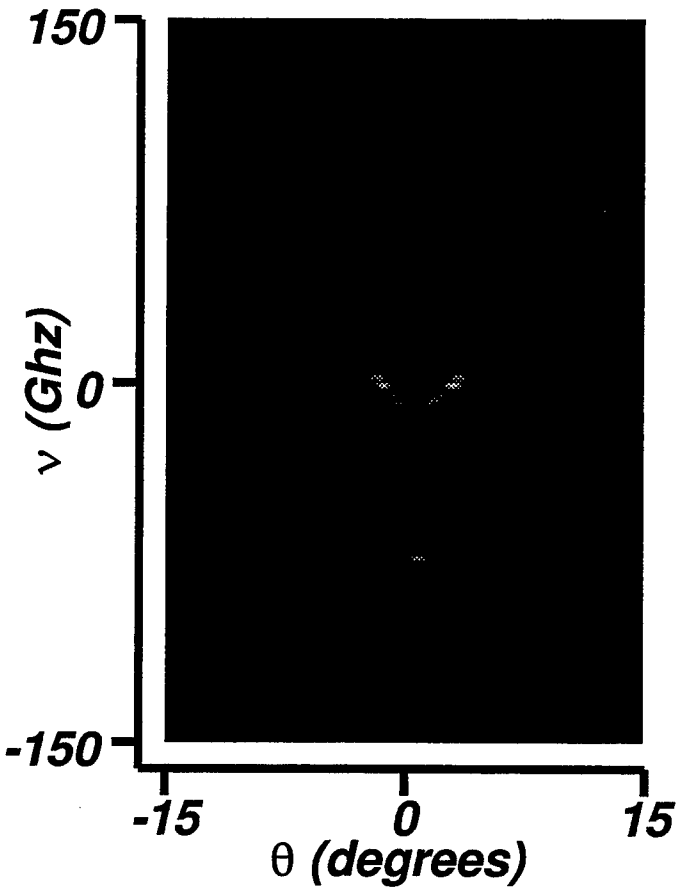
Fig. 5. Schematic of a Master Oscillator Power Amplifier (MOPA)

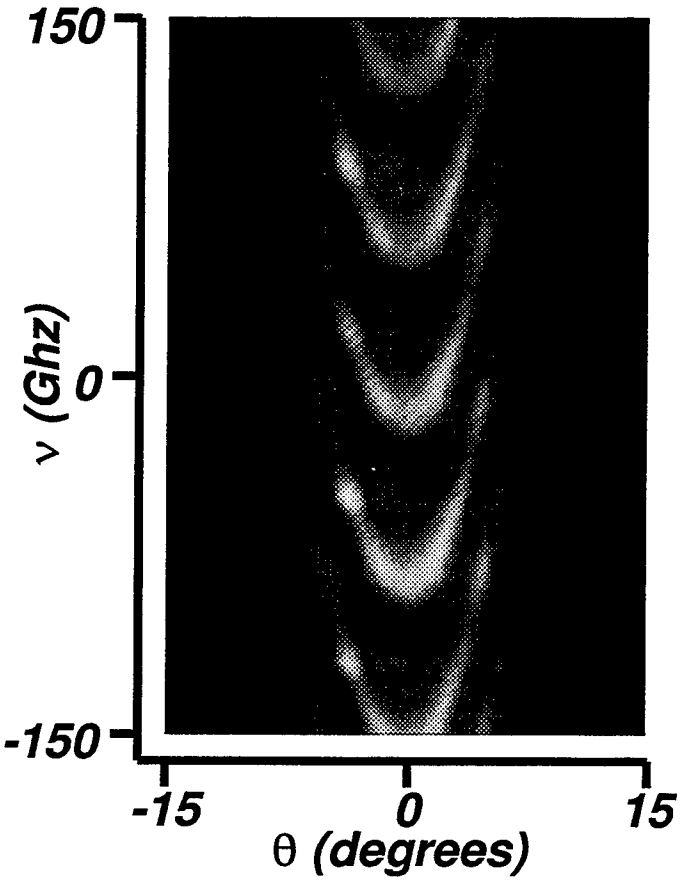

g (X10^3/cm)
3
2
1
0
-1
-2
Δn(X10^-2)
0
-2
-4
-6
-8
-10
0
2
4
6
8
N (X10^12/cm^2)

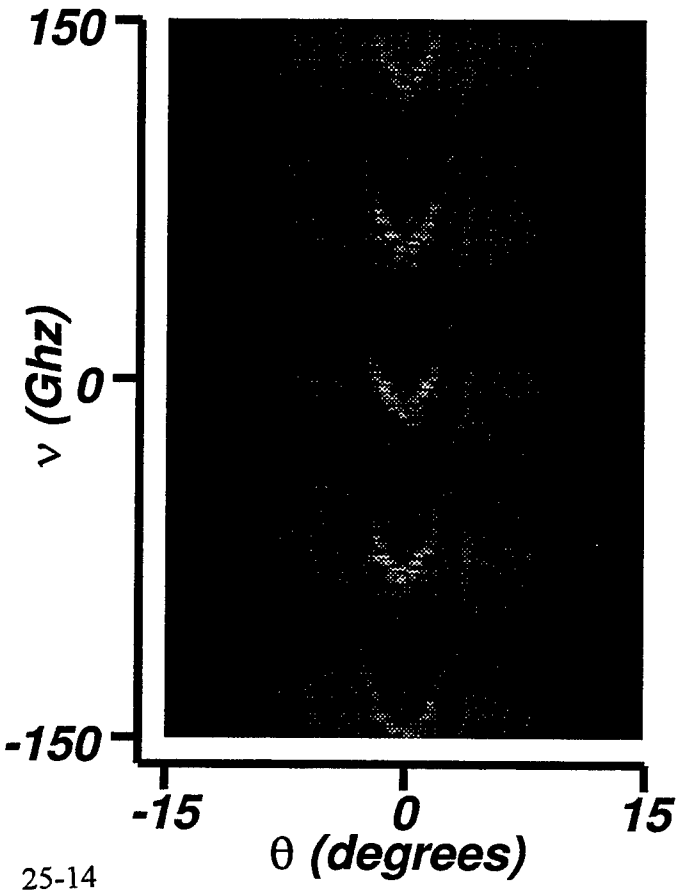

X
X
Z
Z
T
T + dT
Forward Field
Backward Field
Carriers

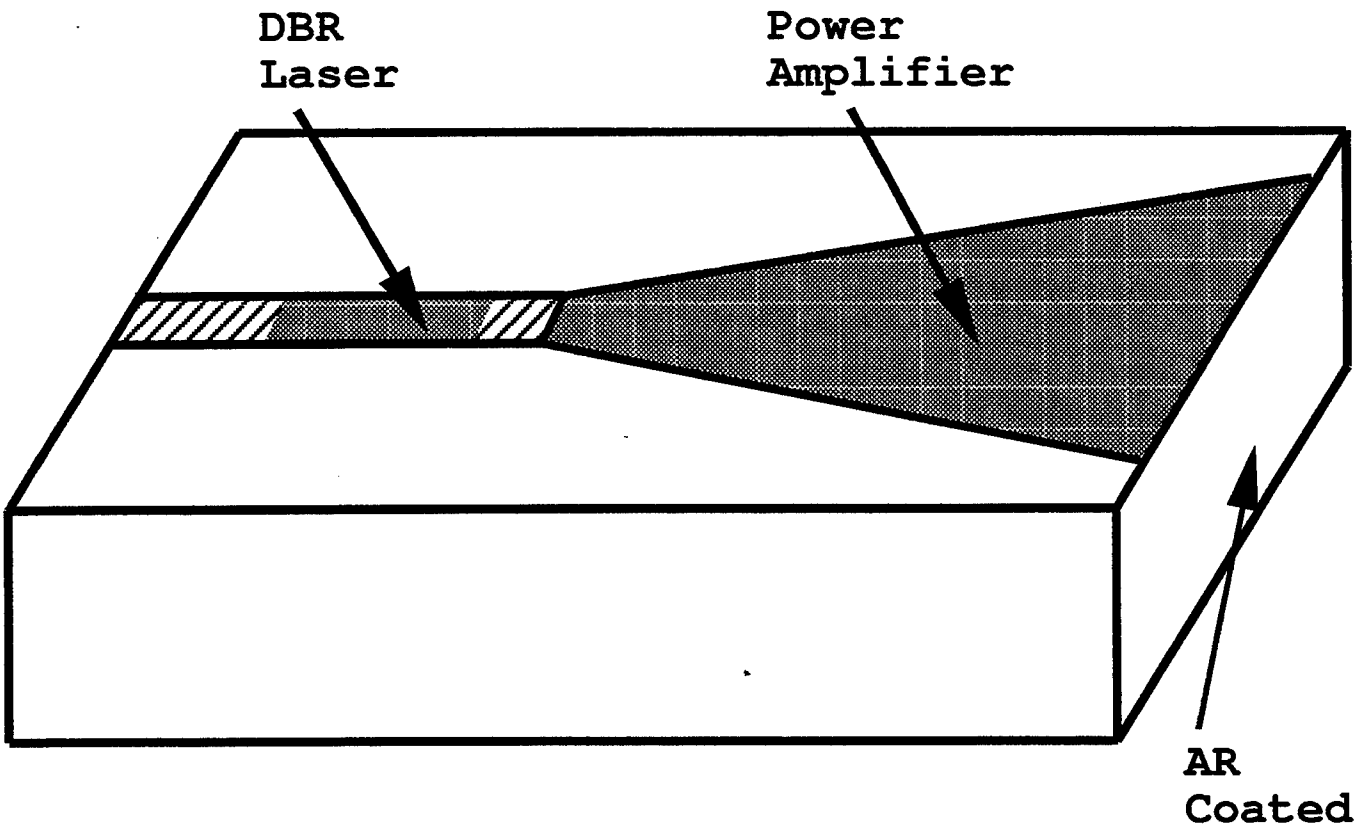

Current Stripe
Cleaved Facet
W
L
d
Cleaved Facet


150
ν (Ghz)
0
-150
-15
0
15
θ (degrees)


150
ν (Ghz)
0
-150
-15
0
15
θ (degrees)


150
ν (Ghz)
0
-150
-15
0
15
θ (degrees)


150
ν (Ghz)
0
-150
-15
0
15
θ (degrees)


DBR
Laser
Power
Amplifier
AR
Coated