LOAN DOCUMENT

PHOTOGRAPH THIS SHEET

AD-A261 992

DTIC ACCESSION NUMBER

LEVEL

1

INVENTORY

AFOSR-TR-93-0114

DOCUMENT IDENTIFICATION

Dec 92

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

DISTRIBUTION STATEMENT

HANDLE WITH CARE

SION FOR
NS GRA&I ☑
NC TRAC ☐
ANNOUNCED ☐
STIFICATION

TRIBUTION/
AILABILITY CODES
STRIBUTION | AVAILABILITY AND/OR SPECIAL
A-1

DISTRIBUTION STAMP

DTIC
ELECTE
MAR 10 1993
S C D

DATE ACCESSIONED

DATE RETURNED

93 3 10 005

~~93 3 4 074~~

DATE RECEIVED IN DTIC

93-05123

REGISTERED OR CERTIFIED NUMBER

PHOTOGRAPH THIS SHEET AND RETURN TO DTIC-FDAC

DTIC FORM 70A JUN 90

DOCUMENT PROCESSING SHEET

PREVIOUS EDITIONS MAY BE USED UNTIL STOCK IS EXHAUSTED.

LOAN DOCUMENT

UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1992

SUMMER FACULTY RESEARCH PROGRAM (SFRP) REPORTS

VOLUME 5A

WRIGHT LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 UPLANDER WAY

CULVER CITY, CA 90230-6608

SUBMITTED TO:

LT. COL. CLAUDE CAVENDER
PROGRAM MANAGER

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

BOLLING AIR FORCE BASE

WASHINGTON, D.C.

DECEMBER 1992

REPORT DOCUMENTATION

AGENCY USE ONLY	28 Dec 92	Annual 1 Sep 91 - 31 Aug 92

AND SUBTITLE

1992 Summer Faculty Research Program (SFRP)
Volumes 1 - 16

F49620-90-C-0076

Mr Gary Moore

0114

Research & Development Laboratoreis (RDL)
5800 Uplander Way
Culver City CA 90230-6600

MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/NI
110 Duncan Ave., Suite B115
Bldg 410
Bolling AFB DC 20332-0001
Lt Col Claude Cavender

SUPPLEMENTARY NOTES

DISTRIBUTION AVAILABILITY STATEMENT

UNLIMITED

ABSTRACT (Maximum 200 words)

The purpose of this program is to develop the basis for cintinuing research of interest to the Air Force at the institution of the faculty member; to stiumlate continuing relations among faculty members and professional peers in the Air Force to enhance the research interests and capabilities of scientific and engineering educators; and to provide follow-on funding for research of particular promise that was started at an Air Force laboratory under the Summer Faculty Research Program.

During the summer of 1992 185 university faculty conducted research at Air Force laboratories for a period of 10 weeks. Each participant provided a report of their research, and these reports are consolidated into this annual report.

SUBJECT TERMS

SECURITY CLASSIFICATION OF REPORT	SECURITY CLASSIFICATION OF THIS PAGE	SECURITY CLASSIFICATION OF ABSTRACT	LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1992

SUMMER FACULTY RESEARCH PROGRAM (SFRP) REPORTS

VOLUME 5A

WRIGHT LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES
5800 Uplander Way
Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Lt. Col. Claude Cavender

Program Manager, RDL
Billy Kelley

Program Administrator, RDL
Gwendolyn Smith

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Bolling Air Force Base
Washington, D.C.

December 1992

PREFACE

This volume is part of a 16-volume set that summarizes the research accomplishments of faculty, graduate student, and high school participants in the 1992 Air Force Office of Scientific Research (AFOSR) Summer Research Program. The current volume, Volume 5A of 16, presents part one of the final research reports of faculty (SFRP) participants at Wright Laboratory.

Reports presented herein are arranged alphabetically by author and are numbered consecutively -- e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

Research reports in the 16-volume set are organized as follows:

VOLUME	TITLE
1	Program Management Report
2	Summer Faculty Research Program Reports: Armstrong Laboratory
3	Summer Faculty Research Program Reports: Phillips Laboratory
4	Summer Faculty Research Program Reports: Rome Laboratory
5A	Summer Faculty Research Program Reports: Wright Laboratory (part one)
5B	Summer Faculty Research Program Reports: Wright Laboratory (part two)
6	Summer Faculty Research Program Reports: Arnold Engineering Development Center; Civil Engineering Laboratory; Frank J. Seiler Research Laboratory; Wilford Hall Medical Center
7	Graduate Student Research Program Reports: Armstrong Laboratory
8	Graduate Student Research Program Reports: Phillips Laboratory
9	Graduate Student Research Program Reports: Rome Laboratory
10	Graduate Student Research Program Reports: Wright Laboratory
11	Graduate Student Research Program Reports: Arnold Engineering Development Center; Civil Engineering Laboratory; Frank J. Seiler Research Laboratory; Wilford Hall Medical Center
12	High School Apprenticeship Program Reports: Armstrong Laboratory
13	High School Apprenticeship Program Reports: Phillips Laboratory
14	High School Apprenticeship Program Reports: Rome Laboratory
15	High School Apprenticeship Program Reports: Wright Laboratory
16	High School Apprenticeship Program Reports: Arnold Engineering Development Center; Civil Engineering Laboratory

1992 FACULTY RESEARCH REPORTS

Wright Laboratory

Wright Laboratory (cont'd)

Wright Laboratory (cont'd)

Wright Laboratory (cont'd)

VALIDATION OF A HYPERSONIC NONEQUILIBRIUM CODE FOR NOZZLE FLOW

Brian M. Argrow
Assistant Professor

School of Aerospace and Mechanical Engineering
University of Oklahoma
Norman, Oklahoma 73019-0601

Final Report for:
AFOSR Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

July 1992

VALIDATION OF A HYPERSONIC NONEQUILIBRIUM CODE FOR NOZZLE FLOW

Brian M. Argrow
Assistant Professor
School of Aerospace and Mechanical Engineering
University of Oklahoma

ABSTRACT

A computer code, has been validated for the external flow of hypersonic nonequilibrium air over a blunt body. The full Navier-Stokes equations with nonequilibrium chemistry and a five-species dissociation model are solved using a Roe flux-difference splitting scheme. Modifications for the computation of hypersonic nonequilibrium nozzle flows have been partially completed. Work is continuing on developing characteristic-based inflow and outflow boundary routines that allow for subsonic inflow and mixed supersoni -subsonic outflow conditions. Once boundary routines are completed, the code will be validated using data from the Boeing 30-Inch Hypersonic Shock Tunnel. Results are shown for an internal flow test of the flow solver assuming a perfect gas. Application of the second law of thermodynamics in the form of a numerical entropy production is also briefly discussed.

VALIDATION OF A HYPERSONIC NONEQUILIBRIUM CODE FOR NOZZLE FLOW

Brian M. Argrow

1. INTRODUCTION

The research effort associated with the development of the National Aero-Space Plane (NASP) in the United States and proposed hypersonic vehicles of other countries has generated a renewed interest in hypersonics. During this time of renewed interest, computational fluid dynamics (CFD) has undergone a rapid maturation process. Numerical algorithms have been developed to the point that some problems are considered solved to virtually any degree of engineering accuracy. The only limitations being the simplifying assumptions used in the equations of motion and boundary conditions. Examples of which are laminar, incompressible flows and some simple Euler flows.

One of the current goals of CFD research is to develop methods to solve the full Navier-Stokes equations to this same degree of accuracy. Hypersonic flow problems present a great challenge because most of the simplifying assumptions, such as that of a perfect gas, break down for high Mach number. The high temperatures encountered during hypersonic flight cause the gas (air) to dissociate and react, and at high enough speeds ionization also becomes important. Modeling these phenomenon introduces an additional level of complexity in the governing equations. In addition, the equations that govern the chemical and internal nonequilibrium rate processes are primarily empirical and are often based on questionable simplifying assumptions.

Validation of numerical codes is generally accomplished by comparison of computed results to experimental results. A substantial experimental data base exists for flows of moderate Mach numbers. This is not the case for hypersonic flows, however. Very specialized and expensive experimental facilities are required to produce hypersonic flow. Only a few specialized vehicles (publicly acknowledged) have operated in this regime and most of those are spacecraft that briefly pass through this regime and have no sustained hypersonic operations. As a result, there has been little demand for experimental hypersonic facilities. Additionally, the experimental methods of obtaining data are often inferential and sometimes questionable. Thus, the current state of affairs is that the experimental data base for hypersonic flows is relatively small.

The goal of this project is to adapt a code validated for the flow of hypersonic air in chemical nonequilibrium and vibrational equilibrium past a blunt body (Josyula et al., 1991) to that of hypersonic, nonequilibrium nozzle flow. The experimental data to be used for comparision was obtained at the Boeing 30-Inch Hypersonic Shock Tunnel Facility (Barr, 1992; Cassady et al., 1990; Barr and Syberg, 1990).

2. NUMERICAL METHOD

2.1 Governing Equations

The original computer code (designated as VIBEQL) was written to compute the two-dimensional or axisymmetric hypersonic flow field over a blunt body in dissociating air with vibrational equilibrium. For axisymmetric flow, the equations of motion are written as

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial}{\partial x}(\mathbf{F} + \mathbf{F}_v) + \frac{1}{r}\frac{\partial}{\partial r}(\mathbf{G} + \mathbf{G}_v) = \mathbf{H}$$

where x and r are the axial and radial coordinates and t is time. The vector of conserved variables is designated as $\mathbf{U}$. The inviscid portion of the flux vectors are designated as $\mathbf{F}$ and $\mathbf{G}$, where $\mathbf{F}_v$ and $\mathbf{G}_v$ are the viscous part of the flux vectors and $\mathbf{H}$ is the source vector containing the species source term and an azimuthal stress component. Complete details of this formulation can be found in Josyula et al. (1991).

VIBEQL solves the full Navier-Stokes equations using a finite-volume, Roe flux-difference splitting scheme for the inviscid terms. Derivatives in the viscous terms are evaluated using central differences and the solution is advanced in time using a predictor-corrector method with global or local time stepping.

The chemical source term is evaluated using the law of mass action and a five species dissociation model where the reactions are described by

$$N_2 + M \leftrightarrows 2N + M$$

$$O_2 + M \leftrightarrows 2O + M$$

$$NO + M \leftrightarrows N + O + M$$

$$N_2 + O \leftrightarrows NO + N$$

$$NO + O \leftrightarrows O_2 + N$$

Where N and O are the chemical symbols for nitrogen and oxygen and M is a collision partner that represents any of the species. The vibrational internal energy is assumed to in in equilibrium with the translational mode and is modeled using a simple harmonic oscillator.

2.2 Boundary Conditions

Boundary conditions associated with internal flow are generally more complicated and more difficult to implement than those for an external flow. Figure 1 shows the flow geometries and computational spaces with the appropriate boundaries indicated. The original external flow code has all freestream conditions specified at the supersonic inflow

and extrapolation is used at the outflow. The no slip condition is enforced at the wall with the assumption of a zero normal pressure gradient. A zero flux condition is enforced at the symmetry axis.

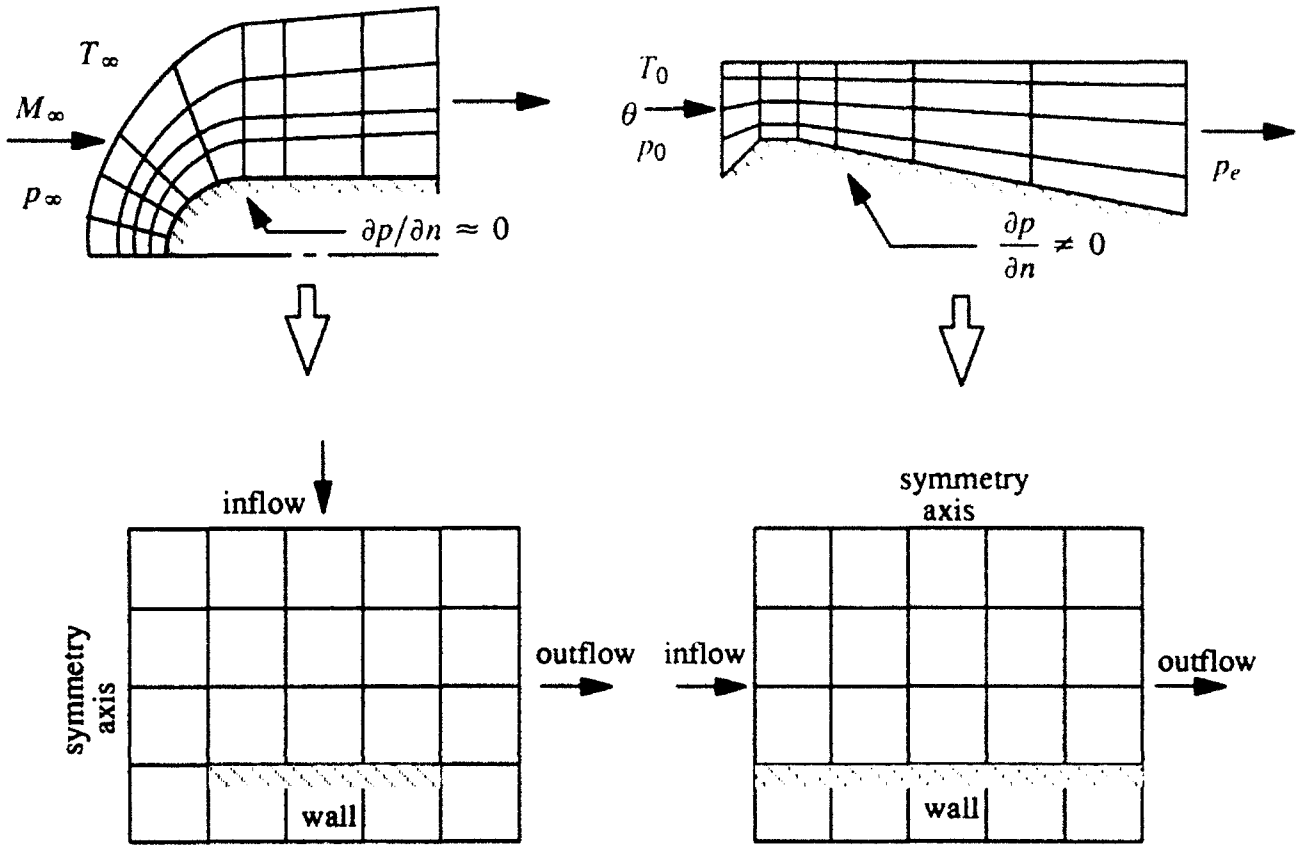


Figure 1 Boundary Conditions

These simple boundary conditions are not appropriate for internal flow. A subsonic inflow is required at the nozzle inlet, thus specification of the inflow flow variables is not permitted. For nozzle flow, the usual inflow condition is the specification of the stagnation temperature and pressure T_0, p_0 and the flow angle θ. The other flow variables are then determined by using an iterative reference-plane characteristic scheme (Marcum and Hoffman, 1987; Cline, 1981; Serra, 1972). This scheme allows conditions in the inflow plane at time $t + \Delta t$ to be determined from conditions at time t_0 by following the characteristic path as shown in Fig. 2 where ζ and η are the transformed axial and radial coordinates. The compatibility relation and characteristic equation that connects the interior points to the boundary is given by Cline (1981) as

$$dp - \varrho a du = (\psi_4 + a^2\psi_1 - \varrho a\psi_2)dt \quad \text{on} \quad d\zeta = \omega(u - a)dt$$

where ϱ is the density, a is the speed of sound, and the ψ and ω terms contain various derivatives with respect to ζ.

The iterative procedure begins by assuming a Mach number M, then computing temperature and pressure T and p from the isentropic relations

$$\frac{T_0}{T} = 1 + \frac{\gamma - 1}{2} M^2$$

$$\frac{p_0}{p} = \left(1 + \frac{\gamma - 1}{2} M^2 \right)^{\frac{\gamma}{\gamma - 1}}$$

where γ is the ratio of specific heats. The density ϱ is then computed from the state equation and u computed from the compatibility equation. Finally v, the radial velocity component is found using the flow angle. This process is repeated until the change in M is less than some prescribed amount.

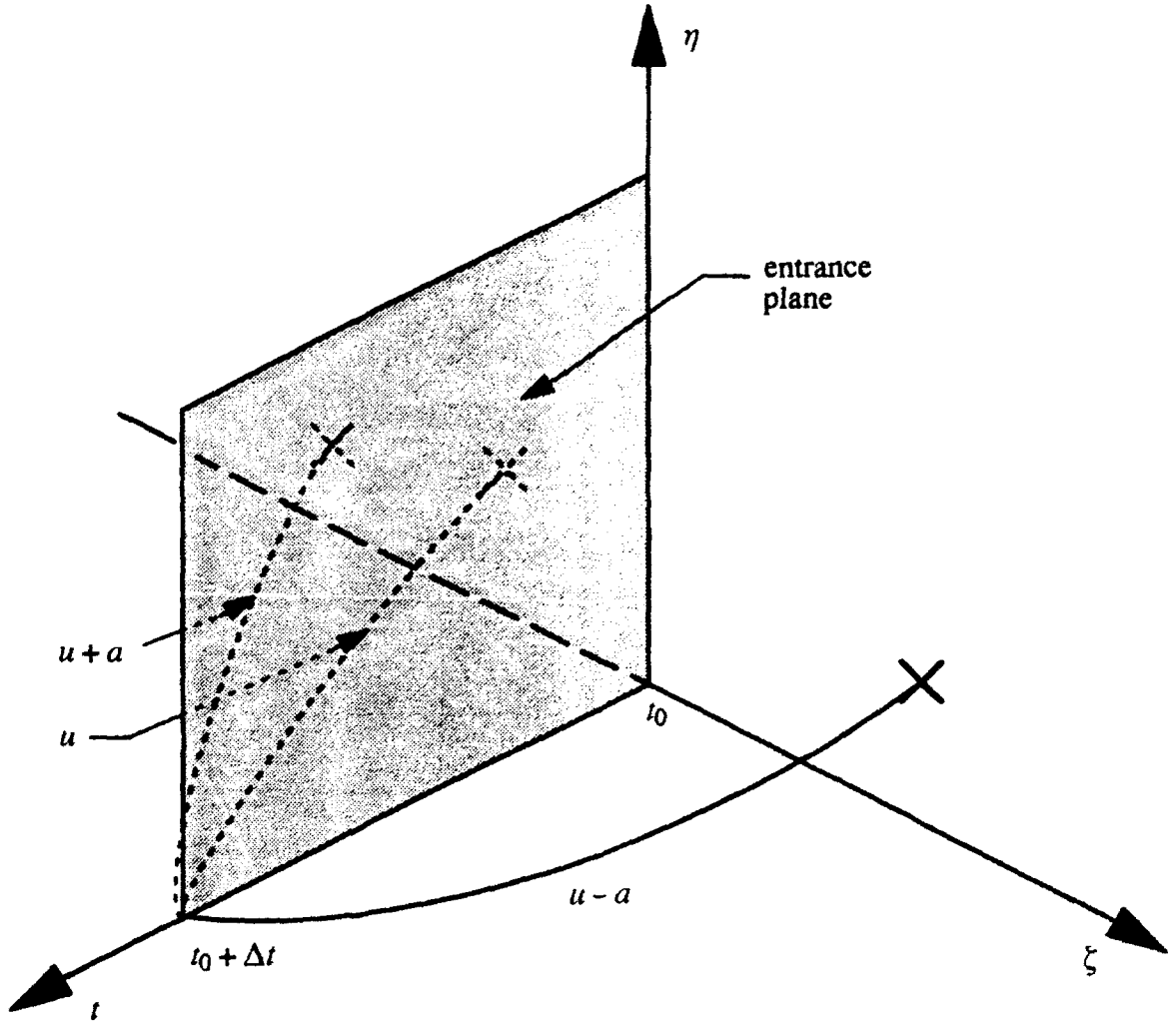


Figure 2 Physical representation of characteristic boundary condition.

The natural exit boundary condition for a nozzle is the specification of the back pressure p_e. Since the exit flow contains a subsonic boundary layer and a supersonic (hypersonic) core, this boundary condition also requires a reference-plane characteristic scheme similar to that at the inlet to solve for the subsonic portion of the exit boundary, while the supersonic portion can be determined by extrapolation. This type of boundary condition can be set up to allow for back flow that may occur if the boundary layer separates at the exit.

Because the nozzle flow is confined, the zero normal pressure gradient at the wall used in the external flow case is usually not a good assumption. The usual procedure for a no-slip wall is to solve the mass conservation and normal momentum equations using the interior point operator and fictitious wall points. The symmetry axis does not present any additional difficulties and is treated the same in both cases.

2.3 Code Revisions, Additions and Results

Several modifications have been made to VIBEQL. As shown in Fig. 1, boundary locations have been changed to agree with internal flow, however as of this writing, the boundary conditions remain those of the external flow. As mentioned earlier, work is continuing to incorporate the appropriate boundary routines before validation can commence.

A subroutine has been added to compute a quasi-one-dimensional initial solution surface. This was shown to aid in accelerating convergence. In addition, it will prevent computation of a "started nozzle" when such a solution is physically impossible for a given geometry and inlet conditions. This routine is also used to compute supersonic inflow conditions from specified stagnation conditions.

An interpolation and grid generating routine has been added to generate a simple grid from tabulated wall data. The routine allows clustering near the wall and at a given axial location. It has been used to generate the grid for the Boeing nozzle. No flow field computations have been attempted for the Boeing nozzle at this time. This simple routine will be replaced later with a more sophisticated routine, but is not a priority item at this time.

To test the flow solver before the appropriate boundary routines were completed, the nozzle geometry (Cline, 1981; Cuffel et al. 1969) shown in Fig. 3 was used for a perfect gas computation. Stagnation conditions T_0 = 300 K, p_0 = 70 psia and flow angle θ = 0° were specified at the inlet and forced to produce a supersonic inflow corresponding to the same area ratio as the designed subsonic inlet. Since the inlet is forced to be supersonic, the nozzle is actually performing as a diffuser. Time did not permit computing this flow field to convergence but the computation was advanced far enough to see the important flow features developing, including an oblique shock generated by the rapid compression turn near the entrance. Figures 4–6 show the pressure, temperature and Mach number distributions. The anomalous behavior at the symmetry axis indicates that this boundary condition is not being treated properly. Also note that the temperature contour values are not appropriately normalized. Real gas computations will begin as soon as we are satisfied that the simple boundary conditions produce realistic solutions.

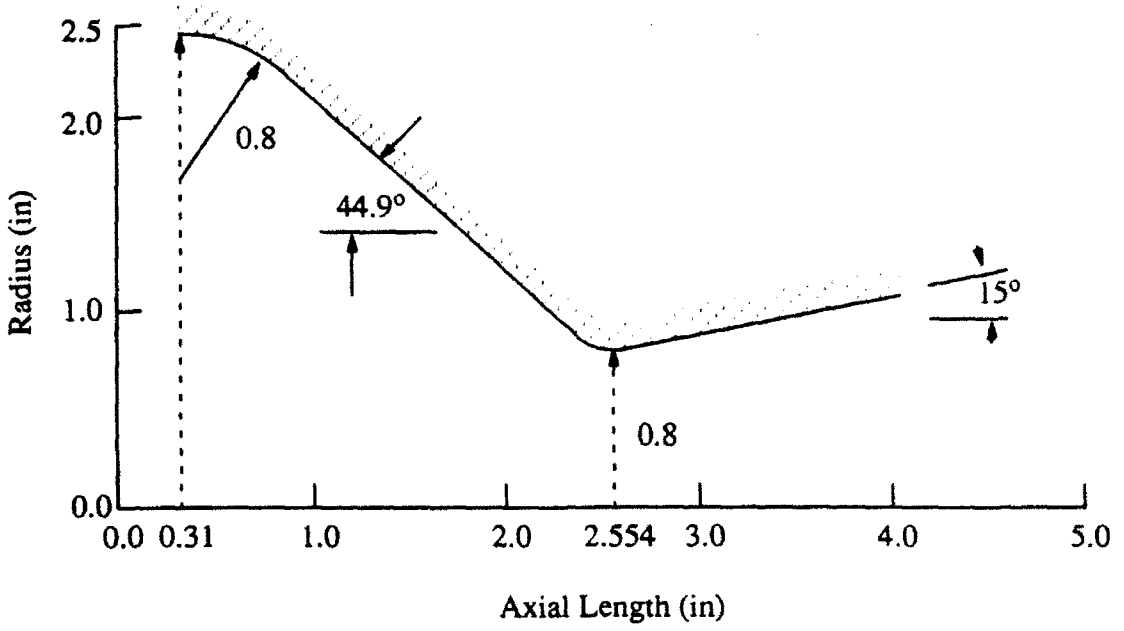


Figure 3 Nozzle geometry.

3. NUMERICAL ENTROPY PRODUCTION

Computation of hypersonic, axisymmetric blunt body flow often produces numerical anomalies near the symmetry axis due to the behavior of the eigenvalues of the inviscid portion of the equations of motion and from the transformation Jacobian. Josyula and Shang (1990) show that in some cases this anomalous behavior that appears as a bulge in the shock wave near the symmetry axis can be grid dependent, thus can be controlled with appropriate grid refinement in the region. For the Roe scheme, the entropy condition is enforced by cutting off the eigenvalues in a manner that is also dependent on the local behavior of the grid (Josyula et al., 1991).

In his original paper, Roe (1981) indicates that his scheme will violate the entropy condition for certain types of expansions. Although he does not present a method of dealing with this problem, he does indicate that an empirical cure should not be too difficult to formulate. This then leaves the question of what these empirical "fixes" actually do to the solution, even if their influence is confined only to the region where the anomalous behavior is occurring.

Borth and Argrow (1991) show how the second law of thermodynamics can be formulated in terms of a rate of entropy production and used to track artificial flow phenomena (numerical anomalies) and separate purely numerically generated phenomena from physical processes. The methods are developed for general application, but are only applied to very simple test flow fields. The same type of analysis should be applicable to the more complex flow phenomena asso-

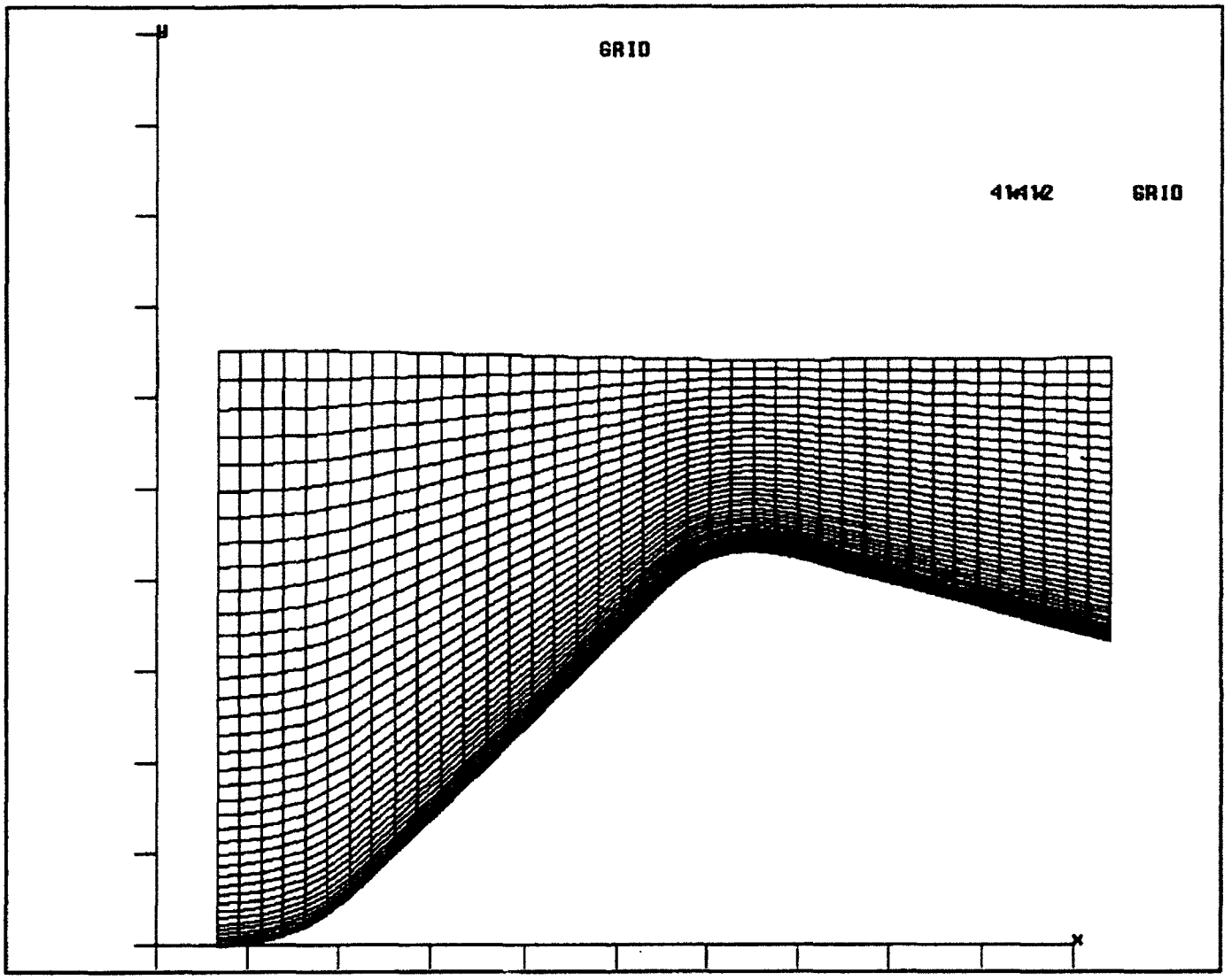


Figure 3 Nozzle/diffuser grid.

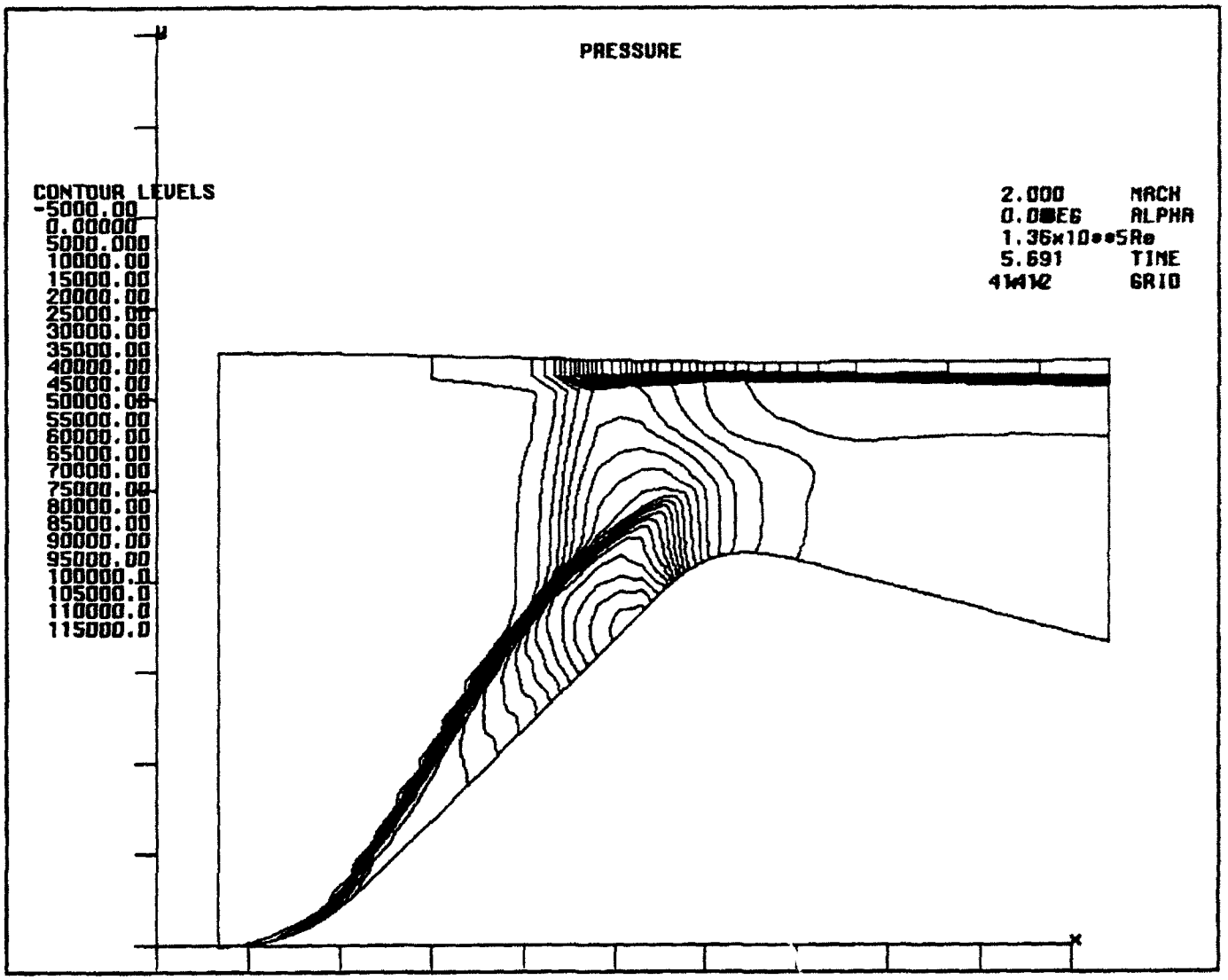


Figure 4 Nozzle/diffuser pressure distribution.

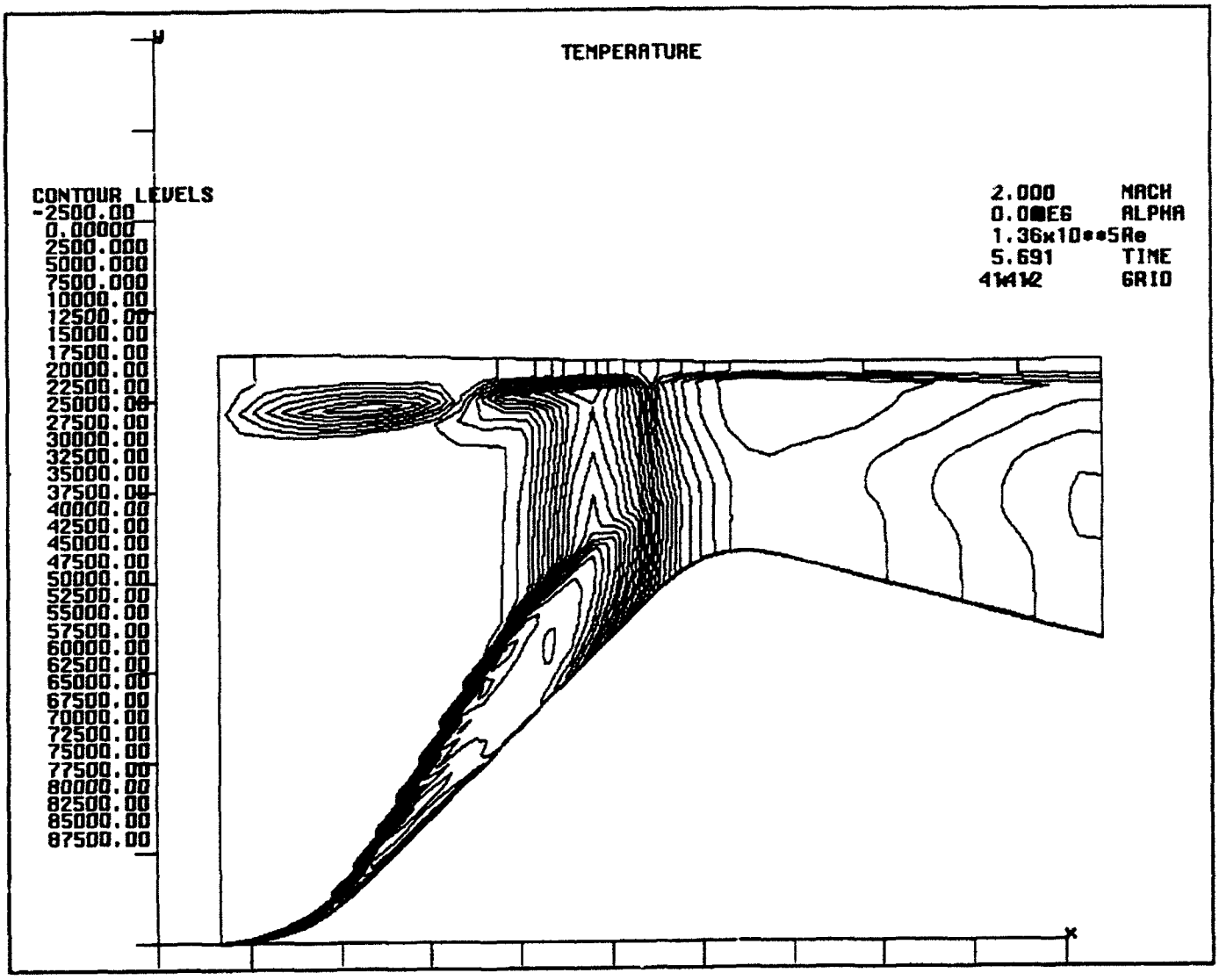


Figure 5 Nozzle/diffuser temperature distribution.

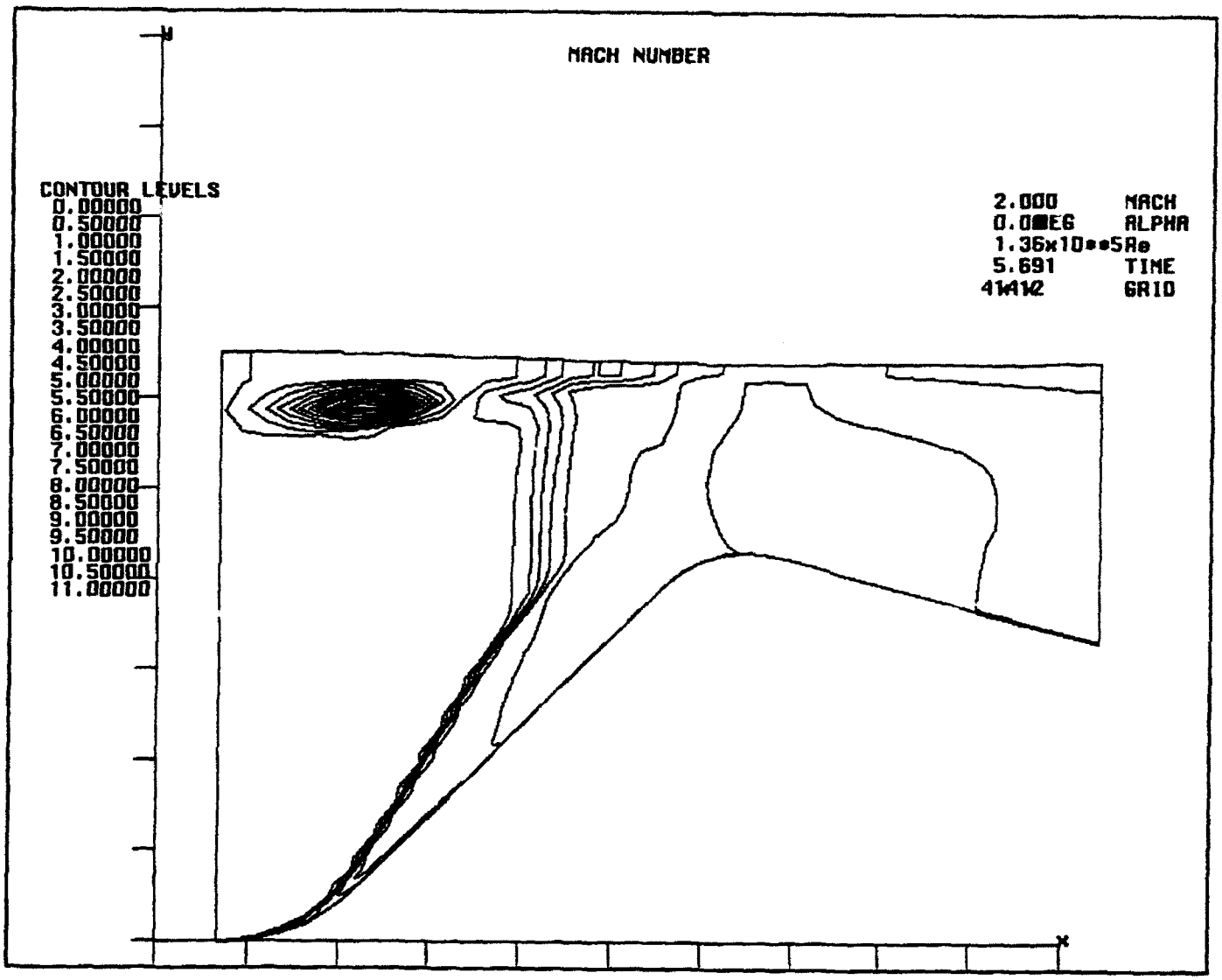


Figure 6 Nozzle/diffuser Mach number distribution.

ciated with nonequilibrium hypersonic flow. The method of accounting for the physical processes presented by Argrow et al. (1987) can be combined with that numerical entropy production to answer some of the questions of what is actually occurring. These methods can be used to quantify the numerical errors in terms of a physical entropy production. In addition, it is of interest to determine the contributions of the various modes of physical sources to the overall production rate.

4. SUMMARY AND CONCLUSIONS

A computer program to compute hypersonic, nonequilibrium external flow has been partially modified to compute hypersonic, nonequilibrium nozzle flow. Boundary routines are being written to employ a reference-plane characteristic scheme for the inflow and outflow boundaries, appropriate wall and symmetry axis conditions are also being implemented.

A sample case of a supersonic diffuser was run to test the flow solver for internal flow. Perfect gas results indicate that the computed solution shows all the complex flow features expected including resolution of an oblique shock near the entrance. Tests will continue in an attempt to use subsonic inflow conditions while the characteristic boundary routines are constructed. Once these routines are completed, validation tests using data from the Boeing 30-Inch Hypersonic Tunnel will begin to test the full real-gas capabilities of the code.

The production of numerical anomalies seen in some previous external flow cases was addressed in terms of applying a second law analysis. Because of time constraints, this was not pursued any further than some initial formulations.

Future work will incorporate the routines developed for VIBEQL into a later version of the code that contains a more complex real gas model, including thermal nonequilibrium (Josyula and Shang, 1992). The full development of a second law analysis will also be completed.

ACKNOWLEDGEMENTS

The author acknowledges the assistance of Mr. Eswar Josyula and Dr. Datta Giatonde of the Computational Aerodynamics Group, Wright Lab.

REFERENCES

1. Argrow, B. M., Emanuel, G. and Rasmussen, M. L., "Entropy Production in Nonsteady General Coordinates," *AIAA J.*, Vol. 25, Dec., 1987, pp. 1629-1631.

2. Barr, D. and Syberg, J., "Reacting Gas Experimental Data in Low Density Flow, Task IV, Numerical Validation of the Experimental Data Over Models," Flight Dynamics Laboratory, Wright Research and Development Center, WRDC-TR-91-3095, Jan., 1992.

3. Barr, D. and Syberg, J., "Reacting Gas Experimental Data in Low Density Flow, Task II, Numerical Validation of Nozzle Experimentation," Flight Dynamics Laboratory, Wright Research and Development Center, WRDC-TR-90-3008, March 1990.

4. Borth, C. J. and Argrow, B. M., "Evaluation of Entropy Production and Numerical Entropy Change in Flowfield Solutions," ASME Winter Annual Meeting, Dec. 1991, AES-Vol. 25/HTD-Vol. 191, *Second Law Analysis - Industrial and Environmental Applications*, pp. 101-107.

5. Cassady, P., Lieberg, S., Barr, D. and Syberg, J., "Reacting Gas Experimental Data in Low Density Flow Final Report for Task I; Boeing 30 Inch Hypersonic Shock Tunnel Nozzle Experimental Survey," Flight Dynamics Laboratory, Wright Research and Development Center, WRDC-TR-90-3010, May 1990.

6. Cline, M. C., "VNAP2: A Computer Program for Computation of Two-Dimensional, Time-Dependent, Compressible, Turbulent Flow," Los Alamos National Laboratory, LA-8872, Aug. 1981.

7. Cline, M. C., "Computation of Steady Nozzle Flow by a Time-Dependent Method," *AIAA J.*, Vol. 12, No. 4, April 1974, pp. 419-420.

8. Cuffel, R. F., Back, L. H. and Massier, P. F., "Transonic Flow-Field in a Supersonic Nozzle with Small Throat Radius of Curvature," *AIAA J.*, Vol. 7, No. 7, July 1969, pp. 1364-1366.

9. Josyula, E. and Shang, J. S., "Computation of Hypersonic Flowfields in Thermal and Chemical Nonequilibrium," AIAA 27th Thermophysics Conference, July 1992, AIAA Paper 92-2874.

10. Josyula, E., Gaitonde, D. and Shang, J. S., "Nonequilibrium Hypersonic Flow Solutions using the Roe Flux-Difference Split Scheme," AIAA 22nd Fluid Dynamics, Plasma Dynamics and Lasers Conference, June 1991, AIAA Paper 91-1700.

11. Marcum, D. L. and Hoffman, J. D., "Numerical Boundary Condition Procedures for Euler Solvers," *AIAA J.*, Vol. 25, Aug. 1987, pp. 1054-1062.

12. Roe, P. L., " Approximate Riemann Solvers, Parameter Vectors, and Difference Schemes," *J. of Computational Physics*, Vol. 43, 1981, pp. 357-372.

13. Serra, R. A., "Determination of Internal Gas Flows by a Transient Numerical Technique," *AIAA J.*, Vol. 10, May 1972, pp. 603-611.

MOTION SEGMENTED OBJECT IDENTIFICATION USING 1-D SIGNAL ANALYSIS AND A HETEROASSOCIATIVE COMPLEX NEURAL NETWORK

Abdul Ahad S. Awwal
Assistant Professor
Department of Computer Science & Engineering

Wright State University
Dayton, Ohio 45435

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D. C.

September 1992

MOTION SEGMENTED OBJECT IDENTIFICATION USING 1-D SIGNAL ANALYSIS AND A HETEROASSOCIATIVE COMPLEX NEURAL NETWORK

Abdul Ahad S. Awwal
Assistant Professor
Department of Computer Science & Engineering
Wright State University

Abstract

A complete and novel pattern recognition system is proposed which utilizes techniques of motion analysis coupled with a complex associative memory in a neural network architecture. Motion was used for segmenting the scene, obtaining a continuous edge information, and calculating chain code data from a sequence of moving images. The FFT of the sampled edge co-ordinates was used as an input to a complex inner product associative memory. The complex inner product associative memory model was formulated based on an efficient optical amplitude modulated phase only filter to serve as the decision making module. An adaptive threshold was used to indicate a match or identify a mismatch. The research has pointed out new directions for tracking, use of motion analysis in making a library of templates which may yield predictable matching behavior, use of other cues to reduce the search space, and other problems for segmentation using motion for further investigation.

MOTION SEGMENTED OBJECT IDENTIFICATION USING 1-D SIGNAL ANALYSIS AND A HETEROASSOCIATIVE COMPLEX NEURAL NETWORK

Abdul Ahad S. Awwal

1 Introduction

Two of the important areas of motion analysis [1, 2, 3] are identification and tracking of moving objects. An image in motion conveys much more information then when it is in a static condition. For example a moving pattern can be easily extracted from the background and then more attention can be given to the matching process since it automatically gets rid of the background which basically can be classified as noise. Thus the recognition becomes much simpler. Using motion more information can be gathered about an object. Such information can be used to eliminate unplausible matches: for example, by matching an object recorded from several viewpoints or from several consecutive frames of the same viewpoint one is able to reconfirm a match or eliminate an incorrect hypothesis. Tracking is continuous identification of the object even when it is undergoing different affine transformation due to motion.

2 Problems in Motion Analysis

Application of motion analysis may be found in medicine, tomography, autonomous navigation, communications, television, video conferencing, dancing and choreography, meteorology and so on. Motion analysis could be classified based on the type of motion: 1-D, 2-D or 3-D motion. Again two types of classification could be based on the method of attacking the problem of motion analysis: feature based or optical flow based. In an optical flow based approach the velocity is estimate ' from the apparent velocities of the brightness pattern. Yet another classification could be based on the desired parameters to-be-estimated: structure, shape, aspect angle, velocity, direction, instantaneous displacement (tracking), identification of the moving object. Other problems such as motion correspondence, where correspondence between two or more consecutive frames are established; (examples are corner detection, correlation based pixel matching,) and motion segmentation, where objects of coherent motion are grouped or extracted, are more global problem that are addressed or solved as a part of the above

problems or are active areas of research by their own merits. Motion correspondence allows one to estimate the velocity of moving object by tracking the points through a sequence of images. Different cases of the above could also be studied such as object motion with static background, both object and camera in motion, object motion with or without occlusion. This work was initially limited to object motion with static background and without occlusion.

The following are some of the goals of motion analysis: object identification, object tracking, Scene segmentation, Shape or structure from motion, Aspect angle from motion, Velocity from motion, Modelling of motion, and Modelling of moving objects.

3 Motivation

Most of the two dimensional signal processing applications are plagued by the humongous amount of data present in the images. In a real world, objects are essentially three-dimensional whereas images are only two-dimensional. More interestingly, although images contain two-dimensional information, because they are real they can be represented as one-dimensional complex data. Then necessary features may be extracted from these data to simplify the process of matching, tracking, identification etc.

Fourier descriptors provide a compact one-dimensional representation of two-dimensional information using the external edges of an image. From pattern recognition literature it is known that for correlation discrimination, edges are found to be a more important factor than the image gray levels. The two dimensional pattern matching using optical correlation techniques mainly rely on the edges to differentiate between images. For example in optical phase only filters [4], the edges of images are extracted (as a result of high pass filtering done by the phase only filter). Edges also govern the phase variation more vigorously. Even some matched filtering technique only matches the edges of one image to the edges of the template resulting in higher correlation and discrimination.

Now digitally, the two-dimensional Fourier transform operation is expensive. However, the above discussion points out that although optically Fourier transform can be done in a very fast and and inexpensive way using simple lenses, there is a tendency amongst researchers in the optics community to enhance correlation peak by taking only edge information. Thus for digital, one may just consider edge information. The advantage for digital is the gain of tremendous speed by reducing

the 2-D information to 1-D and enhanced performance. Thus it may be a practical way to reduce the information burden. Critics may argue that edges are more sensitive to noise etc., however, motion will be used not only to reduce information but also to reduce noise. Combination of motion analysis tools with those of signal processing may thus make the recognition process much easier.

A second important part of recognition is the decision making process. Usually, a sequential matching followed by thresholding is used to search for the presence of an object in a library of templates and then to decide whether it matches with one particular one. In a sequential way practically all the templates have to be brought one after another and then matched with the acquired object. The matching process can be speeded up by manifold if an associative memory processor [7] is employed instead of a sequential memory. In an associative memory, a parallel match of the input with all the stored patterns are executed and the match is also identified in parallel. A further important part of the associative memory is the nonlinearity associated with each elementary neurons. This very important element not usually emphasized by the neural network community warrants serious investigation.

The present research effort has several goals. A broad goal is to use an algorithm that could be executed in real time with minimal hardware. The other goal is to use motion to simplify or enhance the pattern recognition system. To these effect, motion is used to reduce the amount of data to-be-processed by segmenting the moving object from the background. Features are extracted in terms of Fourier descriptors and then the matching is performed by means of a simple heteroassociative memory. The complex associative memory described in this report is proposed for the first time.

4 Fourier Descriptors

Fourier descriptor is defined as the 1-D Fourier transform of a complex sequence formed by the boundary pixels of an object outline or shape taken in a consecutive fashion. It has been argued that edges of an image is the main reason why two objects can be discriminated. Let $f(k)$ be the complex sequence representing the boundary pixels of an edge image, then the Fourier descriptor is expressed as the following summation,

$$F_n = \sum_{k=0}^{N-1} f(k) e^{-(j2\pi/N)nk} \qquad n = 0, 1, ..., N-1 \qquad (1)$$

For the Fourier descriptors to work in our model, a zero padding is necessary before taking FFT.

5 Motion for Segmentation

This is a very important part of the whole recognition process. For our purpose, we used images acquired by a camcorder. Two sequence of images are chosen to form an absolute difference image. This provides a region of interest, from which local edge operators such as prewitt combined with thresholding, and median filtering, yields an edge image of the original object. Then the chain code of this edge image is obtained. This process works successfully if several constraints are satisfied. For example, if the camera was steady, and smearing caused by the image is minimal. Sometimes, the algorithm failed because the object recorded by the hand held camera apparently moved backwards exhibiting large smearing effect due to camera motion. This distorted the edges completely. A way to overcome the smearing effect is also proposed later.

The selection of threshold is very important in the segmentation process. The actual threshold value was obtained by trial and error. Note that there are several thresholds that could be used. Our experiment with limited number of data suggested that the threshold value was not correlated well with those suggested by different histogram program. More dynamic thresholding techniques may be developed. The filling of discontinued edges are done now automatically. Since this is a prerequisite for the chain code detection routine to work.

The chain coded edge image is then converted to border pixel data. It may be noted that the border pixels may be obtained in a much faster way. For example, using horizontal or vertical scan lines. However, the sequence of co-ordinates that are obtained are not consecutive locations on the edge image. Although one can take Fourier transform of the resulting sequence and generate the image back after an inverse transform, the problem is when the image is rotated, then the sequence is totally altered. There does not exist a simple relationship between the complex sequence of the original object boundary and its rotated version, unless samples are taken in a consecutive fashion.

The other advantage of using motion is that instead of a single image, we work with multiple images. Thus if a sequence of only two images does not work we are free to choose from other image sequences, since we acquire thirty frames per seconds.

6 Sampling the Chain Data

For a chain-coded image the consecutive pixel locations on the edges are not equidistant. For example, two neighboring pixels on the diagonal is at distance of $\sqrt{2}$ where as those that are apart horizontally or vertically are one pixels apart. Often images of varying sizes are acquired resulting in chain code sequences of varying length. As a result the number of samples that are obtained are not equal for different sized objects. However, to make the processing faster usually a 1-D FFT algorithm is applied which requires that the sequence be either a power of two or some other number. Thus the sequence may have to be resampled to make it a power of two. Wallace and Mitchell [5] have proposed sampling the contour with equal interval. Now usually when the image is sampled along the contour it may give points which are fractions when represented digitally it may not represent an exact point. In the current approach, we choose two options, depending on the size of the contour. When the size of the contour is small, we take an equal interval sample; when the size of the contour is more than the required number (N), then we propose an error diffusion technique which selects discrete points closest to the equi-interval points. The total error is minimized by the error diffusion technique. Thus a set of discrete points on the edge image are obtained.

Alternate Suggestion to sampling: A similar approach to sampling may be taken in the frequency domain. In this approach, a power of two FFT may be taken after zero padding and then the spectrum may be interpolated, scaled and sampled to get a spectrum of desired size. If this is practical, then the time in sampling the original data could be saved.

7 Calculation of the Fourier Descriptors

Next a 1024 point FFT is taken on the stored border pixels data after padding it with 512 zeros. Wallace et al. [5] have provided a normalization procedure which scales, rotates the edge data and then fixes the same starting point so that when the input is compared with the template they are at the same reference. Next the comparison is made using a mean square error method. Since in their approach mean square error is used slight variation of scale, rotation or starting point will introduce large comparison error. In our case, we propose to compare the input with the stored templates using a newly proposed associative memory.

8 A Complex Inner Product Associative Memory

A novel complex inner product associative memory (CIPAM) neural network model is proposed for associative pattern identification from noisy input data. Simulation of the model for object recognition problem is provided to illustrate the usefulness of the proposed implementation. An adaptive thresholding scheme is developed which permits one step matching. The inherent parallelism of this neural network architecture may be exploited for a parallel optical or digital array processor implementation and hence enhance the performance of the pattern recognition system.

The major advantage of the proposed architecture is the fact that by only taking an inner product, we can detect the match with an intelligent adaptive threshold formulation. In fact, this does not require us to perform processing for scaling, rotation and translation and variation in the starting point. Next the storage prescription, input representation and pattern retrieval mechanism of the complex content addressable memory model are presented.

9 The Storage and Matching prescription

The complex associative memory model which is an extension of our previously proposed trinary associative memory [8, 9, 10] model consists of a collection of stored (learned) vectors represented in complex form. The input vector V^{in}, of length L is represented in complex form. For a 3-D object the input is the Fourier descriptor of its edge data. The storage prescription is a 1-D variant of our earlier proposed amplitude modulated phase only filter [11]. This filter has already received attention from the optical pattern recognition community as optimum and more efficient than the phase only filter [12]. Note that the 1-D variation of the 2-D filter is more advantageous for saving both space and time.

The edge data is obtained from sampling the chain coded image. Then the data is zero padded to make a double length sequence. If $F(n)$ is the discrete Fourier transform of this sequence and $exp(j\theta(n))$ is the phase part of $F(n)$, then the kth stored template is

$$V_n^{st,k} = \frac{exp(-j\theta(n))}{|F_k(n)| + a} \tag{2}$$

Here $0 < a$ and $|.|$ denotes a modulus operation. The parameter a is useful in (i) overcoming the indeterminate condition and (ii) reducing noise.

In order to search for a stored vector in the associative memory, the memory is addressed by an input vector, which could be either a complete or a partial representation of one of the stored vectors. The input vector, V^{in}, forms M inner products (IP) with the M stored complex vectors, $V^{st,k}$. These inner products are thresholded by an adaptive threshold function, T_τ and are given by W^k

$$W^k = T_\tau \left[\sum_{j=1}^{L} V_j^{st,k} V_j^{in} \right] \tag{3}$$

Zero thresholding (τ = threshold = 0) [6], average thresholding (τ =average IP) scheme was used to [8, 9] to overcome negative and incorrect convergences of the Hopfield memory [13]. A more intelligent adaptive thresholding scheme formulated recently, enabled us to retrieve more than 95% information with more than 60% missing information [10]. In the present work, the effect of threshold is to accept or reject a hypothesis. The threshold operation, T_τ selects the match or unmatch which is given by:

$$\begin{aligned} T_\tau[w] &= 1(\text{match}) \quad \text{if} \quad w - \tau \geq a\tau, \\ &= 0(\text{no match}) \qquad \text{otherwise} \end{aligned} \tag{4}$$

here a is a constant. Note that the above threshold does not simply select the maximum, but selects only if it exceeds the noise margin by certain amount which is again proportional to the noise level. This threshold formulation is similar to the threshold formula adopted for efficient character recognition problem [10]. Note here that the inner product is usually normalized using the 2-norm of the product vector $V_j^{st,k} V_j^{in}$.

The thresholded output then works as an input to a hetero-associative memory, which contains the information about the orientation, type of the stored object. For a actual implementation, the location of the thresholded output will be a pointer to the information table which is stored in the memory.

The pattern matching in CIPAM mainly involves multiplication of the input complex vector by the stored matrix containing the library of vectors. For optical implementation of the inner product processor, electronically addressable spatial light modulators are suitable. If an electronic processing is desired array processors could used to perform 1-D FFT (10 ms for a 1024 point complex FFT) and signal processing chips such as TMS320 family could be used for matrix vector multiplication.

10 Computer Simulation

The edge images for the stored objects are given in Fig. 1. These images were captured by a video camera, and five different view of the same plane were used as the template. The input object was multiplied point by point with the stored memory complex vectors and then an inverse FFT was taken. The zero order of this sequence is the inner product. The following results show that even when the starting point is shifted, the inner product is still high detect the match. Scaling, translation and rotation produced exactly same result as those of the unshifted inputs (inner product = 0.999).

Input pl23 (unshifted)

	Normalized inner product				
2-norm	1024	3178	2636	1661	4721
Template	pl23	pl25	pl26	pl27	pl28
order					
0	0.999508	0.162913	0.212990	0.310848	0.105691
1	0.001052	0.026222	0.020027	0.009519	0.029689
2	0.000862	0.012553	0.015579	0.016386	0.028344
3	0.001010	0.012257	0.008560	0.014683	0.026138
4	0.001065	0.057627	0.005979	0.005916	0.015674
5	0.000872	0.013869	0.050668	0.004551	0.031553

Decision form CIPAM: The match is with pl 23

The following tables shows the effect when the input is not aligned exactly. When the starting point is shifted by 10 pixels:

Input pl 23 (shifted)

The stored	Normalized Inner product
pl23	0.483682
pl25	0.182784
pl26	0.208999
pl27	0.253237
pl28	0.137422

Decision form CIPAM: The match is with pl23

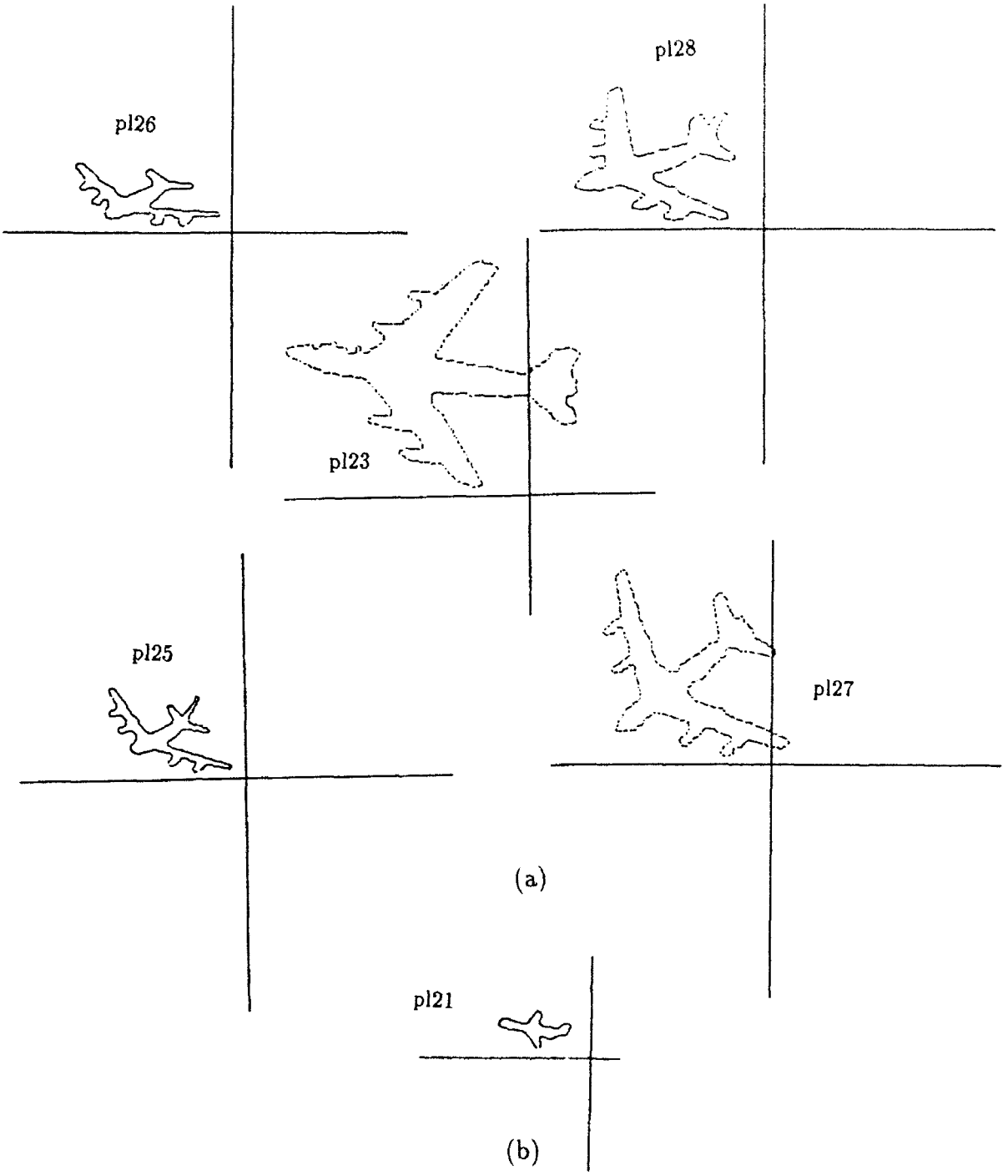


Fig. 1. (a) The templates stored in the CIPAM (b) the unknown object.

When the starting point is shifted by 60 pixels

Input pl 23(shifted)

The stored	Normalized Inner product
pl23	0.437496
pl25	0.121118
pl26	0.152302
pl27	0.277716
pl28	0.135051

Decision form CIPAM: The match is with no. pl23

Experiment with input pl 25

The stored	Normalized Inner product		
	No shift	100 shift	200 shift
pl23	0.36489	0.294045	0.304005
pl25	0.99950	0.456012	0.445188
pl26	0.29118	0.240720	0.227515
pl27	0.42509	0.383282	0.385069
pl28	0.24473	0.226236	0.232428

Decision from CIPAM: Match with pl25.

When the input is pl21 (not one of the templates)

Input pl 21

The stored	Normalized Inner product
pl23	0.664064
pl25	0.627993
pl26	0.631322
pl27	0.678711
pl28	0.519280

Decision from CIPAM: No match found, unknown object.

11 Tracking

The tracking could be done using motion in several ways. Simple centroid based tracking tracks object without identifying them. Here we propose and simulate an optical joint Fourier transform based tracker [4]. In a joint Fourier based processor [14, 19], real time matching may be done quite simply. To perform a matching, the template and the to-be-tracked object are placed side by side on an input plane say a liquid crystal TV. Then the plane is illuminated through a coherent light (say a diode laser). A Fourier transform of the joint image is performed via a lens and the joint intensity transform is recorded via a TV camera. This pattern is then displayed on a second or the same TV and another Fourier transform is taken by a second lens. A cross correlation of the input with the template is performed in real time and detected in the output plane.

Joint Fourier based processors are hybrid compromise between the optical and electronic technology. It avoids the difficulty of optically constructing a holographic matched filter and aligning it accurately. In a dynamic situation, the object may change shape as it moves through the background. In a joint Fourier based processor this difficulty is alleviated as by displaying the object at Δt time earlier. Thus as the shape of the object changes, so does the template. A sample simulation was run using Khoros in the model based vision lab. Figure 2 illustrates a sample result. Note that as the object is moving away from the template. the tracking point is also moving at the same pace.

12 Building a library of object images in real time

As it is important to get a better image for the input object so is it important to build a proper library of templates. For example, building a library with equal increment of aspect angle (or any geometrically equal intervals) may not be the most optimum for better object recognition. Depending on the matching co-efficient (inner product), unequal interval may be chosen which produces a fixed degradation of matching co-efficient between two neighboring templates. Thus depending on the recognition process, the templates are decided. This will ensure predictable results when matching with the actual unknown object. Since we have control over the way the libraries are made, we should utilize as much available information as possible to optimize the performance of moving object tracker/recognizer.

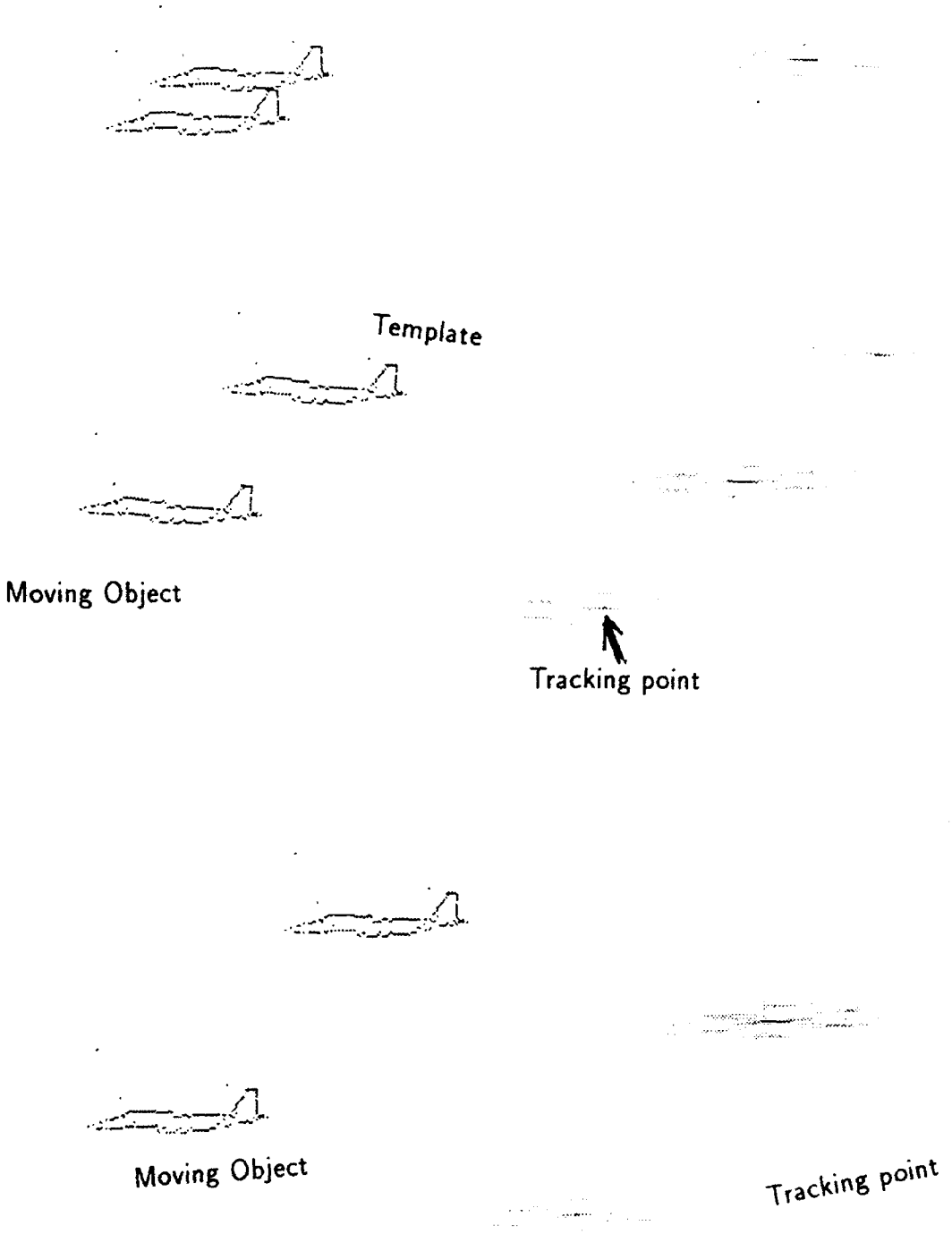


Fig. 2. Moving image tracking using Joint Fourier Transform set up.

13 Reduction of search space and increase the accuracy of detection

In any decision making system, the final decision is seldom made with only one input. In critical cases this could spell disaster when some unforseen situation arises. Thus intelligence gathered through several sources may be more reliable. Topological features could be coupled with our CIPAM model to increase the degree of certainity. For handwritten character recognition such features are used only in the final classification stage. For example, to differentiate between a 6 and 9, the topological features may be used to differentiate between them only after either a 6 or 9 is detected [15]. Here we propose to use such information before and after the final classification and also as a separate decision module. We believe such transfer of information between different levels of decision making could not only confirm a certain decision but also reduce the search space consequently reducing the overall processing time.

14 Conclusion

An approach toward a simplified moving object recognition problem is attacked. A novel complex inner product associative memory (CIPAM) neural network scheme with an adaptive thresholding measure is proposed and its application for the real-time moving object recognition is presented. The proposed memory model results in a highly parallelizable neural net architecture. Consequently, a possible parallel electronic implementation may be identified which will result in a high speed automatic object recognition. In the proposed scheme there is no need for training, since learning is achieved simply when the new complex vector is added to the associative memory.

Motion analysis tools are used as an integral part of the whole system. Motion segmentation is actively used to simplify the matching process. Note that without motion it is not possible to implement the scheme since the object cannot be extracted without it.

The third important aspect is the proposition that multiple source of information may be used to enhance and accelerate the automatic object recognition process. Also priori information may be used as much as possible to maximize the efficiency of the library of template.

15 Future Works

1. Most often a research is usually applied to a limited set of data set, an algorithm which works in an ideal situation may not work for unlimited data set or a non ideal data set. Thus the attaining robustness of the algorithm is a very important extension of any research work.

2. One important error that may come into the edge data is the degradation due to motion and the finite integration time (of CCD detectors) or phosphor persistence. This problem may baffle even the most sophisticated object detection algorithm. Our goal is to model the image acquisition system and undo the effect of motion degradation and acquire a clean data which does not require computationally expensive and time consuming detection algorithm. Using a rough estimate of the velocity we may decide whether to bypass the degradation model or not bypass it [16, 17, 18].

3. To extend our scheme to infrared imageries, specially for ground objects acquired from a moving airplane. For IR images, the image degradation in a dynamic situation is much more severe than those in a static situation. Our model had earlier characterized such effects for night vision equipments and proposed models for restoration of images based on such models [16]. These models could be extended to FLIR (forward looking infrared radar) images. Then the same recognition algorithm could be used with such images.

4. More work needs to be done on the process of making the library. The recognition is dependent on what we store and how we store them. Also the work may be extended to perform real time template library construction.

5. Theoretical work needs to be done on the CIPAM. To find its convergence and matching processes. The formulation of the CIPAM may be altered or modified to incorporate quantization or random noise reduction capability. This can be done since the original model on which it is based on has such capabilities. Newer electronic neural networks chips may be used to implement the algorithm [20]

6. More work needs to be done to validate our proposition of using other topological information to enhance, and accelerate the recognition process.

7. To extend the CIPAM to interpolate between several stored vectors to recreate a better image of the object input.

8. Explore the alternate sampling strategy, to overcome the problem of getting always fixed number of edge data.

9. The idea presented regarding the adaptive threshold may have far-reaching implications, since it may be applied to other neural network models and extended to other applications of the CIPAM model. It may be noted that for any neural network the parallel connection provides speed enhancement, whereas the non-linear operations make it an intelligent decision making system. Thus intelligent modification of the nonlinear behavior of the neurons will have more impact in altering or tailoring the neural network to a particular application or enhancing its capabilities.

Finally, it may be emphasized that the area of pattern recognition is highly interdisciplinary in nature. References [21, 22] points out that pattern recognition community may get positive enhancements form Image Understanding (IU) community, so is also true for the optical pattern recognition community. Optics has many things to offer in terms of its parallel processing capability also in terms of its noise and background reduction capability. For example, many digital recognition algorithm only works if and only if the complete segmentation of the object is done from the scene. In optical matched filtering this is done automatically. Again electronics is famous for its unsurpassed flexibility. Intermingling of all these technologies may only deliver a robust, optimum system.

Acknowledgement I like to acknowledge active co-operation and participation of Mr. Greg Power of WPAFB in this project. The summer support of AFOSR and WLAA is gratefully acknowledged.

16 References

1. *Advances in Computer Vision and Image Processing*, T. S. Huang Ed., Vol. 3. Time Varying Image Analysis. JAI Press Inc., 1988.

2. *Time-Varying Image Processing and Moving Object Recognition 2* V. Cappellini, Ed., Elsevier, Amsterdam, 1990.

3. *Proc. on Workshop on Motion Representation and Analysis* IEEE Computer Society Press, 1986.

4. M. A. Karim, and A. A. S. Awwal *Optical Computing: An Introduction*. Wiley, New York, 1992.

5. T. P. Wallace and O. R. Mitchell, "Analysis of Three-Dimensional Movement Using Fourier Descriptors" IEEE PAMI, pp. 583-588, 1980.

6. A. A. S. Awwal, M. A. Karim, and H. K. Liu, "Associative Memory : An Optimum Binary Neuron Representations" Optical Pattern Recognition, H. K. Liu Ed., Proc SPIE-1053, pp. 17-29, 1989.

7. H. K. Liu, S. Y. Kung, and J. A. Davis, "Real-Time optical Associative retrieval technique," Opt. Engg. Vol. 25, no. 7, 853-856, 1986.

8. A. A. S. Awwal, M. A. Karim, and H. K. Liu, "Machine parts Recognition Using a Trinary Associative Memory" Opt. Engg., Vol. 28, no. 5, pp. 537-543, 1989.

9. M. A. Karim and A. A. S. Awwal, "Convergence Mechanism of Associative Memory" Proc. NAECON Conf, Dayton, Ohio, pp. 698-704, 1989.

10. F. Ahmed, A. A. S. Awwal, and P. Chen, "Character Recognition Using a Trinary Inner Product Associative Memory," IEEE 1992 NAECON Conference Proc., pp. 905-909, Dayton, Ohio, May 1992.

11. A. A. S. Awwal, M. A. Karim, and S. R. jahan, "Improved correlation discrimination using an amplitude modulated phase only filter," Applied Optics, Vol. 29, pp. 233-236, 1990.

12. L. P. Yaroslavsky, "Is the phase-only filter and its modifications optimal in terms of the discrimination capability in pattern recognition ?" Applied Optics, Vol. 31, no. 11, pp. 1677-1679, 1992.

13. J. J. Hopfield, "Neural networks and physical systems with emergent Collective computational abilities," Proc. Natl. Acad. Sci. USA, Vol. 79, no. 8, 2554-2558, 1982.

14. M. S. Alam, A. A. S. Awwal and M. A. Karim, "Improved Correlation Discrimination Using Joint Fourier Transform Optical Correlator," Microwave and Optical Technology Letters, Vol. 4, No. 3, pp. 103-106, Feb. 1991.

15. M. Shridhar and A. Badreldin, "High Accuracy Character Recognition Algorithm Using Fourier and Topological Descriptors", Pattern Recognition, Vol. 17, no. 5, pp. 515-524, 1984.

16. A. A. S. Awwal, M. A. Karim, A. K. Cherri, and D. L. Moon, "Dynamic Modulation Transfer Function of a Display System," Applied Optics, Vol. 30, No. 2, pp. 201-205, Jan. 1991.

17. A. A. S. Awwal, Standardization of Nondiscrete Displays, In Electro-Optical Displays (M. A. Karim Editor), Chapter 11, pp. 447-474, Marcel Dekker, Inc., New York, NY 1992.

18. A. A. S. Awwal, Restoration of Dynamically Degraded Images in Displays, In Electro-Optical Displays (M. A. Karim Editor), Chapter 12, pp. 475-494, Marcel Dekker, Inc., New York, NY 1992.

19. A. A. S. Awwal, A. D. McAulay, Xin xu, and J. Wang, "Device Specific Modelling of the Nonlinear Joint Transform Correlator Using LCLV" IEEE 1992 NAECON Conference Proc., pp. 1135-1138, Dayton, Ohio, May 1992.

20. 80170NX- Electronically Trainable Analog Neural Network chip, *Technical Literature*, Intel Corporation, June 1991.

21. J. D. Wald, D. B. Krig, and T. DePersia, "ATR: Problems and Possibilities for the IU community," Proc. Image Understanding workshop, pp. 255-264, 1992.

22. B. Bhanu and T. L, Jones, "Image Understanding Research for Automatic Target Recognition," Proc. Image Understanding workshop, pp. 249-254, 1992.

ANALYTICAL GUIDANCE LAWS AND INTEGRATED GUIDANCE/AUTOPILOT FOR HOMING MISSILES

S. N. Balakrishnan
Associate Professor
Department of Mechanical and Aerospace Engineering and Engineering Mechanics
University of Missouri-Rolla
Rolla, MO 65401

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Eglin Air Force Base, FL

September 1992

ANALYTICAL GUIDANCE LAWS AND INTEGRATED GUIDANCE/AUTOPILOT FOR HOMING MISSILES

S. N. Balakrishnan
Department of Mechanical and Aerospace Engineering and Engineering Mechanics
University of Missouri-Rolla
Rolla, MO 65401

ABSTRACT

An approach to integrated guidance/autopilot design is considered in this study. It consists of two parts: 1) recognizing the importance of polar coordinates to describe the end game in terms of problem description and measurement acquisition, the terminal guidance problem is formulated in terms of polar coordinates; 2) through a commonly used approximation on time-to-go and a coordinate transformation, a family of proportional navigation optimal guidance laws is obtained in a closed form. A typical element of such a guidance law is combined with the autopilot dynamics to result in a feedback control law in terms of output variables. Numerical simulations are in progress.

ANALYTICAL GUIDANCE LAWS AND INTEGRATED GUIDANCE/AUTOPILOT FOR HOMING MISSILES

S. N. Balakrishnan

I. INTRODUCTION

The most popular homing missile guidance is based on a control law called proportional navigation [1]. The basic notion is that if the line-of-sight rate is annulled, then (for a non-maneuvering, constant velocity target) the missile is on a collision course. If the target is considered smart or maneuvering, then variations to the proportional navigation have been shown to result in better miss distances. These variations have been given optimal control foundations through linear quadratic Gaussian (LQG) formulations [2-5].

There are, however, a few problems with the use of such guidance laws. First is that the measurements in an end game are nonlinear (bearing angle, range, and range rate) in cartesian coordinates. As a consequence, there is linearization in the filtering update process. The measurements are linear in a polar coordinate based state space. However, the propagation between the measurement updates in this case leads to nonlinear equations. Therefore, the states used in the guidance law are suboptimal. The second problem lies with the guidance law which is formulated assuming separability of the guidance (control) law and the estimators which do not hold. It is usually formulated in cartesian coordinates for linearity [3, 4, 6]. The third problem is that the autopilot is usually designed independent of the estimator and the guidance law. Whether it is designed based on classical control or modern control theory (LQG/LTR, H_∞), linearized dynamics are

assumed at different operating points and the autopilot design really does not take the guidance law into account. As a result, there is considerable scope for research in improving the missile performance in terms of estimator, guidance, and autopilot in an intercept scenario [7].

The research in this study is focused on obtaining improvements with a properly posed controller for guidance and its use in an integrated guidance/autopilot design. A few studies have been presented in the area of integrated design of guidance and autopilot [8-10]. The difference here is that we approach the problem from proper formulation of the intercept kinematics. Such a view will enable us to integrate the estimator in the loop in an optimal way and help us address the three problems mentioned earlier in an integrated manner. The central idea here is that the polar coordinates present natural coordinate system for a missile engagement. In order to obtain a closed form solution for the commanded accelerations, the radial and transverse coordinates are decoupled. The decoupling of the coordinates leads to a two point boundary value problem with linear time-varying coefficients. However, with a nonlinear transformation, a class of closed form solutions are obtained which yield several proportional guidance laws.

The rest of the paper is organized as follows: the optimal guidance problem is developed in polar coordinates in Section II. It is further shown to decompose into two decoupled optimal control problems where a closed form control solution is available in the radial direction and a time-varying linear dynamic system has to be solved for control in the transverse direction. A commonly used approximation for time-to-go and a transformation are shown to lead to a class of proportional navigation-type feedback guidance laws in

Section III. In section IV, a feedback control law is shown to result by combining the kinematics of the optimal guidance law with the dynamics of the autopilot. The conclusions are summarized in Section V.

II. OPTIMAL GUIDANCE IN DECOUPLED POLAR COORDINATES

The dynamics of a target-intercept geometry are expressed by a set of coupled nonlinear differential equations in an inertial polar coordinate system as (Figure 1)

$$\ddot{r} - r\dot{\theta}^2 = a_{T_r} - a_{M_r} \tag{1}$$

$$r\ddot{\theta} + 2\dot{r}\dot{\theta} = a_{T_\theta} - a_{M_\theta} . \tag{2}$$

In Eqs. (1) and (2) r is the relative range between the target and the missile, θ is the bearing angle and a_{T_r} and a_{M_r} are respectively the target and missile accelerations in the line-of-sight (radial) direction. Similarly, a_{T_θ} and a_{M_θ} represent the target and missile accelerations respectively in the transverse directions. Dots denote differentiations with respect to time. Note that if the analysis is carried out in three dimensions, there will be another equation involving elevation angle.

Line-of-Sight (Radial) Commanded Acceleration

It can be easily observed that Eq. (1) and (2) are coupled. In order to decouple the dynamics, a pseudo-control in the radial direction, $a_{M_{r1}}$ is defined as

$$a_{M_{r1}} = a_{M_r} - r\dot{\theta}^2 . \tag{3}$$

This definition decouples the radial coordinate from the transverse coordinate. It facilitates a state space, y, in the line-of-sight direction as $y = [r, \dot{r}, a_{T_r}]^T$ and describe their dynamics as

$$\dot{y}_1 = y_2 \tag{4}$$

$$\dot{y}_2 = y_3 - a_{M_{rl}} \tag{5}$$

$$\dot{y}_3 = -\lambda_r y_3 \, . \tag{6}$$

where λ_r is the time constant associated with target acceleration.

The optimal guidance law in the radial direction is obtained as a solution to minimizing the performance index, J_r, where

$$J_r = \frac{1}{2} S_{f_r} y_{1_f}^2 + \frac{1}{2} \int_0^{t_f} \gamma \, a_{M_{rl}}^2 \, dt \, . \tag{7}$$

In Eq. (7), y_{1_f} the value of the relative range (miss-distance) at the final time, t_f, S_{R_f} is the weight on the miss distance, and γ is the weight on the pseudo-control effort. The final time, t_f, which is the time-to-go is approximated as $r/|\dot{r}|$. The minimizing control is

$$a_{M_{rl}}(t) = \frac{t_f}{\gamma} \lambda_1(t) \tag{8}$$

and

$$\lambda_1(t) = S_{f_r}(y_1(t) + t_f \, y_2(t) + (1/\lambda_r^2) a_{T_r}(\exp(-\lambda_r t_f) + \lambda_r t_f - 1))/(1 + t_f^3 S_f/3\gamma). \tag{9}$$

In Eqs. (8) and (9), λ_1 is a Lagrangian multiplier which adjoins the state in Eq. (5) to the performance index in Eq. (7). The actual missile acceleration can be obtained from Eq. (3) as

$$a_{M_r}(t) = a_{M_{r1}}(t) + r(t)\,\dot{\theta}^2(t) .$$

The instantaneous values of the relative range, r(t), and relative range rate, $\dot{r}(t)$, can be solved for by integrating Eq. (4)-(6).

Transverse Acceleration

The equation of motion in the transverse direction in Eq. (2) can be rewritten as

$$\ddot{\theta} = -\frac{2\dot{r}\dot{\theta}}{r} + \frac{1}{r}\,a_{T_\theta} - \frac{1}{r}\,a_{M_\theta} . \qquad (10)$$

Note that since r and $\dot{r}$ are known from Eqs. (4)-(6) they can be treated as functions of time. Consequently, Eq. (10) can be expressed as a time-varying linear differential equation as

$$\ddot{\theta} = f(t)\,\dot{\theta} + g(t)\,a_{T_\theta} - g(t)\,a_{M_\theta} \qquad (11)$$

where $f(t) = -\frac{2\dot{r}}{r}$ and $g(t) = -\frac{1}{r}$.

With a first-order dynamics for target acceleration, Eq. (11) can be expressed in a state space $z = [\theta,\ \dot{\theta},\ a_{T_\theta}]^T$ as

$$\dot{z}_1 = z_2 \tag{12}$$

$$\dot{z}_2 = f(t)\, z_2 + g(t)\, z_3 - g(t)\, a_{M_\theta} \tag{13}$$

and

$$\dot{z}_3 = -\lambda_\theta\, z_3 \,. \tag{14}$$

where λ_θ is the time-constant associated with the transverse target acceleration.

A performance index, J_θ, similar to Eq. (7) for the transverse direction is

$$J_\theta = \frac{1}{2} S_{f_\theta} z_{2_f}^{\,2} + \frac{1}{2} \int_0^{t_f} (\gamma_1 z_2^2 + \gamma_2 a_{M_\theta}^{\,2})\, dt \tag{15}$$

where S_{f_θ}, γ_1, and γ_2 are the weights.

The optimization process to yield the controller minimizing Eq. (15) leads to a two-point boundary value problem:

$$\begin{bmatrix} \dot{z}_2 \\ \dot{\lambda}_2 \end{bmatrix} = \begin{bmatrix} f(t) & -g^2(t) \\ -\gamma_1 & -f(t) \end{bmatrix} \begin{bmatrix} z_2 \\ \lambda_2 \end{bmatrix} + \begin{bmatrix} g(t) z_{30}\, \exp(-\lambda_\theta t) \\ 0 \end{bmatrix} \tag{16}$$

where z_{30} is assumed known and $\lambda_{2_f} = S_{f_\theta}/z_{2_f}$. In Eq. (16), λ_2 represents the Lagrangian multiplier which adjoins Eq. (13) to the performance index, J_θ. This system can be solved either numerically by techniques such as the shooting method or analytically if functional forms of f(t) and g(t) are known. The minimizing control in the transverse direction is given by

$$a_{M_\theta}(t) = \lambda_2(t)\, g(t)/\gamma_2 \,. \tag{17}$$

III. A CLASS OF PROPORTIONAL NAVIGATION GUIDANCE LAWS THROUGH TRANSFORMATIONS

Although Eqs. (16) and (7) represent a general decoupled solution, interesting analytical solutions for the terminal guidance problem and a feedback guidance law can be obtained through a transformation of coordinates. For comparison with existing results, the target acceleration is assumed zero.

Note that the final time (time-to-go) calculation involves an assumption that the closing velocity (relative range rate) is constant. This assumption can be translated to

$$r(t) = \dot{r}\,(t_f - t) \tag{18}$$

where t is the current time. In a feedback rule, this assumption is not very restrictive since $\dot{r}$ is updated at each instant. By using Eq. (18) in Eq. (10), we get

$$\frac{d}{dt}\,\dot{\theta} = \frac{2\dot{\theta}}{t_f - t} - \frac{a_{M_o}\,t_f}{r_o\,(t_f - t)} \,. \tag{19}$$

This equation is difficult to integrate numerically since $(t_f - t)$ appears in the denominator. Hence, define a variable u as

$$u = (t_f - t)^2\,\dot{\theta} \,. \tag{20}$$

The differential equation for u is (after some algebra)

$$\dot{u} = - a_{M_o}\,\frac{\dot{r}_f}{r_o}\,(t_f - t) \,. \tag{21}$$

Note that Eq. (21) is devoid of any expression in u on the right hand side and $(t_f - t)$ in the denominator.

The optimal control problem is now solved through the use of the new variable u. Consider a performance index, $J_{\theta 1}$, given by

$$J_{\theta 1} = \int_0^{t_f} \frac{1}{2} \gamma_2(t) \, a_{M_\theta}{}^2 \, dt \, . \tag{22}$$

Although this performance index seems simpler than J_θ in Eq. (15), it is shown later that Eq.(22) can accommodate a variety of designs by assuming different functional representations for the weight $\gamma_2(t)$.

The Hamiltonian, H, of this system is given by

$$H = \frac{1}{2} \gamma_2(t) \, a_{M_\theta}{}^2 - \lambda \, a_{M_\theta} \frac{t_f}{r_o} (t_f - t) \, . \tag{23}$$

The propagation of the Lagrangian multiplier, λ, is governed by

$$\dot{\lambda} = 0 \, . \tag{24}$$

Hence, λ is a constant. The optimality condition leads to

$$a_{M_\theta} = \frac{t_f(t_f - t)}{r_o \, \gamma_2(t)} \lambda \, . \tag{25}$$

By using Eq. (25) in the propagation equation for u in Eq. (21), we get

$$\dot{u} = \frac{- t_f^2 \, (t_f - t)^2}{r_o^2} \frac{\lambda}{\gamma_2(t)} \, . \tag{26}$$

We will now derive a family of proportional-navigation laws. Let

$$\gamma_2(t) = (t_f - t)^{-k} \tag{27}$$

where k is a positive integer. The implication of this time-varying weight is that the control effort should achieve most of the trajectory shaping before the time-to-go reaches the last second.

With this expression for $\gamma_2(t)$, Eq. (26) can be integrated. The Lagrangian multiplier consequently can be solved for as

$$\lambda = \frac{r_o^2 \dot{\theta}_o (k + 3)}{t_f^{k+3}} . \tag{28}$$

With Eq. (28) we can solve for u(t) from Eq. (26) as

$$u(t) = u(o) + \frac{\dot{\theta}_o}{t_f^{k+1}} \left[(t_f - t)^{k+3} - t_f^{k+3}\right] . \tag{29}$$

The control acceleration, $a_{M_e}(t)$ and the line-of-sight rate, $\dot{\theta}(t)$ can be obtained as explicit functions of time as

$$a_{M_e}(t) = (k+3)\, \dot{r}_o \dot{\theta}_o \,(1 - (t/t_f))^{k+1},\ k \neq -3 \tag{30}$$

and

$$\dot{\theta}(t) = \dot{\theta}_o \,(1 - t/t_f)^{k+1} . \tag{31}$$

By varying k ($\neq$ -3), we can obtain a family of proportional navigation guidance laws. In particular, let k = 0, in Eq. (30) and (31). We get

$$a_{M_e}(t) = 3\dot{r}_o \dot{\theta}_o \,(1 - t/t_f) \tag{32}$$

and

$$\dot{\theta}(t) = \dot{\theta}_o \,(1 - t/t_f) . \tag{33}$$

If we assume t = 0 as the current time, we get

$$a_{M_e} = 3\, \dot{r}_o \dot{\theta}_o . \tag{34}$$

This is the standard proportional navigation guidance law. By assuming non-zero values for k which are greater than zero, we can get the line-of-sight rate decreased sooner as desired.

It can be argued that since the performance index, $J_{\theta 1}$, does not contain any state dependent ($\dot{\theta}$-dependent) term, we do not have any control over the line-of-sight rate, $\dot{\theta}$. The line-of-sight rate is of dual importance. The aim of the guidance law is to nullify the line-of-sight rate so that missiles can ride the line-of-sight to the target. However, from observability aspects involved in the design of the estimator, it is a desirable feature to have a non-zero line-of-sight rate (in a passive case). Hence, it becomes imperative to weight the line-of-sight rate to achieve both objectives. A good strategy is to have acceptable line-of-sight rates at the initial period of the end game to help estimate the target positions accurately and drive the line-of-sight to zero at the latter part of the engagement. The presence of state related terms in the performance index, however, leads to more involved optimization problems. A way to obviate such problems is to design a time-varying weight on the control so as to incorporate the constraints on the line-of-sight rate. For example, consider a weight, $\gamma_2(t)$, given by

$$\gamma_2(t) = \frac{a_o\, t_r}{\left[a_1 + a_2 \left(\frac{1}{2} - t_r\right)^3 + a_3\, t_r \left(\frac{1}{2} - t_r\right)^2\right]} \tag{35}$$

where a_i, $i = 0, 1, 2, 3$ are constants and $t_r = t/t_f$. This weight will result in a line-of-history rate, $\dot{\theta}(t)$ of a form given by

$$\dot{\theta}(t) = a_1 + a_2 \left[\frac{1}{2} - t_r\right]^3 \tag{36}$$

and a control acceleration history, $a_{M_\theta}(t)$ in a term given by

$$a_{M_\theta} = b_1\, \dot{\theta}(t) + b_2\, t_r \left(\frac{1}{2} - t_r\right)^2 . \tag{37}$$

Note that Eq. (36) implies that line-of-sight rate can be kept high initially in order to help the estimator and as t_r reaches unity, towards the end, line-of-sight rate becomes small and tends to zero resulting in an intercept. These correlations are easy to calculate from Eq.(25) only because the Lagrangian multiplier is a constant in our formulation.

IV. INTEGRATED GUIDANCE LAW/AUTOPILOT DESIGN

As indicated earlier, an avenue to improve the homing missile performance is to consider an integrated design of the guidance law and autopilot instead of separate designs. In this manner, the kinematics of the engagement geometry used in the guidance law development and the dynamics of the airframe as reflected in the linearized equations of motion and used in the autopilot design can be brought together.

Consider the geometry of the engagement in Figure 1 and the relationship between the flight path angle and the pitch angle and angle of attack as shown in Figure 2.

One way to formulate the integrated approach is to relate the guidance law to the dynamics of the airframe directly and solve for the commanded control surface deflection.

For this purpose, let us assume a conventional proportional navigation guidance law given by

$$a_{M_o}(t) = k(1 - t_r) \qquad (38)$$

where $k = 3\dot{\theta}_o \dot{r}_o$

and $t_r = t/t_f$.

This transverse acceleration can be approximated to be perpendicular to the flight path of the missile so that

$$V_m \dot{\gamma} = k(1 - t_t) \tag{39}$$

where V_m = missile velocity

and γ = flight path angle.

The flight path angle is related to the pitch angle and angle of attack as

$$\gamma = \sigma - \alpha \tag{40}$$

or $$\dot{\gamma} = \dot{\sigma} - \dot{\alpha} = q - \dot{\alpha} \tag{41}$$

where q = pitch rate.

The linearized dynamics of the airframe in a short period mode is described by

$$\dot{\alpha} = Z_\alpha \alpha + q + Z_\delta \delta_c \tag{42}$$

$$\dot{q} = M_\alpha \alpha + M_q q + M_\delta \delta_c \tag{43}$$

where Z_α, Z_δ, M_α, M_q, and M_δ are the dynamic stability derivatives of an airframe.

From Eq. (39)-(43), a feedback law for the control surface deflection can be obtained as

$$\delta_c = (C_1 + C_2)q - C_1 \dot{\alpha} + C_3 \dot{q} \tag{44}$$

or

$$\delta_c = C_1 \frac{3\dot{\theta}_o \dot{r}_o}{V_m} + C_2 q + C_3 \dot{q} \tag{45}$$

where $C_3 = \dfrac{Z_\alpha}{Z_\alpha M_\delta - M_\alpha Z_\delta}$ (46)

$$C_2 = -M_q C_3 \tag{47}$$

and $C_1 = \dfrac{M_\alpha}{Z_\alpha M_\delta - M_\alpha Z_\delta}$. (48)

Note that the control here considers control deflection and not its rate, $\dot{\delta}_c$, as in [9]. If needed, it can be obtained either with proper use of Eq. (41) or reformulation of the

acceleration and kinematics terms in Eq. (41), (42), and (43). It should be observed, however, that such rate commands will lead to need for estimates of line-of-sight acceleration [9] and hence, may be less reliable when they are coupled with estimators.

Note that the variables in Eq. (44) and (45) can be picked up through measurements. Numerical experiments of these controls are being conducted.

V. CONCLUSIONS

A class of proportional navigation guidance laws have been derived through an approximation of time-to-go and a transformation of state variables. A closed form feedback autopilot control law has also been derived by the use of the guidance law in the dynamics of the airframe.

ACKNOWLEDGEMENTS

The author wishes to express his gratefulness to Dr. James R. Cloutier and the excellent group in MNAG Division at the Armament Laboratories at the Eglin Air Force Base for the insights they provided, discussions, and help. The author is thankful to the Air Force Office of Scientific Research for providing the Summer Faculty Research Fellowship to work at the Eglin Air Force Base, Florida that made this work possible.

REFERENCES

1. Adler, F. P., "Missile Guidance by Three-Dimensional Proportional Navigation," Journal of Applied Physics, Vol. 27, 1956.

2. Fiske, P. H., "Advanced Digital Guidance and Control Concepts for Air-to-Air Tactical Missiles," U.S. Air Force Armament Lab., Armament Development and Test Center, Eglin AFB, FL, Rept. AFATL-TR-77-130, December 1977.

3. Sammons, J. M., Balakrishnan, S., Speyer, J. L., and Hull, D. G., "Development and Comparison of Optimal Filters," Air Force Armament Lab., U.S. Air Force Systems Command, Eglin AFB, FL, Rept. AFATL-TR-79-87, October 1979.

4. Balakrishnan, S. N. and Speyer, J. L., "A Coordinate Transformation-Based Filter for Improved Target Tracking," Journal of Guidance, Control, and Dynamics, Vol. 9, November-December 1986, pp. 704-709.

5. Bryson, A. E. and Ho, Y. C., Applied Optimal Control, New York: Hemisphere, 1975.

6. Balakrishnan, S. N., "An Extension of Modified Polar Coordinates and Application with Passive Measurements," AIAA Journal of Guidance, Control, and Dynamics, November-December 1989, pp. 906-912.

7. Cloutier, J. R., Evers, J. H., and Feeley, J. J., "Assessment of Air-to-Air Missile Guidance and Control Technology," IEEE Control Systems Magazine, 9:6, 1986.

8. Yueh, W. R. and Lin, C. F., "Optimal Controller for Homing Missile," AIAA Journal of Guidance, Control, and Dynamics, 8:3, 1985.

9. Evers, J. H., Cloutier, J. R., Lin, C. F., Yueh, W. R., and Wang, Q., "Application of Integrated Guidance and Control Schemes to a Precision Guided Missile," 1992 American Control Conference, Chicago, Illinois, June 24-26, 1992, pp. 3220-3224.

10. Evers, J. H., Cloutier, J. R., and Lin, C. F., "A LTV Dynamics Model for Missile Guidance and Control in the End Game," 1992 AIAA Guidance, Navigation, and Control Conference, Part 2, Hilton Head, South Carolina, August 10-12, 1992, pp. 1024-1029.

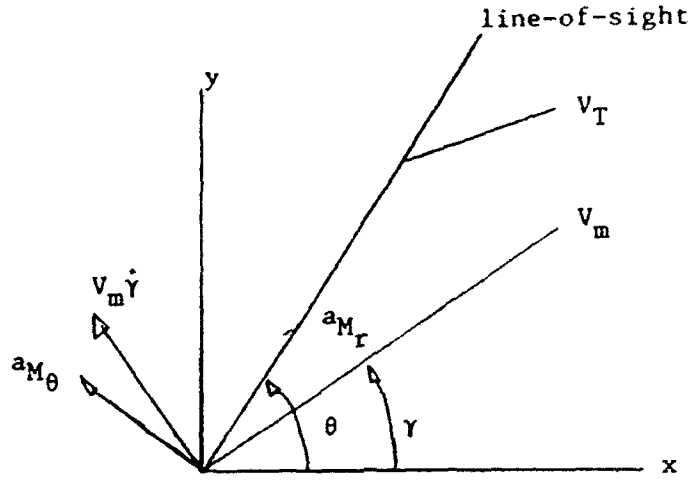


Figure 1. Engagement Geometry.

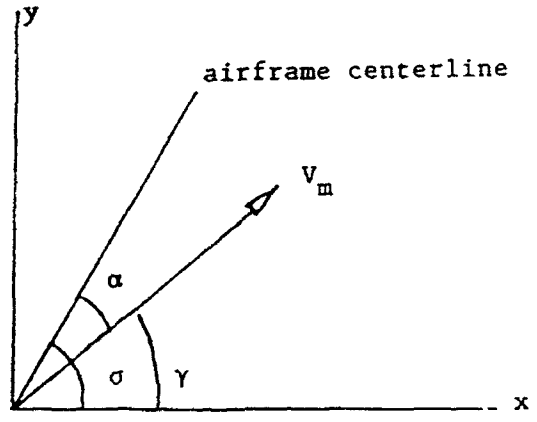


Figure 2. Schematic of Angular Relationship.

ESTIMATION OF ASPECT ANGLES OF TARGETS IN FLIR IMAGES

Prabir Bhattacharya

Associate Professor

Department of Computer Science & Engineering

University of Nebraska-Lincoln

Ferguson Hall, Rm 214A

Lincoln, NE 68588-0115

Final Report for:

Summer Research Program

Wright Laboratory WL/AARA

Sponsored by:

Air Force Office of Scientific Research

Bolling Air Force Base, Washington, D.C.

July, 1992

ESTIMATION OF ASPECT ANGLES OF TARGETS IN FLIR IMAGES

Prabir Bhattacharya

Associate Professor

Department of Computer Science & Engineering

University of Nebraska-Lincoln

Abstract

We develop a model-based approach to estimate the aspect angle of a target in a FLIR image. From a set of given 3-D models represented in the voxel form, a database of binary images matching algorithm is then used to match a silhouetted version of the given target with the images in the database. The angle of rotation corresponding to the matched image in the database from the database is equal to the required aspect angle of the target.

1 INTRODUCTION

Automatic Target Recognition (ATR) is a practical problem of considerable interest to the military. One direct application of the ATR is to guide pilots of high performance aircraft flying at low altitudes in night or in bad weather and to recognize targets. The ATR is also of serious interest to other branches of the military, such as the navy. Extensive work has been done on the recognition of ships using radar imagery (e.g., [3]). However, detecting a target like a tank from an aircraft is much more complex than detecting a (large) ship in the open ocean. For, the ship would appear to be a long thin strip and the segmentation of the image is easy because the image can be modelled as being *two-class* – the target being either brighter or darker than the scene [3].

There are three standard types of sensors which are available: the forward-looking infrared (FLIR), the synthetic aperture radar (SAR) with three primary imaging modes (range, intensity and Doppler) and the multiple mode laser radar. ATR using the FLIR imagery has been studied extensively. Although the interpretation of FLIR images is complex, the geometry of a FLIR image could be used successfully for the purpose of recognition of the target. One information about the target of significant interest is the determination of the **aspect angle** (see e.g., [3],) of the target which is defined to be the orientation of the target about its vertical axis. Let us introduce some more relevant ATR terminologies: the *depression angle* of the target is the angle below the horizontal plane of the line from the sensor. The *pose* of the target is the combination of the information consisting of the depression angle and the aspect angle. For low altitude views the depression angle is relatively insignificant. Therefore, it is then important problem to estimate the aspect angle of the target.

The **Model-Based Vision** (MBV) approach is an important methodology for 3-D object recognition and is widely used in computer vision research. In the MBV approach, the unknown image is matched with some known models and in this manner it can be *recognized*. The MBV

approach has been investigated extensively at the Wright Lab, WL/AARA, for Automatic Target Recognition, see e.g., [11], [12], [13]. In this project we have used the MBV approach to estimate the aspect angle of a target.

2 PRELIMINARIES

There are essentially three general approaches to represent rigid solids: the surface or boundary representation, the generalized cylinders representation and the volumetric (*voxel*) representation. Here, we shall use the voxel representation of solids. There are several methods to store and process the voxel representation of solids, we shall consider here an algebraic approach using polynomials [7]. First, given a binary image say, of size $r \times s$ we associate with it a polynomial in two variables : $\sum_{i,j} a_{ij} X^i Y^j$ where the summation is over all the pixels of the image and a_{ij} is 1 or 0 according as (i, j) corresponds to a black or white pixel respectively. Next, a 3-D object given by the voxel representation is encoded by a polynomial in three variables:

$$\sum_{i,j,k} p_{ijk} X^i Y^j Z^k$$

where $p_{ijk} = 1$ if the voxel (i, j, k) of the array lies completely inside A, otherwise $p_{ijk} = 0$.

The *rotation procedure* used in our proposed algorithm, is an adaptation of the general method of *geometric transformation* given in [10]. Consider a 3D object S with the polynomial representation $\sum_{ijk} p_{ijk} X^i Y^j Z^k$ and let S' be the solid obtained by rotating S through an angle θ about the OZ axis. When θ is a multiple of 90° the polynomial representation of S' is easy to obtain from that of S. For example, if $\theta = 180°$, then we have that

$$S' \equiv \sum_{ijk} p_{ijk} X^{-1} Y^{-1} Z^{-1}$$

When θ is not necessarily a multiple of 90° S' is obtained as an approximation in the following manner. Suppose that the polynomial representation of S' is $\sum_{rst} q_{rst} X^r Y^s Z^t$. Compute each

(r, s, t) from (i, j, k) using the following transformation:

$$\begin{bmatrix} i \\ j \\ k \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r \\ s \\ t \end{bmatrix} .$$

If $\cos\theta$ or $\sin\theta$ are not integers, we round the solution to integral values. For any r, s, t in the set of all real numbers we get (i, j, k) coresponding to (r, s, t). If $p_{ijk} = 1$ then we assign $q_{rst} = 1$ where (r, s, t) is the solution corresponding to (i, j, k). We refer the reader to Rosenfeld and Kak (1982, 9.3) for details regarding why this indirect method is better than the direct method of computation by the *inverse* of the transformation given above. In a similar manner, we may obtain rotations about other cordinate axes.

Let $P = \sum_{i,j,k} p_{ijk} X^i Y^j Z^k$ be the polynomial represention of a solid A. We shall now determine the polynomial representations of the orthographic projections of A onto the coordinate planes. First, consider the orthographic projection onto the XY plane. Let

$$Q = \sum_{i,j} (\bigvee_{k} p_{ijk}) X^i Y^j$$

where $\vee$ denotes the usual Boolean operation OR. We claim that Q is the polynomial representation of the orthographic projection of A onto the XY plane. For, if two or more terms in the polynomial P have the same factor $X^i Y^j$ then the voxels lying in A corresponding to these terms give the same pixel while projecting orthographically onto the XY plane. So, putting $Z = 1$ for these terms in P and then OR-ing their coefficients gives the projection algebraically. If there is only one term in P with the factor $X^i Y^j$ then clearly the projection of the corresponding voxel onto the XY plane is obtained by putting $Z = 1$ in the term. The orthographic projections onto the other coordinate planes are obtained in a similar manner.

3 PROPOSED ALGORITHM

3.1 Main Ideas

We propose a model-based approach to estimate the aspect angle which is based on our recent work [7]. First, we sketch the main ideas of our algorithm. We assume that the target is an object which is identical (at least approximately) to one of a set of given 3-D *models*.

Step 1 (Pre-processing stage): The given 3-D models are rotated through a series of pre-determined angles about the Z axis, the vertical direction. Then, the projections of these various rotated views onto the XY plane are obtained. Thus, we have now a database of 2-D images obtained from the given models: each such image corresponds to an angle, θ, by which a given model has been rotated, and then projected, to get this image. We can make this database as extensive as possible by choosing more and more angles in the range [0,360], measured in degrees.

Step 2: The (FLIR) image of the target is now segmented (using some standard segmentation procedure) to obtain a silhouetted image. (In practice, the scene would consist of several objects and we could zoom into some specific object, the target, whose aspect angle we are interested to determine.)

Step 3: We now use a matching algorithm to match a silhouetted image of the target with the images in the database obtained in Step 1. When we are able to match the given image with one of the images in the database (upto a certain percentage of accuracy, depending on the number of images in the database), the angle θ corresponding to the image in the database is the aspect angle which we are interested to determine.

The above method is a workable method to compute the aspect angle of a target.

Salient features:

- To implement the various steps, we shall use an algebraic approach to process images which

we have studied extensively earlier, e.g., [1], [2], [5], [6], [7], [8], etc. Our previous work on this algebraic approach was partially supported by the AFOSR Grant AFOSR 90-0046.

- *Speedup*: In Step 3 of the method, if we use an algebraic method for matching developed by us [7], then a tremendous speedup of the run-time is possible if we use the well known *Fast Fourier Transform* (FFT) technique to compute the convolutions which occur in our matching algorithm (the estimates of the speed-up has been given explicitly in [7]).

3.2 Description

For simplification, the models are rotated through steps of 22.5° each: it is straightforward to extend this to more (less) general situations. Our objective, here, is to match an unknown object with one of these models by using a silhoutted image of the target: this would give us the aspect angle, as explained earlier. It is assumed that the target is approximately similar to one of the models.

Suppose we have n models denoted by $A_i,\ 1 \leq i \leq n$ whose polynomial representations are $P_i,\ 1 \leq i \leq n$ respectively.

Algorithm: Aspect Angle Estimation

Step 1. Rotate the given models through 360° by steps of 22.5° about each of the coordinate axes. Thus we obtain a set of polynomials. For $1 \leq i \leq n$, let $P^i_{\lambda(\theta)}$ denote the polynomial representation of the solid obtained from P_i by rotating about the axis $O\lambda$ through an angle θ where $\lambda \in \{X, Y, Z\}$ and θ (measured in degrees) in the given range. (For example, in this notation $P^i_{X(45)}$ corresponds to the solid obtained from P_i by rotating about the X axis by 45°.)

Step 2. Project orthographically all the solids obtained in Step 1 onto the XY plane. Let $Q^i_{\lambda(\theta)}$ denote the polynomial representation of the projection of the solid corresponding to $P^i_{\lambda(\theta)}$. Now,

let $R^i_{\lambda(\theta)}$ be the polynomial obtained in Step 2 of the procedure 3.4 when $A := Q^i_{\lambda(\theta)}$.

Step 3. A single silhouetted image of the target is obtained by using any standard thresholding algorithm (the choice of this is left to the operator). Let $Q \equiv Q(X,Y)$ be the polynomial representation of the resulting binary image. Let

$$S^i_{\lambda(\theta)} := Q * R^i_{\lambda(\theta)}, \quad 1 \leq i \leq n, \quad \lambda \in \{X, Y, Z\}$$

where * denotes the usual multiplication of polynomials.

Step 4. Let $S^i_{\lambda(\theta)} \equiv \sum_{lm} s_{lm} X^l Y^m$. Let us refer to the largest value of s_{lm} as the *weight* of the polynomial. Now, determine the weights of the polynomials obtained in Step 3 and find out the polynomial in each class corresponding to $i = 1 \ldots n$ which has the highest weight (if there is more than one such polynomial then we choose the first one that we come across). Then, compare the weights of the n highest-weight polynomials from each class and select the one which has the largest weight.

Theorem [9]. *If the highest weight polynomial comes from the class i, $(1 \leq i \leq n)$, then the closest match of the unknown object would be the object A_i.*

3.3 Further Refinements and Complexity

If the target is only approximately similar to one of the models, we can compute the percentage of matching of the target with each of the models and then choose the one with the highest percentage of match as the recognition of the target. The angle of rotation of the view with whom the match is being made corresponds to the aspect angle of the target. Also, we can vary the level of thresholding of the FLIR image corresponding to the target and carry out the steps of the algorithm repeatedly until we are able to recognize the test object in a convincingly way (that is, with a satisfactorily high level of percentage of matching), initially we may start with a high thresholding level and then gradually lower the level and run the algorithm with each level until we are able to make a clear

decision.

Complexity of the Algorithm. Whether we use the polynomial representation or the traditional array representation, Step 1 (rotation) has the complexity $O(m^2)$, Step 2 (projection) has the complexity $O(m^3)$, Step 3 (multiplication) has the complexity $O(m^2n^2)$, and Step 4 (comparison) has the complexity $O(n^2)$.

4 EXPERIMENTAL DETAILS

The implementation of the proposed algorithm was done using the standard image processing package *KHOROS* which was run on a *SPRARC station 2* computer.

The model of an M60 tank was created using the *BRL CAD* package. This package was used to obtain a binary silhouette of the tank which was then rotated through steps of 22.5° to get sixteen views which we denote by $t_0, ..., t_{15}$. These were used as the *templates* in our experiments. (We remark that the templates used in our experiments are not all of the same size.)

A series of images were provided for the implementation, each image consisting of a scene with four M60 tanks (signatures) and some background clutter. From each such image, we extracted, one by one one, a signature and matched it, after a suitable thresholding, with the sixteen templates using the "normalized cross correlation" method. If the *best match* occurred with the template t_i, then according to our theory, the aspect angle of the signature was expected to correspond with the angle by which the given binary silhouette (model) had been rotated to obtain t_i.

In general, we found that the results were consistent with the actual values of the aspect angles (which were made available to us in a file called the "Ground Truth Fle") subject to the various limitations of our experiment (see the following section).

In the first image (which was in the directory /usr/data/tv and named cota.0531.arf), the aspect angle of all the signatures is zero. We have included four tables where the values of the matching

of the four signatures are recorded. In each table, u denotes the signature (unknown image) whose dimensions are given on the line beginning with u, then the dimensions of the sixteen templates t_i are given, and subsequently the percentages of the match of u with each t_i are recorded. In each of the four images we find that the best match was with t_0 which is, in fact, the template corresponding to 0° rotation. Another image (cota.0538) is also included for which we found the results (not included here for space limitations) were correct upto an approximation.

5 RECOMMENDATIONS

This report gives a practical method to estimate the aspect angle of a target using the model-based-vision approach. The experimental results presented in this report show good promise of the feasibility of the method in a realistic situation. During the last week of my visit, I gave a presentation at the WL/AARA reporting the results of the project. We intend shortly to submit a Grant Proposal to the AFOSR as a follow-up Proposal to this Summer Visit where we propose to develop our method in much more detail than it was possible in the short duration of the present visit. More, specifically, our future objective would be to develop this method in the following manner:

1. As suggested by E. Zelnio, we would look into the problem of scaling the method so that in the real time situation it is a feasbile method.

2. The segmentation of the given FLIR image of the target would be done in a more efficient manner using more aevanced segmentation algorithms which have been created specifically for FLIR imagery.

3. The database library of the rotated views of the models would be made more extensive.

4. The models would be created in a more realistic manner using some more advanced CAD/CAM generated images. Notice that a lot depends on how well the models are generated.

5. We shall use a FFT package to perform the convolution which is required in our matching subroutine. This would give a tremendous speed-up of the entire method.

6. Extensive further testing of the entire system to estimate the aspect angle would be done using the images to be supplied through the courtesy of the Model-Based Vision Lab, WL/AARA, Wright-Patterson Air Force Base..

ACKNOWLEDGEMENTS

I would like to thank the *Model Based Vision Lab*, WL/AARA of the Wright Laboratory and the AFOSR to give me the opportunity to spend eight weeks working on this project. Special thanks go to Vincent Velten who acted as my Focal Point; the problem investigated in this project was suggested by him and he spent considerable amount of time discussing the project with me at various stages. I would like to thank Ed Zelnio and Mike Bryant sincerely for providing the excellent facilities of the MBV Lab. Thanks are also due to Jim Leonard and Lori Westerkamp of WL/AARA for several useful discussions. The system support staff at the Lab were all very nice and helpful, in particular, Bruce York (of Svendrup) helped me greatly with programming and made some useful observations.

References

[1] P. Bhattacharya, Study of Features of Binary Images using Image Algebra Techniques, AFOSR Grant No. 90-0046, , 1990. Final Report.

[2] P. Bhattacharya and K. Qian, Polynomial Approach for Developing Vision Systems, *Proc. of IEEE Internat. Conf. on Systems, Man & Cybernetics*, 1991, pp. 35-37.

[3] D. N. Kato *et al.*, Ship classification and aimpoint maintenance, *Infrared Systems and Components*, SPIE vol. 890, pp. 174-181, 1988.

[4] K. Qian and P. Bhattacharya, Binary image processing by polynomial approach, *Pattern Recognition Letters*, vol. 11, pp. 395-403, 1990.

[5] K. Qian and P. Bhattacharya, An algebraic method for processing colored images, *Pattern Recognition Letters*, vol. 12, pp. 805-811, 1991.

[6] K. Qian and P. Bhattacharya, A 3D object recognition system using voxel representation, *Pattern Recognition Letters*, to appear in Sept., 1992.

[7] K. Qian and P. Bhattacharya, A Template Polynomial Approach for image processing and visual recognition, *Pattern Recognition*, to appear.

[8] K. Qian and P. Bhattacharya, Determining Holes and Connectivity in Binary Images, *Computers & Graphics*, to appear.

[9] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, Academic Pr., New York, 1982. vol. 1–2.

[10] V. J. Velten, Signature prediction models for flir target recognition, *Proc. of SPIE on Automatic Object Recognition*, Cocoa Beach, FL, SPIE vol. IS 7, pp. 98-107, 1990.

[11] E. Zelnio, The importance of sensor models to model-based vision applications, *Proc. of Conf. on Automatic Object Recognition*, Cocoa Beach, FL, SPIE vol. IS 7, pp. 98-107, 1990.

[12] E. Zelnio, ATR Paradigm comparison with emphasis on model-based vision, Keynote address: *Proc. of Conf. on Model-Based Vision Developments and Tools*, Boston, MA, SPIE vol. 1609, pp. 2-15, 1991.

```
Image 1 of /usr/data/tv/cota0531.arf
**********************************

          DATA
**********************************


u     56 42

16

t_0   70 80
t_1   58 76
t_2   57 100
t_3   61 109
t_4   58 102
t_5   60 109
t_6   57 100
t_7   58 76
t_8   58 80
t_9   58 76
t_10  58 100
t_11  58 109
t_12  59 102
t_13  58 109
t_14  57 100
t_15  58 76


**********************************
          RESULTS
**********************************

view: t_0     matching: 0.424588
view: t_1     matching: 0.378288
view: t_2     matching: 0.352167
view: t_3     matching: 0.354517
view: t_4     matching: 0.353468
view: t_5     matching: 0.357613
view: t_6     matching: 0.347319
view: t_7     matching: 0.382596
view: t_8     matching: 0.324759
view: t_9     matching: 0.375550
view: t_10    matching: 0.358571
view: t_11    matching: 0.279043
view: t_12    matching: 0.348885
view: t_13    matching: 0.350269
view: t_14    matching: 0.352768
view: t_15    matching: 0.373775
```

```
Image 2 of /usr/data/tv/cota0531.arf

**********************************

         DATA

**********************************

u 56 47

16

t_0  70 80
t_1  58 76
t_2  57 100
t_3  61 109
t_4  58 102
t_5  60 109
t_6  57 100
t_7  58 76
t_8  58 80
t_9  58 76
t_10 58 100
t_11 58 109
t_12 59 102
t_13 58 109
t_14 57 100
t_15 58 76

**********************************

         RESULTS

**********************************

view: t_0    matching: 0.631602
view: t_1    matching: 0.593611
view: t_2    matching: 0.553059
view: t_3    matching: 0.571841
view: t_4    matching: 0.559361
view: t_5    matching: 0.559911
view: t_6    matching: 0.526209
view: t_7    matching: 0.549272
view: t_8    matching: 0.423005
view: t_9    matching: 0.549748
view: t_10   matching: 0.537835
view: t_11   matching: 0.370524
view: t_12   matching: 0.559342
view: t_13   matching: 0.577848
view: t_14   matching: 0.555353
view: t_15   matching: 0.593006
```

```
Image 3 of /usr/data/tv/cota531.arf

*********************************************

               DATA

*********************************************


u   57 76

16

t_0   70 80
t_1   58 76
t_2   57 100
t_3   61 109
t_4   58 102
t_5   60 109
t_6   57 100
t_7   58 76
t_8   58 80
t_9   58 76
t_10  58 100
t_11  58 109
t_12  59 102
t_13  58 109
t_14  57 100
t_15  58 76


***********************************

          RESULT

***********************************


view: t_0     matching: 0.410554
view: t_1     matching: 0.269783
view: t_2     matching: 0.286133
view: t_3     matching: 0.367166
view: t_4     matching: 0.333259
view: t_5     matching: 0.329533
view: t_6     matching: 0.265940
view: t_7     matching: 0.222089
view: t_8     matching: 0.263392
view: t_9     matching: 0.231753
view: t_10    matching: 0.287519
view: t_11    matching: 0.258839
view: t_12    matching: 0.332972
view: t_13    matching: 0.362101
view: t_14    matching: 0.298592
view: t_15    matching: 0.279394
```

```
Image 4 of /usr/data/tv/cota0531.arf

Lower threshold: 180, higher threshold: 255

**********************************

          DATA
**********************************

u     56 52

16

t_0   70 80
t_1   58 76
t_2   57 100
t_3   61 109
t_4   58 102
t_5   60 109
t_6   57 100
t_7   58 76
t_8   58 80
t_9   58 76
t_10  58 100
t_11  58 109
t_12  59 102
t_13  58 109
t_14  57 100
t_15  58 76


**********************************

          RESULTS

**********************************

view: t_0     matching: 0.501475
view: t_1     matching: 0.478977
view: t_2     matching: 0.456418
view: t_3     matching: 0.465282
view: t_4     matching: 0.459112
view: t_5     matching: 0.452731
view: t_6     matching: 0.416492
view: t_7     matching: 0.424945
view: t_8     matching: 0.382590
view: t_9     matching: 0.426914
view: t_10    matching: 0.431403
view: t_11    matching: 0.358797
view: t_12    matching: 0.450462
view: t_13    matching: 0.465738
view: t_14    matching: 0.466891
view: t_15    matching: 0.483811
```

CRACK ARREST IN COMPOSITE PLATES REINFORCED WITH TOUGH LAYERS

Victor Birman
Associate Professor
Engineering Education Center in St Louis

University of Missouri-Rolla
8001 Natural Bridge Road
St Louis, MO 63121

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

July 1992

CRACK ARREST IN COMPOSITE PLATES REINFORCED WITH TOUGH LAYERS

Victor Birman
Associate Professor
Engineering Education Center
University of Missouri-Rolla

Abstract

The paper presents theoretical foundations of a crack arrest technique applied to composite plates. The approach is based on reinforcement of the original structure with thin tough layers. These layers reduce the stress field around the crack tip resulting in its effective arrest.

The theory utilizes the existing solution for a crack in an anisotropic sheet. It is shown that the addition of tough layers results in a decrease of displacements and stresses by a single factor. This factor is evaluated using the principle of virtual work.

All crack tip stresses and displacements, as well as the components of the strain energy, are expressed in terms of real functions. To the best knowledge of the author, this is the first such presentation of these parameters for a general case of an arbitrary directed crack in arbitrary laminated layers.

Introduction

The fracture phenomena in composite materials have been intensively studied due to their increasing applications in structural components. The theoretical approach utilizing complex potential methods is based on the monograph of Lekhnitskii [1], although he did not apply his technique to crack problems. These problems were considered by Sih, Paris and Irwin [2] and Sih and Leibowitz [3] for two-dimensional plane stress. Their solutions utilized the complex stress function analysis of Lekhnitskii in conjunction with the appropriate boundary conditions on the crack edges. A modification of the method

using the conservation laws of anisotropic elasticity was developed in [4] for the analysis of mixed-mode cracks. A generalized complex potential method has recently been presented by Becker [5, 6] for the solution of problems where the bending-extension coupling is present.

It must be noted that the crack tip displacements modes in anisotropic materials under the application of a single load mode are mixed [7]. The exception is found in specially orthotropic materials where a linear crack propagates along one of the principal directions. Therefore, a realistic study of fracture in composites should consider mixed-mode cracks, although single-mode analyses can be used to obtain qualitative results.

An excellent outline of research on cracks in composite plates was given in the book of Sih and Chen [8]. Mentioned here is also a more recent paper of Williams [9].

The present paper introduces theoretical foundation of a technique intended to slow or completely arrest cracks in composite plates. According to this technique, the laminate is "reinforced" by tougher layers. A crack propagating in a damaged layer induces displacements and stresses in the tougher (arrester) layers. In turn, these layers interact with the damaged layer reducing the stresses around the crack tip. Obviously perfect bonding and prevention of the interface fracture are the necessary conditions for successful implementation of this technique. Note that this technique is similar to ARALL laminates that have a remarkable fatigue resistance [10].

It is necessary to emphasize that fracture of three-layered laminates with a through crack was considered in [8, Ch 6] where each layer was assumed isotropic and homogenous. The fracture problem for a plate formed of three isotropic layers with a crack in the center layer was studied by Chen and Sih [11]. In the contrary, in the present paper, all three layers are anisotropic, although in practical applications it may be advantageous to use

quasi-isotropic or isotropic arrester layers that would provide equal resistance to the crack propagation in all directions.

Analysis

Consider a three-layered plate with a through-the-thickness straight crack in the central layer (Fig 1). Two outer layers are symmetric about the central layer. The analysis is based on the following assumptions.

- The materials of all layers are elastic;
- Plane stress state exists throughout the laminate;
- Bonding between the layers is perfect and there are no interlaminar cracks;
- Due to a small thickness of the outer layers, in-plane displacements in these layers are independent of the thickness coordinate.

1. Plane Stress in Anisotropic Layer

This problem solved by Lekhnitskii [1] is outlined here for convenience.

The static plane equilibrium equations, in the absence of body forces, are:

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} = 0$$

$$\frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \sigma_y}{\partial y} = 0 \qquad (1)$$

For a generally orthotropic layer, the strain-stress relationships read

$$\begin{Bmatrix} \epsilon_x \\ \epsilon_y \\ \gamma_{xy} \end{Bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{16} \\ a_{12} & a_{22} & a_{26} \\ a_{16} & a_{26} & a_{66} \end{bmatrix}_k \begin{Bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{Bmatrix} \qquad (2)$$

where a_{ij} are elastic compliances and the subscript "k" indicates the number of a layer. The linear in-plane kinematic equations are

$$\epsilon_x = \frac{\partial u}{\partial x} \qquad \epsilon_y = \frac{\partial v}{\partial y}$$

$$\gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \tag{3}$$

u and v being displacements in the x and y directions, respectively.

The equilibrium equations are identically satisfied using the stress function that is introduced via

$$\{\sigma_x,\ \sigma_y,\ \tau_{xy}\} = \left\{\frac{\partial^2}{\partial y^2},\ \frac{\partial^2}{\partial x^2},\ -\frac{\partial^2}{\partial x \partial y}\right\}\psi \tag{4}$$

The stress function must be determined from the compatibility equation

$$\frac{\partial^2 \epsilon_x}{\partial y^2} + \frac{\partial^2 \epsilon_y}{\partial x^2} = \frac{\partial^2 \gamma_{xy}}{\partial x \partial y} \tag{5}$$

that, upon the substitution of equations (2) and (4), yields the well-known result:

$$a_{11}\frac{\partial^4 \psi}{\partial y^4} - 2\,a_{16}\frac{\partial^4 \psi}{\partial x\,\partial y^3} + (2\,a_{12} + a_{66})\frac{\partial^4 \psi}{\partial x^2\,\partial y^2} - 2\,a_{26}\frac{\partial^4 \psi}{\partial x^3\,\partial y} + a_{22}\frac{\partial^4 \psi}{\partial x^4} = 0 \tag{6}$$

Introducing a complex variable $z = x + \mu y$ where μ is a complex number, one can reduce equation (6) to an algebraic fourth-power equation for μ that has two pairs of conjugate complex or

imaginary roots [1]. Straightforward transformations yield the following expressions for the stresses and displacements:

$$\sigma_x = 2\ Re\left[\mu_1^2\phi_1'(z_1) + \mu_2^2\ \phi_2'(z_2)\right]$$

$$\sigma_y = 2\ Re\left[\phi_1'(z_1) + \phi_2'(z_2)\right]$$

$$\tau_{xy} = -2\ Re\left[\mu_1\ \phi_1'\ (z_1) + \mu_2\ \phi_2'\ z_2)\right]$$

$$u = 2\ Re\left[p_1\ \phi_1\ (z_1) + p_2\ \phi_2\ (z_2)\right]$$

$$v = 2\ Re\left[q_1\ \phi_1\ (z_1) + q_2\ \phi_2\ (z_2)\right] \qquad (7)$$

Where $\phi\ (z) = \psi'\ (z)$, and z_1 and z_2 are the roots of the above mentioned algebraic equation (not a conjugate pair), the prime indicates differentiation with respect to the complex variable, and complex constants are defined by

$$p_j = \mu_j^2\ a_{11} + a_{12} - \mu_j\ a_{16}$$

$$q_j = \mu_j\ a_{12} + a_{22}\ \mu_j^{-1} - a_{26} \qquad (8)$$

The functions $\phi_j(z_j)$ must be specified from the boundary conditions.

2. Three-Layered Anisotropic Plate With a Crack in the Central Layer

In the absence of adjacent layers, the crack tip stresses and displacements in a generally orthotropic layer subjected to exter·l in-plane loads are [2]:

$$\sigma_x = \frac{K_I}{\sqrt{2\pi r}}\ Re\left[\frac{\mu_1\mu_2}{\mu_1 - \mu_2}\ (\mu_2 F_2 - \mu_1 F_1)\right] + \frac{K_{II}}{\sqrt{2\pi r}}\ Re\left[\frac{1}{\mu_1 - \mu_2}\ (\mu_2^2 F_2 - \mu_1^2 F_1)\right]$$

$$\sigma_y = \frac{K_I}{\sqrt{2\pi r}}\ Re\left[\frac{1}{\mu_1 - \mu_2}(\mu_1 F_2 - \mu_2 F_1)\right] + \frac{K_{II}}{\sqrt{2\pi r}}\ Re\left[\frac{1}{\mu_1 - \mu_2}\ (F_2 - F_1)\right]$$

$$\tau_{xy} = \frac{K_I}{\sqrt{2\pi r}} Re\left[\frac{\mu_1 \mu_2}{\mu_1 - \mu_2}(F_1 - F_2)\right] + \frac{K_{II}}{\sqrt{2\pi r}} Re\left[\frac{1}{\mu_1 - \mu_2}(\mu_1 F_1 - \mu_2 F_2)\right]$$

$$\frac{\pi u}{\sqrt{2\pi r}} = K_I Re\left[\frac{1}{\mu_1 - \mu_2}\left(\frac{\mu_1 p_2}{F_2} - \frac{\mu_2 p_1}{F_1}\right)\right] + K_{II} Re\left[\frac{1}{\mu_1 - \mu_2}\left(\frac{p_2}{F_2} - \frac{p_1}{F_1}\right)\right]$$

$$\frac{\pi v}{\sqrt{2\pi r}} = K_I Re\left[\frac{1}{\mu_1 - \mu_2}\left(\frac{\mu_1 q_2}{F_2} - \frac{\mu_2 q_1}{F_1}\right)\right] + K_{II} Re\left[\frac{1}{\mu_1 - \mu_2}\left(\frac{q_2}{F_2} - \frac{p_1}{F_1}\right)\right] \quad (9)$$

where K_I and K_{II} are the stress intensity factors and

$$F_j = (\cos\theta + \mu_j \sin\theta)^{-1/2} \quad (10)$$

The study of Williams [9] resulted in the conclusion that, with the exception of extremely anisotropic materials, these factors can be taken as in the isotropic fracture mechanics, i.e.

$$K_I = \sigma\sqrt{\pi a}, \qquad K_{II} = \tau\sqrt{\pi a} \quad (11)$$

where σ and τ are applied stresses and a is a half-length of the straight crack.

Williams indicated that the stresses given by (9) can be represented in terms of real functions if μ_1 and μ_2 are imaginary numbers (the situation that exists in a subcase of specially orthotropic plates). However, such a representation is possible for arbitrary laminated plates in a general case where

$$\mu_j = a_j + i\, b_j \quad (12)$$

The substitution of μ_j into F_j given by (10) yields

$$F_j = (A_j + i\, B_j)^{-1} \quad (13)$$

where A_j and B_j are real numbers given by

$$A_j = \left\{\frac{1}{2}\left[c_j + \left(c_j^2+d_j^2\right)^{\frac{1}{2}}\right]\right\}^{\frac{1}{2}}$$

$$B_j = \left\{\frac{1}{2}\left[-c_j + \left(c_j^2+d_j^2\right)^{\frac{1}{2}}\right]\right\}^{\frac{1}{2}} \qquad (14)$$

In equation (14),

$$c_j = \cos\theta + a_j \sin\theta$$

$$d_j = b_j \sin\theta \qquad (15)$$

The substitution of F_j into the expressions for the stresses and displacements (9) yields the following result:

$$\sigma_x = \frac{K_I}{\sqrt{2\pi r}} S_1 + \frac{K_{II}}{\sqrt{2\pi r}} S_2$$

$$\sigma_y = \frac{K_I}{\sqrt{2\pi r}} S_3 + \frac{K_{II}}{\sqrt{2\pi r}} S_4$$

$$\tau_{xy} = \frac{K_I}{\sqrt{2\pi r}} S_5 + \frac{K_{II}}{\sqrt{2\pi r}} S_6$$

$$\frac{\pi u}{\sqrt{2\pi r}} = K_I S_7 + K_{II} S_8$$

$$\frac{\pi v}{\sqrt{2\pi r}} = K_I S_9 + K_{II} S_{10} \qquad (16)$$

Real functions $S_1(\theta), S_2(\theta), \ldots, S_{10}(\theta)$ are given in Ref. 12 Equations (16) were obtained for an anisotropic plate with a through crack. If this plate is a central layer in a symmetric three-layered laminate, the crack tip stresses and displacements will be reduced due to the presence of tougher outer layers.

$$U_1\left(u_i' u_j\right) = \sum_{k=1}^{3} h_k \int_A \left(\sigma_x^0 \epsilon_x + \sigma_x \epsilon_x^0 + \sigma_y^0 \epsilon_y \right. \tag{25}$$

$$\left. + \sigma_y \epsilon_y^0 + \tau_{xy}^0 \gamma_{xy} + \tau_{xy} \gamma_{xy}^0\right)_k dA$$

where A is the laminate planform area, k is the number of a layer, h_k is its thickness and the terms with the superscript "0" are obtained for the intact plate.

It is shown in [12], that the right side of equation (25) can be represented in the form:

$$U_1\left(u_i' u_j\right) = 2\int_A \sum_{k=1}^{3} \left\{\sigma_x^{0k} h_k \left[\frac{K_I}{\sqrt{2\pi r}}\left(a_{11}^{(k)} S_1^{(k)} + a_{12}^{(k)} S_3^{(k)} + a_{16}^{(k)} S_5^{(k)}\right)\right.\right.$$

$$\left. + \frac{K_{II}}{\sqrt{2\pi r}}\left(a_{11}^{(k)} S_2^{(k)} + a_{12}^{(k)} S_4^{(k)} + a_{16}^{(k)} S_6^{(k)}\right)\right]$$

$$+\sigma_y^{0k} h^{(k)}\left[\frac{K_I}{\sqrt{2\pi r}}\left(a_{12}^{(k)} S_1^{(k)} + a_{22}^{(k)} S_3^{(k)} + a_{26}^{(k)} S_5^{(k)}\right)\right. \tag{26}$$

$$\left. + \frac{K_{II}}{\sqrt{2\pi r}}\left(a_{12}^{(k)} S_1^{(k)} + a_{22} S_4^{(k)} + a_{26}^{(k)} S_6^{(k)}\right)\right]$$

$$+\tau_{xy}^{0k} h_k\left[\frac{K_I}{\sqrt{2\pi r}}\left(a_{16}^{(k)} S_1^{(k)} + a_{26}^{(k)} S_3^{(k)} + a_{66}^{(k)} S_5^{(k)}\right)\right.$$

$$\left.\left. + \frac{K_{II}}{\sqrt{2\pi r}}\left(a_{16}^{(k)} S_2^{(k)} + a_{26}^{(k)} S_4^{(k)} + a_{66}^{(k)} S_6^{(k)}\right)\right]\right\} dA$$

All terms and functions in equation (26) are real. Note that equation (26) is obtained by assumption that the ratio of the plate length to the crack length is so large that the effects associated with finite plate dimensions can be disregarded.

The evaluation of the integral in the denominator in the right side of equation (22) is simplified by using the expressions for the strain energy density function associated with crack tip deformations [8, p. xxxi]:

$$U_2\,(u_i\ u_j) = 2\sum_{k=1}^{3} h_k \int_A \frac{s_k}{r}\,dA \tag{27}$$

where

$$s_k = A_{11}^{(k)}\ K_I^2 + 2\ A_{12}^{(k)}\ K_I\ K_{II} + A_{22}^{(K)}\ K_{II}^2 \tag{28}$$

The functions $A_{ij}^{(k)}$ in equation (28) are given in [8]. In the present paper, the functions $A_{ij}^{(k)}$ are evaluated in terms of real functions $S_i^{(k)}$, $S_2^{(K)}$, . . ., $S_6^{(k)}$ defined in [12] for a general case where μ_j and F_j are given by equations (12) and (13), respectively:

$$A_{11} = \frac{1}{4}\left(a_{11}\ S_1^2 + a_{22}\ S_3^2 + a_{66}\ S_5^2 + 2a_{12}\ S_1\ S_3 + 2a_{16}\ S_1\ S_5 + 2a_{26}\ S_3\ S_5\right)$$

$$A_{12} = \frac{1}{4}\left[\ a_{11}\ S_1\ S_2 + a_{22}\ S_3\ S_4 + a_{66}\ S_5\ S_6 + a_{12}\ (\ S_2\ S_3 + S_1\ S_4\ )\right.$$

$$\left. + a_{16}\ (\ S_2\ S_5 + S_1\ S_6\ ) + a_{26}\ (S_3\ S_6 + S_4\ S_5\ )\right]$$

$$A_{22} = \frac{1}{4}\left(a_{11}\ S_2^2 + a_{22}\ S_4^2 + a_{66}\ S_6^2 + 2a_{12}\ S_2\ S_4 + 2a_{16}\ S_2\ S_6 + \right. \tag{29}$$

$$\left. 2a_{26}\ S_4\ S_6\ \right)$$

The superscripts (k) indicating that the corresponding quantities must be calculated for the k-th layer are omitted in equations (29).

Example. Correction Factor for Symmetrically Laminated Cross-Ply Plates Subjected to Mode I Loading

In the particular case shown in Fig. 2 the inner layer is specially orthotropic. The outer arrester layers are also arranged so that the fibers direction coincides with one of the plate coordinate axes.

The correction factor can be calculated from the condition that the applied force must be equal to the force acting in the

y-direction in the cross section y=0. In the case of a central crack shown in Fig 2 this condition reads:

$$f \int_0^{(h+h_1)} \int_0^{(L-a)} (\sigma_y + \sigma_0)\, dr\, dz = \sigma_0\, L\, (h + h_1) \qquad (30)$$

where σ_y is the crack tip stress at y=0, h and h_1 are the outer layer thickness and the inner layer half-thickness, respectively, and the coordinate x is replaced with the radial coordinate counted in the same direction from the crack tip. The finiteness factor f is introduced to account for the finite dimensions of the plate.

The total stresses in the inner layer in front of the crack, calculated in the absence of the arrester layers, are

$$\sigma_x^{(2)} = f \frac{K_I}{\sqrt{2\,\pi\, r}}\, S_I(0)$$

$$\sigma_y^{(2)} = f \frac{K_I}{\sqrt{2\,\pi\, r}}\, S_3(0) + \sigma_0 \qquad (31)$$

where

$$S_1\,(0) = b_1\, b_2 - a_1\, a_2$$

$$S_3\,(0) = 1 \qquad (32)$$

The finiteness factor evaluated for the inner layer ($h_1 = 0$) from equation (30) is

$$f = \left[1 - \frac{a}{L} + \sqrt{2\,\frac{a}{L}\left(1 - \frac{a}{L}\right)}\right]^{-1} \qquad (33)$$

Now the vector of in-plane stresses in the outer layers can be calculated as a function of the vector of stresses in the inner layer:

$$\{\bar{\sigma}\} = n\,[c]_0\,[a]_i\,\{\bar{\sigma}\}_i \tag{34}$$

where $[c]_0$ and $[a]_i$ are the matrix of stiffnesses for the outer layer and that of compliances for the inner layer, respectively. Equation (34) yields

$$\sigma_y^{(1)} = nf\sigma_0 \left(\frac{a}{2r}\right)^{\frac{1}{2}} C \tag{35}$$

where the superscript (1) refers to the outer layers and

$$C = \left[c_{12}^{(1)}\, a_{11}^{(2)} + c_{22}^{(1)}\, a_{12}^{(2)}\right] S_1(0) + c_{12}^{(1)}\, a_{12}^{(2)} + c_{22}^{(1)}\, a_{22}^{(2)} \tag{36}$$

The correction factor can be calculated using equation (30):

$$n = \frac{1 + \bar{h}}{1 + \bar{h}\,C} \tag{37}$$

where $\bar{h} = h_1/h$.

Note that equation (37) was obtained by assumption that in-plane stresses are applied to all layers of the three-layered plate. This means that the external load is increased proportionally to the thickness of the added arrester layers. If the total external load is not increased with the addition of outer layers, the correction factor given by equation (37) must be divided by $1 + \bar{h}$, *i.e*,

$$n_1 = \frac{n}{1 + \bar{h}} \tag{38}$$

The correction factor decreases with an increase of the arrester layers thickness. As follows from equation (37), even very thin (but tough) arrester layers result in a dramatic drop of stresses in the inner layer. An increase of C corresponding to relatively tougher arrester layers is very efficient for a stress reduction in the inner layer.

Example. Application to Annular Plates

Correction factors were calculated for an isotropic annular plate subjected to axisymmetric in-plane loading by the stresses

p_0 and reinforced by thin isotropic outer layers (Fig 3). The solution for a single-layered plate is well-known [13].

Contrary to the previous discussions, in this problem there is no need to subdivide displacements into those for the intact plate and those due to the pressure of the hole. Otherwise, the approach to the calculation of the correction factor does not differ from that in the case of a crack. Transformations yield

$$n = \left(1 + \frac{h_1}{h}\right) \frac{1 + \nu_i + \left(\frac{b}{a}\right)^2 (1 - \nu_i)}{1 + \nu_i + \left(\frac{b}{a}\right)^2 (1 - \nu_i) + \frac{h_1}{h} \frac{E_o}{E_i} \left[\frac{(1 + \nu_i)^2}{1 + \nu_o} + \left(\frac{b}{a}\right)^2 \frac{(1 - \nu_i)^2}{1 - \nu_o}\right]} \qquad (39)$$

where the subscripts i and o indicate the inner and outer layers, respectively, h is a half-thickness of the central layer, h_1 is the outer layer thickness, and E and ν are the elasticity moduli and the Poisson's ratios.

Note that equation (39) was obtained by the assumption that the in-plane stresses are applied to all layers of the three-layered plate. If the total load remains unchanged after an addition of the outer layers, the correction factor can be calculated by equation (38).

The effect of the outer layers is shown in Fig 4. As follows from Fig 4, even the addition of very thin outer layers, whose total thickness is equal to only 1% of the inner layer thickness, can result in a significant reduction of stresses in the latter layer, provided the elasticity moduli ratio is large. Excellent results can be achieved using thicker outer layers, even if their moduli of elasticity are only 2 or 3 times larger than those of the inner layer (curve 3). Note that if the total load on the plate is not increased, as compared to that for a single inner layer, the correction factors shown in Fig 4 should be reduced even more.

It is interesting to note that as follows from equation (39) the outer-to-inner radii ratio does not affect the correction factor, if the Poisson's ratios $\nu_i = \nu_0$.

Conclusions

Theoretical foundations of a crack arrest technique based on incorporation of thin tough layers into the original composite plate are considered. Stresses and displacements in both damaged and arrester layers can be obtained in terms of real functions. Expressions for these functions are obtained in a general case of arbitrary oriented cracks in arbitrary laminated layers. An analytical expression for the correction factor that accounts for the effect of arrester layers is obtained in terms of real functions using the principle of virtual work.

References

1. Lekhnitskii, S. G., Theory of Elasticity of an Anisotropic Elastic Boyd, Holden-Day, San Francisco, 1966.

2. Sih, G. C., Paris, P. C. and Irwin, G. R., "On Cracks in Rectilinearly Anisotropic Bodies," International Journal of Fracture Mechanics, Vol 1, pp. 189-196.

3. Sih, G. C., and Leibowitz, H., "Mathematical Theory of Brittle Fracture," in Fracture, Vol. II, ed Leibowitz, H., Academic Press, New York, pp. 67-190, 1968.

4. Wang, S. S., Yau, J. F. and Corten, H. T., "A Mixed-Mode Crack Analysis of Rectilinear Anisotropic Solids Using Conservation Laws of Elasticity," International Journal of Fracture, Vol 16, pp 247-259, 1980.

5. Becker, W., "A Complex Potential Method for Plate Problems with Bending Extension Coupling," Archive of Applied Mechanics, Vol 61, pp 318-326, 1991.

6. Becker, W., "Closed-Form Analytical Solutions for a Griffith Crack in a Non-Symmetric Laminate Plate," Composite Structures, Vol 21, pp 49-55, 1992.

7. Parhizgar, S. Zachary, L. W. and Sun, C. T., "Application of the Principles of Linear Fracture Mechanics to the Composite Materials," International Journal of Fracture, Vol 20, pp 3-15, 1982.

8. Sih, G. C. and Chen, E. P., Cracks in Composite Materials, Martinus Nijhoff Publishers, The Hague, 1981.

9. Williams, J. G., "Fracture Mechanics of Anisotropic Materials," in Application of Fracture Mechanics to Composite Materials, ed. Frederick, K., Elsevier, Amsterdam, pp 3-38, 1988.

10. Pindera, M. J., Williams, T. O. and Macharet, Y., "Time-Dependent Response of Aramid-Epoxy-Aluminum Sheet, ARALL, Laminates," Polymer Composites, Vol 10, pp 328-336, 1989.

11. Chen, E. P. and Sih, G. C., "Stress Intensity Factor for a Three-Layered Plate With a Crack in the Center Layer," Engineering Fracture Mechanics, Vol 14, pp. 195-214, 1980.

12. Birman, V. and Nagar A., "Crack Arrester Layer Technique in Composite Plates," 1992 ASME Winter Annual Meeting, Proceedings. In press.

13. Rice, J.R., "Mathematical Analysis in the Mechanics of Fracture," in Fracture, Vol III, ed. Leibowitz, H., Academic Press, New York, pp. 191-311, 1968.

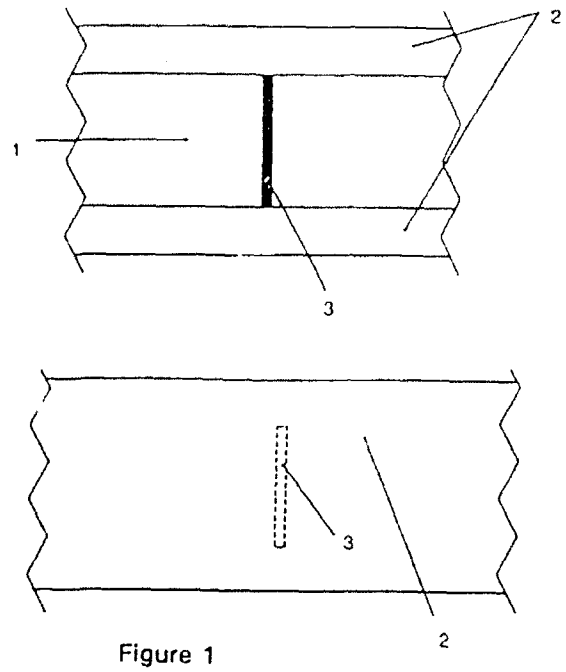


Figure 1

Three-layered plate. 1 = inner, cracked layer; 2 = outer arrester layers; 3 = straight crack.

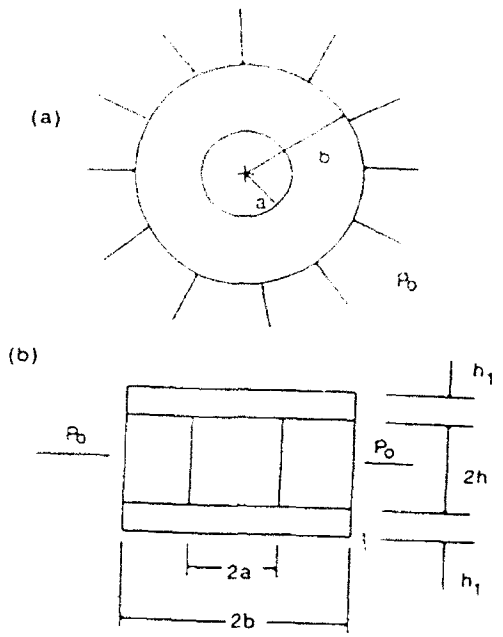


Figure 3

Annular plate subjected to axisymmetric in-plane loading. (a) = plan view without outer layers; (b) = cross section with outer layers.

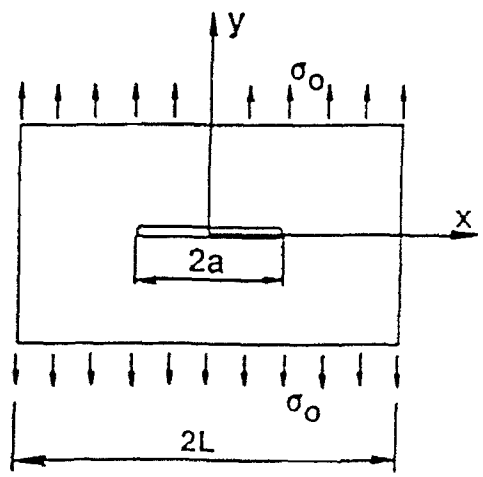


Figure 2

Plate with a central crack subjected to Mode I loading

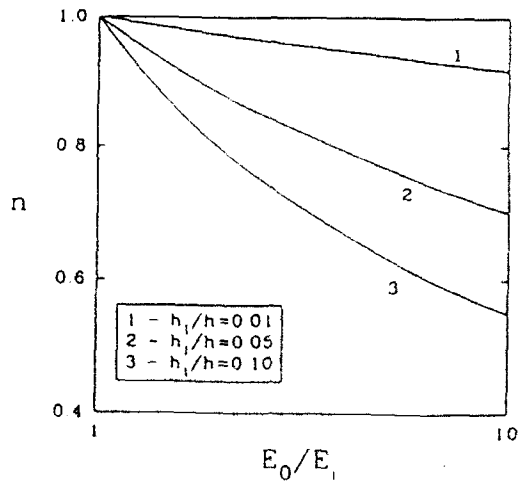


Figure 4

Correction factors for three-layered annular plates, b/a = 2, $\nu_1 = \nu_0 = 0.3$.

SOME RESULTS IN MACHINE-LEARNING

Mike Breen
Assistant Professor
Department of Mathematics

Tennessee Technological University
Box 5054
Cookeville, Tennessee 38505

Final Report for:
AFOSR Summer Faculty Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

August, 1992

SOME RESULTS IN MACHINE-LEARNING

Mike Breen
Assistant Professor
Department of Mathematics
Tennessee Technological University

Abstract

The research here focused on pattern-based machine-learning. A machine's performance is modeled by a function. Previously, results have been obtained for functions of less than ten variables. For functions with more variables, new techniques are needed because of the huge amounts of memory required. We researched the topic of column multiplicity and obtained results that may make the learning task practical for these larger functions. We also described how some future problems involving noise might be reframed to resemble problems already solved by Dr. Tim Ross and the AART-2 group. We also assisted in an experiment to determine the relationship between human-learning and machine-learning.

SOME RESULTS IN MACHINE-LEARNING

Mike Breen

INTRODUCTION

The System Concepts Group (WL-AART-2) studies Pattern-Based Machine-Learning (PBML) in an attempt to have avionics systems learn a function as best as they can with as little information about the function as is possible. The performance of the systems is represented by a function in that the system receives an input and when that input is received a certain output is determined. It is either impractical or impossible to list all possible inputs that the system will encounter. So, by examining a portion of this list, we hope to find a pattern that exists in the entire list of inputs.

Dr. Tim Ross has been running experiments of this nature for two years. The Function Extrapolation by Recomposing Decompositions (FERD) technique has performed on functions as diverse as the world itself. Some are mathematical formulas, some are visual patterns, some involve data compression. In all the cases, FERD learns the desired function when presented with only a small amount of information and learns it without being told what type of function it is. FERD has consistently outperformed other machine-learning techniques. This paper addresses some issues involved with FERD.

I would like to thank Dr. Tim Ross and all the people in AART-2 for the opportunity to work on this project. It is a rewarding and enjoyable experience. I am also extremely grateful to AFOSR and RDL for the chance to participate in the program. Thanks also to Melvin Richardson who did all the work in the experiment mentioned at the end of this paper.

MAIN RESULTS

I. Column Multiplicity

FERD measures the simplicity of a function by a measure called the Decomposed Function Cardinality (DFC). For details on this and the background to what follows, see [4]. We call the function that we are trying to learn, f where $f : \{0, 1, \ldots, 2^n - 1\} \longrightarrow \{0, 1\}$. So, f is defined on all n- digit binary numbers and thus on n binary variables. This situation may not look general, but many applications can be translated to this situation.

To measure DFC, first the n variables, $x_1, x_2, \ldots x_n$ are partitioned into two sets called the row variables and the column variables. The function is then written as a matrix where

if $f(x) = y$, then y appears in the position of the matrix that corresponds to the values that the row and column variables have in the binary representation of x.

Once a partition has been chosen, the next step in finding the DFC is determining how many columns in the matrix are equal. All other things remaining the same, the fewer the number of columns, the smaller the DFC. The number of columns is signified by ν.

Most often, the value for ν will depend on the partition. However, some functions will have the same value for ν regardless of the partition used. The proof is omitted because of its tedium.

THEOREM 1. *If a function, f, has the property that ν is constant for all partitions, P, then either f is*

(1) *constant ($\nu = 1$), or*
(2) *a parity function ($\nu = 2$), or*
(3) *such that either $|\{x : f(x) = 1\}| = 1$ or $|\{x : f(x) = 0\}| = 1$ ($\nu = 2$).*

There is a relationship between rows and columns in the partition matrix. Let the partition, P, have r row variables and $n - r$ column variables. Let ρ be the number of different rows. Then $\rho \leq 2^{\nu}$. By symmetry, $\nu \leq 2^{\rho}$, so $\log \rho \leq \nu \leq 2^{\rho}$, and $\log \nu \leq \rho \leq 2^{\nu}$, where log denotes the logarithm base two. Suppose we are interested in moving one of the row variables, x_i, in the partition P so that it is one of the column variables, in a new partition $\hat{P}$. Then $\hat{P}$ has $r - 1$ row variables and $n - r + 1$ column variables. Denote the column multiplicity of P and $\hat{P}$ by ν_P and $\nu_{\hat{P}}$, respectively. Then,

$$2\sqrt{\nu_P} \leq \nu_{\hat{P}} \leq 2\nu_P \text{ and } \frac{\nu_{\hat{P}}}{2} \leq \nu_P \leq \frac{\nu_{\hat{P}}^2}{4}.$$

The right side of the first inequality and the left side of the second are established in [5]. The other sides are easily established.

For machine–learning, we don't know the entire function, that is what we are trying to learn. So it is impossible to fill in the entire partition matrix with zeros and ones. We must also have values that we don't know or don't care about. They are often symbolized with an x, and are called **don't cares.**

Given two columns, c and k, we say that c and k are **incompatible** if $\exists i$ such that $c_i = 0$ and $k_i = 1$ or $c_i = 1$ and $k_i = 0$. Otherwise, c and k are called **compatible.** *Consider the following partition,* P.

$$\begin{array}{ccccc} 1 & 1 & x & x & 0 \\ x & 0 & x & 0 & x \\ 0 & x & 1 & x & 1 \\ 1 & x & 1 & x & 1 \end{array}$$

If we start with column 1 and proceed left to right, we arrive at the representative 1001^t when the first two columns are used. Column 3 is not compatible with this representative so it starts a new class. Column 4 is compatible with either representative. Let us put it in class 1. The representative doesn't change. Column 5 is compatible with the second class and the representative is $0x11^t$. Thus, we would conclude that $\nu_P = 2$. However, if we had started with Column 2 at the very beginning, and checked the columns in the order 2,3,4,5,1; then we would get $1011^t, 0x11^t$, and $1x01^t$ for our three representatives and thus $\nu_P = 3$. **The procedure we use affects the column multiplicity.**

Suppose P was part of a much larger partition matrix Q a $2^r \times 2^{n-r}$ matrix, where P appeared on the diagonal blocks of Q and the vector, v, is any column vector not compatible with the columns of P, for example $v = 0100^t$, appeared everywhere else. Assume we have p copies of P where $p = \min([\frac{2^{n-r}}{5}], \frac{2^r}{4}), r > 1$. Then the maximum column multiplicity we could arrive at would be $3p$ and the minimum would be $2p$. The difference is p which is a significant difference for large n.

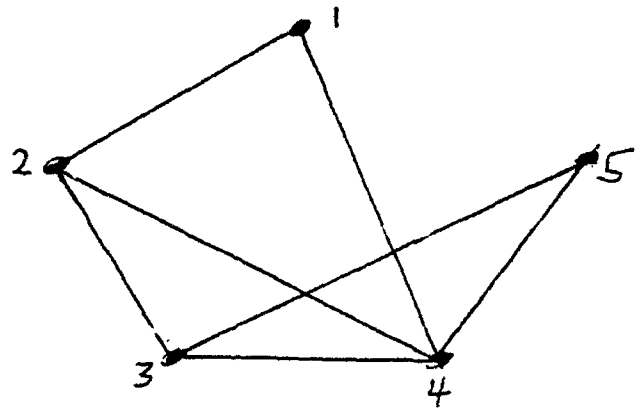
The following is a partial solution to the problem of finding the minimum column multiplicity for a given partition P.

Let $C_1, C_2, ..., C_m$ be the columns of P. Define an undirected graph, G_P, as follows. The vertices of $G_P, V(G_P)$, are the columns of P. There is an edge from C_i to C_j if and only if the two columns are compatible. A graph is called **complete** if there is an edge between any pair of distinct vertices, that is each pair of distinct vertices is adjacent. Define a **complete cover**,$\mathcal{G}$, of a graph G as a collection of graphs $\{G_i\}_{i=1}^n$ with the property that $\bigcup_{i=1}^n V(G_i) = V(G)$, and for each i the edges of G_i are edges of G. That is, $E(G_i) \subseteq E(G)$.

PROPOSITION 2. *The minimal column multiplicity for P is the minimal cardinality of all complete covers of G.*

Proof: The compatibility relation we have used to determine ν is an equivalence relation and the resulting classes are equivalence classes. Therefore, the subgraphs that represent the equivalence classes are complete. So, $\min \nu$ is $\min |\mathcal{G}|$. ∎

For our P, G_P is



and three complete covers are

$$\mathcal{G}_1 = \{\{1,2,4\},\{3,5\}\},$$

$$\mathcal{G}_2 = \{\{2,3,4\},\{1\},\{5\}\},$$

$$\mathcal{G}_3 = \{\{1,4\},\{2,3\},\{5\}\},$$

and the minimal column multiplicity is 2.

There is no efficient algorithm to determine $\min|\mathcal{G}|$. As we shall see later, determining this number for a graph is an NP–complete problem [3]. We propose some approaches below. The first set offers choices on which column (vertex) to choose. The second set offers choices on which columns (vertices) to choose to associate with that column (vertex). The **degree** of a vertex is the number of vertices adjacent to it.

SET I. Column choice

A. The vertex of minimal degree. If there is a tie, choose the vertex such that the set of all vertices adjacent to that vertex has a vertex with minimal degree.

B. If the care's have more 0's than 1's, choose the column with the most 1's. If there is a tie, choose the column that has a 1 in a coordinate that disagrees with the most columns possible. The procedure is similar if there are more 1's than 0's in P.

C. The column with the most care's.

D. The column farthest to the left.

SET II. Choice of Columns

A. Once a vertex has been chosen, choose the largest subset of vertices of G_P such that the columns represented by those vertices are compatible with the original vertex and with each other.

B. Once a vertex has been chosen, choose vertices a vertex at a time whose columns are compatible with the current representative as determined now.

C. Once a vertex has been chosen, choose vertices to associate with that one such that the representative for the class will change in as few coordinates as possible.

The approaches above give us $\nu = 2$. However, I-D and II-B are servants whose master is the method of enumeration.

If G_P is encoded as a matrix with $(i, j) = 1$ if and only if column i and column j are not compatible then we expect to have a sparse matrix. We can make some progress by deleting copies of rows (rows that are equal to others).

Experiments on large partitions can tell us which, if any, of these algorithms combine efficiency and cost in the best way.

Now let us see why this problem is NP-complete. Each of the complete graphs is called a **clique**, so we seek the minimal number of cliques covering G. Let $\bar{G}$ be the graph with the same vertices as G but with the property that two vertices in $\bar{G}$ are adjacent if and only if they are not adjacent in G.

A **coloring** of a graph is done so that no two adjacent vertices are colored the same color. There is a connection between the ideas of cliques and coloring.

THEOREM 3. *The minimal number of cliques covering G is equal to the minimal number of colors needed to color* $\bar{G}$.

Proof: Call the numbers m_G and $c_{\bar{G}}$, respectively.

i. $m_G \leq c_{\bar{G}}$

Given a coloring, vertices with the same color in $\bar{G}$ must not be adjacent in that graph. Thus, within a color, all the vertices are adjacent in G, so we have cliques in G. The minimal number of cliques covering G, then, is no more than the minimal number of colors necessary to cover $\bar{G}$.

ii. $c_{\bar{G}} \leq m_G$

Given a collection of cliques covering G, we are presented with a collection of complete graphs whose union contains all the vertices of G. Within each clique, all the vertices are adjacent in G, hence in $\bar{G}$, each of the vertices will not be adjacent. Therefore, they can be colored with the same color. Hence, the minimal number of colors must be no more than this number.

The two parts of the proof give us the desired conclusion. ∎

All the facts in this paragraph and in the numbered list below are from [3]. As stated before, there is no quick algorithm to find c_G for all graphs G. The problem is NP–complete. If a graph G meets certain conditions then c_G can be found in polynomial (in the number of vertices) time, but the compatibility graph or its complement will probably not meet these conditions. In the list below V stands for the set of vertices in a graph.

1. $m_G \leq 1 +$ maximum degree in G ,
2. $m_G \leq 1 + \max_{G' \subseteq G}\{\text{minimum degree in} G'\}$,
3. $m_G \leq 1 +$ the length of the longest path in G ,
4. $m_G \leq \max_{1 \leq i \leq |V|}\{\min\{i, \deg(V_i) + 1\}\}$,
5. $2\sqrt{|V|} \leq m_G + m_{\bar{G}} \leq |V| + 1$,
6. $|V| \leq m_G \cdot m_{\bar{G}} \leq \frac{(|V|+1)^2}{2}$.

The compatibility matrix can be reduced by first eliminating equal rows and columns. Here we are assuming that our compatibility matrix is square, symmetric, and has 1's on the main diagonal. For example if row 1 equals row 5 we would eliminate row and column 1 from the compatibility matrix and remember to identify column 1 and column 5 in the partition matrix. They would be in the same clique.

Once this is done, we define a partial order on the rows of the compatibility matrix as follows: row i $\leq$ row j if and only if $(i, k) = 1$ implies that $(j, k) = 1$ for all k. So that in the partition matrix, every column that column i is compatible with, is also compatible with column j.

Determine the partial order relation on the rows of the compatibility matrix and construct chains of rows ($r_1 \leq r_2 \leq \cdots \leq r_t$). The greatest row in a chain can be identified with the next greatest row, in the sense that whatever clique the next greatest row will belong to, so will the greatest row. The greatest row and its column can be removed from the compatibility matrix. Some rows might be the greatest in more than one chain. In that case, some choice must be made as to which chain to identify it with. Proceed with the chains each time identifying a row with the row immediately less than it in the chain. Then remove the larger row and its column from the compatibility matrix.

When this has been done, only rows that do not have the partial order relation will be left. There will be fewer rows than when we started, but we still have an NP–complete problem.

So, we do not expect to have a realistic way to completely solve the problem of finding the column multiplicity of a function's partition matrix when the function is not completely known.

A rather astounding fact is that given a compatibility matrix for a partition P, we can find the column multiplicity for the partition, $\hat{P}$, obtained by moving one of the column variables in P to the set of row variables, **without finding** $\hat{P}$.

To make the notation easier, we call the column variables in a partition P $x_1, x_2, \ldots, x_c$, labeled bottom to top. The columns are labeled left to right 0,1,$\ldots, 2^c - 1$. Define the compatibility matrix, C_P, as follows: $(C_P)_{ij} = 1$ if and only if $i < j$ and column i is compatible with column j. It may be a bit unconventional but to correspond to our notation, the columns and rows of C_P are numbered 0 through $2^c - 1$. Furthermore, C_P is an upper triangular matrix where all the information is above the main diagonal. There are $2^{c-1}(2^c - 1)$ entries of importance in C_P.

Example 1. The matrix from the earlier discussion would be

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

which we could write as

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ & 1 & 1 & 0 \\ & & 1 & 1 \\ & & & 1 \\ & & & \end{bmatrix}.$$

Now suppose we wish to shift the column variable x_k to the set of row variables. Call this new partition $\hat{P}$. Then the two partition matrices might look as the partition matrices below.

A_1	$\cdots$	$A_{2^{k-1}}$	B_1	$\cdots$	$B_{2^{k-1}}$	$A_{2^{k-1}+1}$	$\cdots$	A_{2^k}	$\cdots$	A_{2^c}	$\cdots$	$\cdots$	B_{2^c}

A_1	$\cdots$	$A_{2^{k-1}}$	$A_{2^{k-1}+1}$	$\cdots$	A_{2^k}	$\cdots$	A_{2^c}
B_1	$\cdots$	$B_{2^{k-1}}$	$B_{2^{k-1}+1}$	$\cdots$	B_{2^k}	$\cdots$	B_{2^c}

The matrix for partition P is the one above. The A's and B's are vectors corresponding to columns where $x_k = 0$ and $x_k = 1$, respectively. For simplicity's sake, we have assumed that the top half of $\hat{P}$ corresponds to rows where $x_k = 0$. Even though this may not be the case, the placing of these rows in this manner does not affect either column multiplicity or the following discussion. In $\hat{P}$ we see that two columns are equal if and only if their corresponding halves from P are equal. This allows us to determine $C_{\hat{P}}$ without determining $\hat{P}$.

Let i and j be two columns in P. They have a binary representation with $c-1$ digits. Let i_0 be the number whose binary representation (with c digits) is the following:

$$(i_0)_d = \left\{ \begin{array}{lll} i_d & \text{if} & d < k \\ 0 & \text{if} & d = k \\ i_{d-1} & \text{if} & d > k \end{array} \right\} .$$

Thus, we obtain the binary expansion for i_0 in the following way: the first $k-1$ digits (reading right to left) of the binary expansion of i are unchanged, digit k is 0, and the remaining digits of the binary expansion of i are shifted one place to the left. Define i_1 to be $i_0 + 2^{k-1}$, so its binary expansion is the same as i_0's except the kth digit is 1. Define j_0 and j_1 similarly. Note that if x_k is the column variable in P that is moved to the group of row variables to form $\hat{P}$, then column i of $\hat{P}$ is formed from columns i_0 and i_1 of P.

Recall that we write iCj to mean that column i is compatible with column j.

PROPOSITION 4. *In $\hat{P}$, iCj if and only if in P, i_0Cj_0 and i_1Cj_1.*

Proof: If iCj in $\hat{P}$, then as we remarked before, the corresponding halves of the two columns must be equal. But these halves are just the places where (using the enumeration in P) all values except those of x_k are as in P and either simultaneously $x_k = 0$ or $x_k = 1$. Thus i_0Cj_0 and i_1Cj_1. The argument is reversible. ∎

Because of the relationship between i_0 and i_1 which also holds for the j's we have

PROPOSITION 5. *In $\hat{P}$, iCj if and only if in P, i_0Cj_0 and $(i_0 + 2^{k-1})C(j_0 + 2^{k-1})$.*

In the language of matrices,

PROPOSITION 6. *In $\hat{P}$, iCj if and only if in C_P, $(i_0, j_0) = (i_0 + 2^{k-1}, j_0 + 2^{k-1}) = 1$.*

Proof: This follows from Proposition 5 and from the definition of C_P. ∎

Example 2. Let

$$C_P = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ & 0 & 1 & 1 & 1 & 0 & 0 \\ & & 1 & 0 & 1 & 0 & 1 \\ & & & 0 & 1 & 1 & 0 \\ & & & & 0 & 0 & 1 \\ & & & & & 0 & 0 \\ & & & & & & 0 \end{bmatrix}$$

obtained from a partition P with three row variables and column variables. Denote the column variables x_1, x_2, x_3. Thus, the columns of the square matrix are numbered 000 through 111. Let P_k denote the partition that comes from P by moving variable x_k to the row variables. We will have four columns in each of these new partitions. Here is a table that may help. It gives the columns numbers in P that correspond to columns 0 through 3 in P_1, P_2, and P_3.

	$k=1$	$k=2$	$k=3$
i_0	0 2 4 6	0 1 4 5	0 1 2 3
i_1	1 3 5 7	2 3 6 7	4 5 6 7

We find C_{P_1}. To get its first row, we must see how the 0th column of P compares one at a time to the second column of P, the fourth column of P, and the sixth column of P; and simultaneously and respectively how the first column of P compares to the third, fifth, and seventh columns of P. The first row of C_{P_1} will be all zeros because $(C_P)_{02} = 0$ and $(C_P)_{13} = 1$; and $(C_P)_{04} = 0$ and $(C_P)_{15} = 1$; and $(C_P)_{06} = 1$ and $(C_P)_{17} = 0$. The rest of the entries in C_{P_1} are below as are the other two matrices.

$$\begin{bmatrix} 0 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ & 1 & 0 \\ & & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ & 0 & 0 \\ & & 0 \end{bmatrix}.$$

The column multiplicities for the three partitions are 4, 2, and 3, respectively.

We now make some remarks about the calculations.

(1) A total of $2^{c-2}(2^{c-1}-1)$ entries of C_P are examined.

(2) The examination is done on diagonals.

(3) To determine an entry in C_{P_k}, a pair is examined and the new matrix entry is 1 if and only if both entries in the pair from C_P are 1.

(4) If the diagonals of C_P are numbered from 1 to 2^{c-1} (if the entire square matrix was used the main diagonal would be diagonal zero, so this numbering is consistent with our earlier notation), from longest to shortest, then the entries of of C_{P_k} can be determined as follows: to find the (i,j)-th entry of C_{P_k}, look at the pair of entries (i_0, j_0) and $(i_1, j_1) = (i_0 + 2^{k-1}, j_0 + 2^{k-1})$. So you are looking at the i_0-th entry of diagonal $j_0 - i_0$ and the entry 2^{k-1} away from it. Only when both are 1 is $(C_{P_k})_{ij}$ equal 1.

(5) Not every entry in C_P is needed to determine $\{C_{P_k}\}_{k=1}^{c}$.

(6) Some entries are examined more than once.

So this can save us some time and computer memory, and these two quantities are the limiting forces when we get to functions of over one hundred variables.

The cases we have talked about have admitted three values in the partition matrix: 0, 1, and x. More generally, the values could be any real number between 0 and 1, inclusive. One interpretation of this case is where each P_{ij} is the probability that $f(x) = 1$ for the x-value represented by the (i,j)-th position in P. Hence, if $P_{ij} = 0$, then $f(x) = 0$ (for that x). If $P_{ij} = 1$, then $f(x) = 1$. Two columns, u and v, are incompatible if and only if they differ by one in at least one coordinate. Assume that each column has m coordinates. We propose a series of measures to describe the relative compatibility between two columns in P.

a. $D(u,v) = \sum_{i=1}^{m}(u_i - v_i)^2$,

b. $M(u,v) = \max_{1 \le i \le m} |u_i - v_i|$,

c. $M'(u,v) = 1 - M(u,v)$,

d. $P(u,v) = \prod_{1 \le i \le m} \left[1 - (u_i - v_i)^2\right]$,

e. $P'(u,v) = 1 - P(u,v)$.

Measures c and d have the advantage that they equal 0 if and only if the columns being compared are incompatible. They also equal 1 if and only if the columns are equal.

Measures a and b are somewhat familiar. They both satisfy the triangle inequality. The others don't. (Ex.: $u = \langle 0,1\rangle, v = \langle 1,0\rangle, w = \langle .5,.5\rangle$. Then $M'(u,w) = .5 > 0 + 0 = M'(u,v) + M'(v,w), P(u,w) = .5625 > 0 + 0 = P(u,v) + P(v,w), P'(u,v) = 1 > .4375 + .4375 = P'(u,w) + P'(w,v)$) The first measure is simply the square of the distance (Euclidean) between two vectors. However, its familiarity is offset by its inability to convey the fact that two columns are incompatible.

Here are some facts about the measures.

(1) If $u_i \leq v_i \leq w_i$ for all i, then $P(u,v) \geq P(u,w)$ so that $P'(u,v) \leq P'(u,w)$.

(2) $P(u+w, v+w) = P(u,v)$,

(3) If $D(u,v) \longrightarrow 0$, then $P(u,v) \longrightarrow 1$ so $P'(u,v) \longrightarrow 0$. However, if $P'(u,v) \longrightarrow 0$ then it may not be the case that $D(u,v) \longrightarrow 0$.

(4) $D(u,v) \longrightarrow 0 \Longleftrightarrow M(u,v) \longrightarrow 0$,

(5) $M(u,v) \leq \sqrt{D(u,v)}$.

We generated thirty 8×8 matrices whose elements were randomly selected from the uniform distribution on [0,1]. Each matrix can be thought of as the partition matrix for a function on six variables with three row variables and three column variables. Then we determined each of the measures' values for each pair of columns (twenty-eight pairs for each measure) in the partition matrix in order to obtain some feel for the measures and the relation between them. Below is one partition matrix where the values have been rounded to three places.

.911	.359	.139	.534	.349	.171	.172	.371
.040	.034	.788	.525	.048	.478	.251	.526
.170	.505	.234	.312	.810	.582	.064	.946
.299	.241	.679	.100	.573	.686	.142	.367
.428	.713	.125	.153	.335	.055	.955	.521
.823	.184	.571	.941	.831	.425	.034	.207
.218	.186	.282	.994	.859	.529	.112	.362
.347	.657	.728	.781	.152	.633	.925	.782

Here are the values for three of the measures on the columns of the above matrix. The other two measures' values are easily obtained from their complementary measures' values. Since each measure is symmetric (for example, $D(u,v) = D(v,u)$) and since each measures' value when the two columns are the same is known (for example, $M(u,u) = 0$), each matrix is a little less than half of an 8×8 matrix. Below and on the next page are the D-, M-, and P-measure matrix, respectively.

1.007	1.608	1.316	1.260	1.535	1.871	1.733
	1.392	1.882	1.473	1.045	0.0445	0.537
		1.214	1.711	0.315	1.652	1.027
			1.192	1.064	2.533	1.572
				0.864	2.991	1.358
					1.837	0.592
						1.246

.640	.772	.776	.641	.740	.789	.777
	.754	.808	.673	.657	.441	.493
		.712	.740	.348	.830	.713
			.629	.586	.907	.733
				.481	.737	.630
					.900	.466
						.883
.302	.110	.182	.209	.155	.075	.104
	.169	.070	.164	.291	.623	.551
		.220	.114	.719	.101	.284
			.247	.294	.010	.130
				.390	.013	.194
					.060	.529
						.134

After these values were determined for each of the thirty partition matrices, the correlation between each of the three measures' values was found. That is, for a given partition, the correlation coefficient for D and M was found (the correlation between the twenty–eight ordered pairs of the form $(D(u,v), M(u,v))$). The other two correlation coefficients (between M and P, and between D and P) were also found. This was done for each partition. Then the average correlation coefficient in each category for the thirty runs was found. We also found the (sample) standard deviation. The results are listed below.

	Average	Standard Deviation
(D,M)	.79519	.081165
(M,P)	−.89341	.041346
(D,P)	−.93297	.032581

This indicates a high correlation so that information about one measure can be used to make a good guess as to the information about another measure. The last two are negative because a small value for the P–measure indicates that the columns are far apart whereas a small value for the D– or M–measure would indicate that the columns are close.

Following these calculations and given a partition matrix and a measure, the values in the measure matrix were assigned their ordinal numbers with 1 indicating the closest pair of columns (which for D is the smallest value and for P is the largest) and 28 indicating the pair of columns that are the farthest apart. This was done for each measure and each

of thirty partitions and once again, the average and standard deviations in these correlation coefficients were found. Those results are below.

	Average	Standard Deviation
(D,M)	.782	.1091
(M,P)	.861	.0759
(D,P)	.971	.0358

Again a high correlation is indicated especially between the D and P measures.

There is also a way to determine these measures' values on $\hat{P}$ when the values for P are known. Recall that $\hat{P}$ is obtained from P by moving one of the column variables from P to the row variables. The results are as follows where the values to the left of the equal signs are associated with $\hat{P}$.

(1) $D(i,j) = D(i_0,j_0) + D(i_1,j_1)$,
(2) $M(i,j) = \max\{M(i_0,j_0), M(i_1,j_1)\}$,
(3) $M'(i,j) = \min\{M'(i_0,j_0), M'(i_1,j_1)\}$,
(4) $P(i,j) = P(i_0,j_0) \cdot P(i_1,j_1)$.

When a row variable from P is moved to become a column variable, then not as much information can be determined about the new partition matrix. We describe some rather unstartling facts about this situation in the list below. Denote the two columns in P by i and j. Then each of these columns is broken into two parts when placed into $\hat{P}$. Call the two parts of i, i_L and i_R, where L and R stand for left and right, since in $\hat{P}$ one part will be to the left of the other. The same labelling is done to j. Note that no information can be obtained from the measures' values on P about any of the measures' values on the pairs (i_L,j_R) or (i_R,j_L) in $\hat{P}$. However,

(1) $D(i_L,j_L) \leq D(i,j)$,
(2) $M(i_L,j_L) \leq M(i,j)$ and either $M(i_L,j_L)$ or $M(i_R,j_R)$ equals $M(i,j)$.
(3) $M'(i_L,j_L) \geq M'(i,j)$,
(4) $P(i_L,j_L) \geq P(i,j)$,
(5) $P'(i_L,j_L) \leq P'(i,j)$.

Similar results hold for the measures on (i_R,j_R).

II. Noise

We now address the following question: given a function, f, how much and what kind of noise can we add to the function's output so that after filtering, FERD can learn the

function? Included in this question is the question of what type of function (in terms of the level of complexity) can FERD recover given that a certain noise function is to be added?

Here are some symbols used and their meanings.

(1) x original function domain values,

(2) t noise–added probability density function values,

(3) t_x the t–value associated with a given x–value,

(4) t_0 one threshold value — the number below which f's ouput is assumed to be 0,

(5) t_1 the other threshold value — the number above which f's output is assumed to be 1, and

(6) p $\frac{|\{x:f(x)=0\}|}{|X|}$.

We first deal with the case where the noise added to the unknown function, f, is uniformally–distributed.

Assume the noise distribution is continuous on the interval $[-a, a]$. Initially, we assign the value m_0 to 0 and m_1 to 1. Then $t_0 = m_1 - a$ and $t_1 = m_0 + a$. If we assume that the signal is the difference between the values assigned to 0 and 1, then $S = m_1 - m_0$. The noise, N, is $2a$, so that the signal–to–noise ratio is $\frac{m_1 - m_0}{2a}$.

If c stands for the proportion of cares then

$$\begin{aligned} c &= P(t_0 < t < t_1) \\ &= p \cdot \left(\frac{m_1 - a - (m_0 - a)}{2a} \right) + (1 - p) \cdot \left(\frac{m_1 + a - (m_0 + a)}{2a} \right) \\ &= \frac{m_1 - m_0}{2a} . \end{aligned}$$

If the distribution is assumed to be discrete, then the calculations are almost the same and we arrive at $c = \frac{m_1 - m_0}{2a+1}$.

As an example, suppose we wanted to add noise that was uniformally distributed on the interval [-6,6] and we assigned the values 5 and 8 to the original function values 0 and 1, respectively. Then the proportion of cares that results is .25. Given the correlation between sample size needed for FERD to learn a function and the complexity of the function, we could conclude that we would only be able to learn relatively simple functions with these constraints.

We now deal with the case where the noise added to the original function is normally distributed. We will assume that the noise added to both 0 and 1 have the same standard deviation (in effect, we are added the same noise to all values) and that the result of adding

the noise to the 0–values is $N(\mu_0, \sigma)$. The result of adding the noise to the 1–values in $N(\mu_1, \sigma)$.

In the uniformally-distributed noise calculations we knew that when $t < t_0$ then the original function value must have been 0. Since the normal distribution is infinite, we can not be certain of such conclusions. We must now set an error level, α, that we can live with. This error level is the probability of misassigning the function value. Thus,

$$\begin{aligned}\alpha &= P(f(x) = 0)P(t_x > t_1) + P(f(x) = 1)P(t_x < t_0) \\ &= p \cdot P(z > \frac{t_1 - \mu_0}{\sigma}) + (1 - p) \cdot P(z < \frac{t_0 - \mu_1}{\sigma}) .\end{aligned}$$

By symmetry, $\mu_0 - t_0 = t_1 - \mu_1$, so that $t_0 - \mu_1 = -(t_1 - \mu_0)$. Therefore,

$$\alpha = p \cdot P(z > \frac{t_1 - \mu_0}{\sigma}) + (1 - p) \cdot P(z < \frac{-(t_1 - \mu_0}{\sigma}) .$$

Because a normal distribution is symmetric about its mean,

$$\begin{aligned}\alpha &= p \cdot P(z > \frac{t_1 - \mu_0}{\sigma}) + (1 - p) \cdot P(z > \frac{t_1 - \mu_0}{\sigma}) \\ &= P(z > \frac{t_1 - \mu_0}{\sigma}) .\end{aligned}$$

If we let $z_\alpha = \frac{t_1 - \mu_0}{\sigma}$, then $t_0 = \mu_1 - z_\alpha \sigma$ and $t_1 = \mu_0 + z_\alpha \sigma$. Also,

$$\begin{aligned}c &= P(f(x) = 0 \text{ and } t_x < t_0) + P(f(x) = 1 \text{ and } t_x > t_1) \\ &= p \cdot P(z < \frac{t_0 - \mu_0}{\sigma}) + (1 - p) \cdot P(z > \frac{t_1 - \mu_1}{\sigma}) \\ &= p \cdot P(z < \frac{t_0 - \mu_0}{\sigma}) + (1 - p) \cdot P(z > -\left(\frac{t_0 - \mu_0}{\sigma}\right) \\ &= P(z < \frac{t_0 - \mu_0}{\sigma}) \\ &= P(z < \frac{\mu_1 - z_\alpha \sigma - \mu_0}{\sigma}) \\ &= P(z < \frac{\mu_1 - \mu_0}{\sigma} - z_\alpha) .\end{aligned}$$

For example, suppose that we wanted $P(\text{error}) = .005$, $\mu_0 = 0, \mu_1 = 1$, and $\sigma = .4$. Then $z_{.005} = 2.575$ and $c = .4701$. In this case, we would have to sample a good deal of the function in order for FERD to have a reasonable chance at learning the function. Note

that for fixed μ_0 and μ_1, increasing σ will result in a decrease in c. For fixed σ, increasing the difference in the two means will result in an increase in c.

III. An Experiment in People vs. Machine

Near the conclusion of the summer, Dr. Ross had developed a system, FLASH (Function Learning and Synthesis Hotbed), that improved on the previous function-learning system, AFD (ADA Function Decomposition). We compare the two systems to "people" learning. This experiment used the Minnesota cards [2] to investigate the relationship between people's performance, DFC, AFD's performance, and FLASH's performance. We first describe the Minnesota cards.

Eighty-one cards are involved. Each card has one, two, or three borders; one type of shape (circle, square, or plus); and one, two, or three instances of this shape. Also, the shapes are all white, grey, or black. The tester creates a concept unknown to the person taking the test. Then, the cards are shown to the person one at a time and the person is told whether the card fits the concept or not. Previously shown cards are left face up in the "fits the concept" or "doesn't fit" pile. The person guesses or learns the concept based on this information.

Melvin Richardson created ten concepts and gave the test with all these concepts to fifteen people, individually. Melvin recorded the number of cards it took each person to learn each concept. The average number of cards required was determined for each concept.

Using a code designed by Dr. Tim Ross, Melvin coded the cards and determined the DFC of each concept. Also with the code, Melvin allowed AFD and FLASH to learn the concepts with the following procedure. Ten random samples of cards were taken for each concept with at least one of the cards fitting the concept. Each machine-learning program was allowed to discover the concept based on these samples. By adding randomly-selected cards when the program did not learn the concept and removing randomly- selected cards when the program learned the concept, the exact number of cards required for each program to learn each concept based on those initial samples was obtained. The average of these learning thresholds for each concept was found for both AFD and FLASH.

The table on the next page gives the DFC for each concept and the averages for each concept for each type of learner (people, AFD, and FLASH).

CONCEPT	DFC	People	AFD	FLASH
1	8	10.27	16.2	16.2
2	4	27.87	18.5	10.6
3	0	4.00	9.5	2.9
4	12	9.07	22.3	18.9
5	8	10.40	17.1	15.2
6	8	28.87	8.4	12.6
7	2	11.53	10.4	3.8
8	12	16.8	35.9	19.3
9	4	24.67	11.2	9.5
10	4	6.36	9.4	13.5

We are interested in the relationship between these measures and whether AFD or FLASH learned better than people learned. For the latter question, the overall people, AFD, and FLASH averages are 15.1, 15.9, and 12.3, respectively.

For the relationship question, the appropriate correlation coefficients from the above table were found as were the appropriate Hotelling-Pabst test statistics. They are given below.

	Correlation Coefficient	Hotelling-Pabst Statistic
(AFD , DFC)	.73	81
(AFD , people)	.03	152
(FLASH , DFC)	.93	17
(FLASH , people)	.04	172
(people , DFC)	.11	137

The Hotelling-Pabst statistic is used to test the hypothesis that two quantities are posi-

tively related [1]. A low number for this statistic indicates a positive correlation. The test concludes that the FLASH and DFC averages are positively correlated at the .001 level of significance and that the AFD averages and DFC are positively related at the .10 level of significance. The other tests are inconclusive and the correlation coefficients don't indicate a strong positive linear relation between the variables.

Perhaps one reason for this inconclusiveness is that the concepts involved were necessarily very easy to learn — that is, they had a low DFC. Therefore, they do not represent a spectrum of all such functions. They are clustered in the "very easy" range. Hence, what may look scattered under the microscope necessary to discern the points in this narrow range may in actuality be more regular when we see the entire picture.

CONCLUSION

FERD continues to do superbly in its function-learning and continues to amaze people by its performance. It seeks no prior knowledge of the type of function — only a portion of the function itself — when it learns. This distinguishes it from every other machine-learning technique that we know of. We hope that some of the results presented here will lead to a greater understanding of FERD and to its ability to handle both "larger" functions and the situation where noise is added to a function.

SOURCES

1. Bradley, James V. *Probability; Decisions; Statistics.* Prentice- Hall, Englewood Cliffs, N.J., 1976.

2. Bruner, Jerome S., Jacqueline Goodnow, and George Austin. *A Study of Thinking.* John Wiley and Sons, Inc., New York, 1956.

3. Chartrand, Gary, and Linda Lesniak. *Graphs and Digraphs.* Wadsworth and Brooks, Monterey, CA., 1986.

4. Ross, Tim, Mike Noviskey, Tim Taylor, and Dave Gadd. "Pattern Theory: An Engineering Paradigm for Algorithm Design". *Tech Report WL-TR-91-1060.* July 26, 1991.

5. Wolper, Jim. "Aspects of Pattern Theory". *AFOSR-SFRP Tech Report.* 1991.

EFFECT OF ANTIOXIDANTS ON THERMAL DECOMPOSITION OF ENERGETIC MATERIALS

Theodore J. Burkey
Associate Professor
Department of Chemistry

Memphis State University
Memphis, TN 38154

Final Report for:
Summer Research Program
Wright Laboratories/Armament Directorate
Munitions Division/Energetic Materials Branch
HERD Facility

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D. C.

August 31, 1992

EFFECT OF ANTIOXIDANTS ON THERMAL DECOMPOSITION OF ENERGETIC MATERIALS

Theodore J. Burkey
Associate Professor
Department of Chemistry
Memphis State University

Abstract

The effect of free-radical inhibitors (antioxidants) on the thermal decomposition of NTO, RDX and TNAZ was studied. In addition the possible role of redox reactions in the thermal decomposition of NTO was investigated. The inhibitor was doped in the energetic material at about 1% by weight. The heat from isothermally heated samples was detected by differential scanning calorimetry. In general, the inhibitor accelerated or increased the heat evolution of the energetic compounds. The effect was more pronounced for inhibitors that were better reducing agents. While these studies are still incomplete the results suggest that electron or hydrogen atom transfer processes are involved in thermal decomposition of some nitro group compounds. Such reactions can ultimately lead to the oxidation of NTO. In related studies, the gas products of NTO oxidation were studied. Oxidation of NTO with $KMnO_4$ results in the decomposition of NTO and formation of carbon dioxide. Oxidation of NTO with ammonium ceric nitrate results in carbon dioxide and nitrogen gas evolution. These results are consistent with a mechanism of NTO decomposition that involves the formation of didehydro-NTO. This compounds is expected to be unstable and undergo a spontaneous electrocyclic ring opening yielding small gaseous molecules. Taken together these studies show that thermal decomposition and oxidative decomposition yield similar results and may have common intermediates and mechanisms.

EFFECT OF ANTIOXIDANTS ON THERMAL DECOMPOSITION OF ENERGETIC MATERIALS

Theodore J. Burkey

Introduction

Essentially two approaches can be used to prepare insensitive explosives. First, the structure of the explosive can be made more stable (*.e.g.* stronger bonds or less strain) so that unimolecular fragmentation reactions are kinetically unfavorable. Second, other compounds may be added to stabilize the explosive compounds. Which method is effective depends on the relative importance of unimolecular and bimolecular mechanisms for decomposition. The second method is the focus of this investigation.

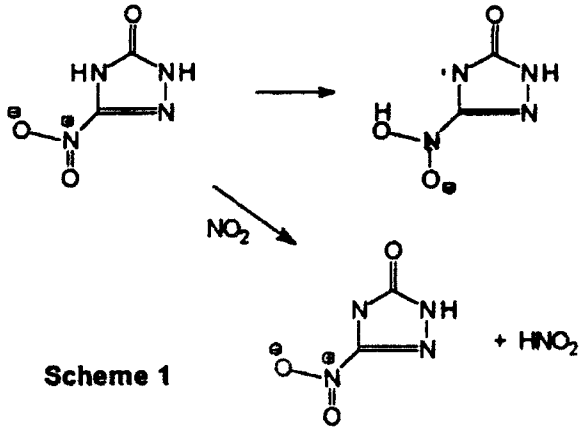
A stabilizer could work by forming a stable complex with the explosive or by intercepting decomposition fragments and keeping them from decomposing further or reacting with more explosive and other fragments. Note that a stabilizer may not work if the primary means of decomposition is strictly unimolecular. Since thermal decomposition often involves radical fragments (eq 1) it may

$$R\text{-}G \longrightarrow R\bullet + G\bullet \qquad (1)$$

$$R\bullet + InH \longrightarrow RH + In\bullet \qquad (2)$$

be possible to stabilize explosives by doping them with free-radical inhibitors to stop further fragmentation or radical chain reactions (eq 2). Inhibition of thermal decomposition by radical inhibitors or the lack of inhibition thus indicates what roles radicals may play in the mechanism of decomposition. Furthermore, a correlation of inhibitor effectiveness with inhibitor structure will provide revealing information about the radical reactions.

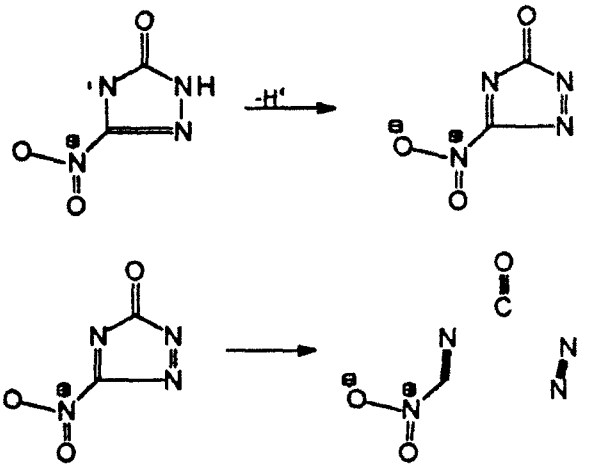
In this investigation we have doped NTO (3-nitro-1,2,4-triazol-5-one), RDX (hexahydro-1,3,5-trinitro-*s*-triazine), and TNAZ (1,3,3-trinitroazetidine) with various free-radical inhibitors to determine if the doped explosives were more stable than the pure explosives. On the contrary, we found that many inhibitors actually sensitized thermal decomposition. The relative effect of the different inhibitors suggests electron transfer may be a key step in the sensitized and undoped thermal decomposition.



Scheme 1

Another part of this investigation involved the oxidation of NTO with chemical oxidants. The mechanism of NTO decomposition is not well

understood. While the C-NO_2 bond is expected to be the weakest in the molecule and the first to be broken in thermal processes, recent isotope effects observed at the Seiler Laboratories indicate the rate determining step for thermal decomposition involves N-H bond cleavage.[1] Isotope effects for thermal decomposition have been observed to correlate with isotope effects of more drastic thermal processes.[2] So we believe that N-H bond cleavage may be important in NTO thermal decomposition and related energetic processes (Scheme 1). Previously, we proposed that complete fragmentation of the NTO ring to gaseous products may follow cleavage of both N-H bonds (Scheme 2). The loss of two hydrogen atoms from NTO constitutes an oxidation of NTO, and we reasoned that an oxidation may initiate the decomposition of NTO at low temperatures. We now have evidence to support this hypothesis. The investigation of the oxidation induced decomposition of NTO may provide information about the mechanism of NTO thermal decomposition if the two processes can be shown to have a common intermediate.



Scheme 2

Methodology

Materials: All materials were reagent grade unless otherwise specified. Ammonium ceric nitrate was Baker Analyzed Reagent (MW = 548.32), and NTO (Olin lot# 002, MW = 130) was recrystallized twice from boiling water (2.2 g/ 30 mL). Inhibitors were obtained from Sigma and Aldrich. $KMnO_4$ was obtained from HERD and is of unknown quality. RDX and TNAZ were obtained from HERD and used without further purification.

Instrumentation: The differential scanning calorimeter was a Perkin Elmer DSC-4. Thermograms were analyzed by Perkin Elmer DSCI and TADS software. Gas chromatographs were obtained with an 8 ft x 1/8 in Poropak Q column and a Varian 3400 chromatograph (conditions: -98° C for 5 min, increase 8° C/min to 200° C).

Oxidation of NTO: In a typical experiment ammonium ceric nitrate (0.471g, 0.858 mM) in 1 mL M H_2SO_4 was placed in a glass insert that fits into a Chemical Reactivity Test (CRT) reactor (a stainless-steel vessel and a cap having two high-vacuum valves). Dissolved gases were removed by three freeze-thaw-evacuate cycles: the solution was frozen in the reactor (dry-ice acetone for 9 min), the gas above the solid was evacuated, both valves were closed, the solid was thawed then frozen, and the remaining

gas was evacuated. A transfer tube (glass, both ends open) containing NTO (11.23 mg, 86.4 μmol) was placed in the insert next to the frozen, degassed ammonium ceric nitrate solution. The insert was returned to the reactor which was subsequently evacuated to about 1 μ Hg pressure (92VI19-4, experiment reference numbers in this report refers to year(92), month(VI), day(19) and experiment #(4)). After closing both valves, the reactor tube was placed in boiling water for 10 min, and the sample was refrozen. The gases over the frozen sample were analyzed by sweeping them with helium into the GC. This procedure was repeated to determine if the reaction was complete. The procedure with $KMnO_4$ was the same except that the reactor tube was not boiled only thawed to room temperature and allowed to react for 9 min. A typical procedure involved 31.07 mg NTO (0.239 mmol) reacting with 30.67 mg $KMnO_4$ (0.194 mmol) in 1 mL of 0.2 M $KMnO_4$ (0.2 mmol) and 1 M H_2SO_4.

Oxidation of oxalic acid: In a typical reaction oxalic acid monohydrate (17.84 mg, 97 μmol) was placed in a transfer tube containing 1 mL of ammonium ceric nitrate (0.2423 g, 442 μmol, in 1 M H_2SO_4, 92VI16-5). The sample was treated as NTO above with 30 min sonication instead of boiling.

Differential scanning calorimetry: Isothermal programs were used to obtain thermograms. Load temperature was 50 ^oC, and samples were heated at 80 ^oC/min. Faster heating rates resulted in unreproducible isothermal temperatures. A control sample of explosive was prepared the same as the doped explosive. Sample weights were kept within 1% of that of controls.

Experiments using aluminum capsules (TA Instruments lid #073191 900794.901 and pan #073191 900793.901) generally involved dissolving 0.6 mg dopant and 60 mg NTO in a minimum amount of acetone (4-5 mL). The acetone was evaporated under a stream of N_2. The dry residue was transferred to a capsule, and the amount was determined gravimetrically. Capsules were crimped containing 0.5 mg of explosive.

Experiments in stainless steel capsules (Perkin Elmer lid #319-1526, pan #319-1525, and O-ring #319-1535) typically were run using a stock solution of 30 mg of NTO and 0.3 mg of dopant dissolved in 2 mL acetone. Using a microliter syringe the NTO/dopant solution was added slowly to a hot (60-65 ^oC) capsule to avoid solvent flash and sample loss. The amount of sample transferred was determined gravimetrically. Stainless steel capsule/cover/o-ring assemblies were crimped. A different procedure was used for Buckminsterfullerene. Buckminsterfullerene (0.3 mg) was sonicated in 2 mL toluene providing a blue/brown solution. The Buckminsterfullerene solution was mixed with solution of 30 mg explosive in 2 mL acetone. One mL more acetone was added since the solution became turbid after a few minutes. Only stainless steel capsules were used with RDX and TNAZ.

Results

Oxidation of NTO: Oxidation with $KMnO_4$ yielded 2.6 mol CO_2 and 0.38 mol CO per mol NTO. No N_2 was detected. Since this amount of CO_2 and CO exceeds the theoretical limit of NTO carbon, experiments with oxalic acid were undertaken to calibrate the method. Unfortunately oxalic acid reacted slowly with $KMnO_4$. Complications occurred with MnO_2 ppt (although boiling $KMnO_4$/oxalic acid solutions may have overcome them), so ammonium ceric nitrate was used. Results with ammonium ceric nitrate were also not quantitative (1.26 mol CO_2/mol $C_2O_2H_2$, 92VI16-5). This was apparently due to inadequate reaction time. Ceric ammonium nitrate oxidation of NTO yielded 0.8 mol CO_2, 0.46 mol N_2, and 0.03 mL NO per mol NTO (92VI17-6). In a second run the yields were 0.70 mol CO_2, 0.69 mol N_2, and 0.04 mol NO per mol NTO (92VI19-4). The nitrogen did not come from the atmosphere since oxygen would have been present and none was detected. Inspection of the insert contents after boiling the sample twice and analyzing the gases each time revealed a precipitate. This result suggests the reaction was incomplete.

Thermal Decomposition of NTO: Small samples (0.5 mg) were used in aluminum capsules to avoid rupture. The most useful temperature was 260 °C. Lower temperatures gave smaller exotherms that were not reproducible, and at higher temperatures, heat evolved so rapidly that thermograms of samples could not be distinguished from that of controls. At 260 °C the exotherm starts as soon as the isothermal temperature is reached (see Figure 1). The magnitude of the leading portion of the exotherm appears to depend on sample size; a result that suggests the leading portion is due to surface catalyzed processes. The results for experiments in aluminum capsules are summarized on Table 1.

Table 1. Effect of antioxidants on thermal decomposition of NTO

antioxidant	onset	max	n	aging	nb ref #
none	2.42	2.74	1	same day	92V15-4
none	2.6(6)	3.2(3)	2	3 days	92V18-3,4
none	2.9(2)	3.2(2)	2	5 days	92V20-12,13
Trolox	1.4(2)	1.8(1)	2	1 day	92V15-6,7
Trolox	1.5(3)	1.90(2)	2	4 days	92V18-2,3
picric acid	2.2(2)	2.6(2)	2	same day	92V18-7,8
dithiothreitol	1.1(3)	1.3(2)	3	1 day	92V19-3,4,5
cysteine	1.7(2)	2.2(3)	3	same day	92V19-7,8,9
phenol	2.97	3.25	1	same day	92V19-11
phenol	2.81(5)	3.5(5)	2	1 day	92V20-6,7
hydroquinone	2.2(4)	2.5(4)	2	same day	92V20-8,9
hydroquinone	2.2	3	1	1 day	92V21-2
octanethiol	2.4(6)	3.2(7)	2	same day	92V20-10,11
octanethiol	2.5	2.8	1	1 day	92V21-4

parenthetical values are the errors in the last digit(s)

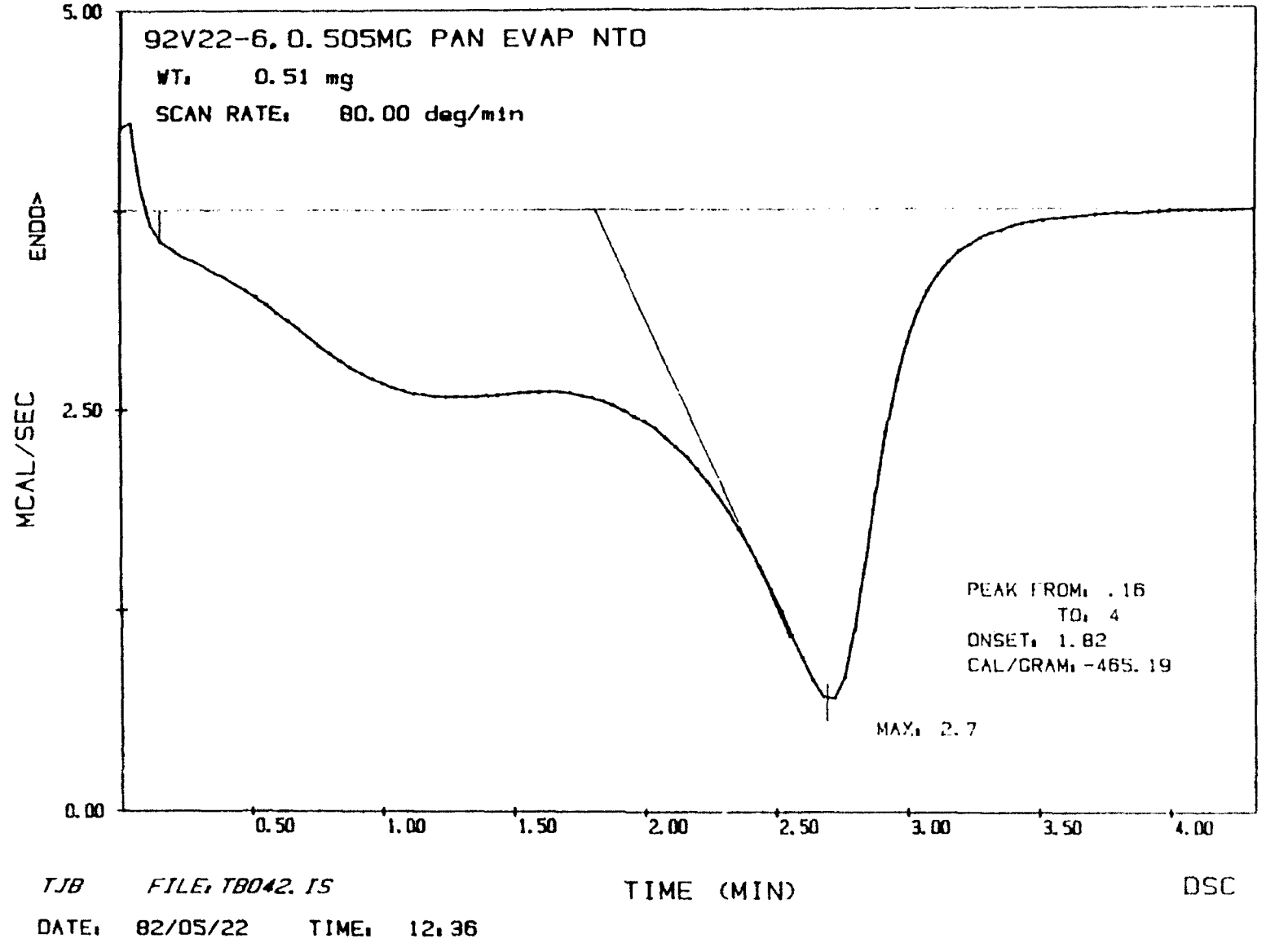

92V22-6. 0. 505MG PAN EVAP NTO
WT: 0.51 mg
SCAN RATE: 80.00 deg/min
5.00
2.50
0.00
ENDO>
MCAL/SEC
PEAK FROM: .16
TO: 4
ONSET: 1.82
CAL/GRAM: -465.19
MAX: 2.7
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
TJB
FILE: TB042. IS
TIME (MIN)
DSC
DATE: 82/05/22
TIME: 12:36

Figure 1

The time to the onset and the maximum of the main exotherm is lowered when NTO is doped with cysteine, Trolox, and dithiothreitol (DTT) (order of increasing effect). Within experimental error no effect is observed for phenol, hydroquinone, and octanethiol. The melting points of cysteine, Trolox, and DTT are 220, 189, and 43 °C, respectively.

Stainless steel capsules were used with NTO to reduce surface catalysis. The thermogram is well behaved with little or no heat evolution when the isothermal temperature is reached (Figure 2). The heat release appears to be faster than the response with the stainless steel capsule, *i.e.* the greater thermal capacity and lower thermal conductivity of the stainless steel capsule versus the aluminum one reduces the rate at which heat transfers into the head. The sample weight and isothermal temperature have not been fully optimized.

Doped and undoped NTO samples at 239 °C evolved small amounts of heat unreproducibly (40-150 cal/g). Capsules containing DDT- and Trolox-doped samples often burst open (presumable due to rapid gas and heat evolution) thus heat evolution could not be quantified. No exotherm or isotherm was observed for neat DTT or Trolox. The onset and the maximum of the NTO exotherm were reproducible and within experimental error occurred earlier when the NTO was doped with DTT, Trolox, and Alcoa 1401 aluminum (see Table 2). For NTO doped with sodium nitrate an unreproducible exotherm was observed as soon as the isothermal temperature was reached. Interestingly the capsules of the sodium nitrate doped NTO or sodium salt of NTO did not burst, and the total heat released was similar to that of undoped NTO.

The effect of Buckminsterfullerene (C_{60}) on NTO decomposition was not reproducible, possibly due to a solvent impurity (92VII8-1, 92VII8-3, 92VII9-8). On the other hand, addition of 1% of 2,3,5,6-tetrachloro-1,4-benzoquinone (TCBQ) to NTO in two samples resulted in a nearly constant evolution of heat for 7 min from the time the isothermal temperature was reached. In addition, 3-4 times more heat evolved from the TCBQ-doped NTO samples than from an NTO control sample (92VII9-4,6, results not shown).

Thermal Decomposition of RDX: Heating RDX (2.6 mg) at 210 °C gives an exotherm that can be detected as soon as the isothermal temperature is reached. This broad exotherm (2 min) course at about 2 minutes. A second sharp exotherm occurs about 30 minutes, and the capsule usually burst open during the second exotherm. Addition of DTT has a pronounced effect on the onset and maximum times for the second exotherm. For 1 mg samples at 220 °C, the first exotherm of doped and undoped RDX show no discernible difference (onset, maximum, or total heat released).

Figure 2

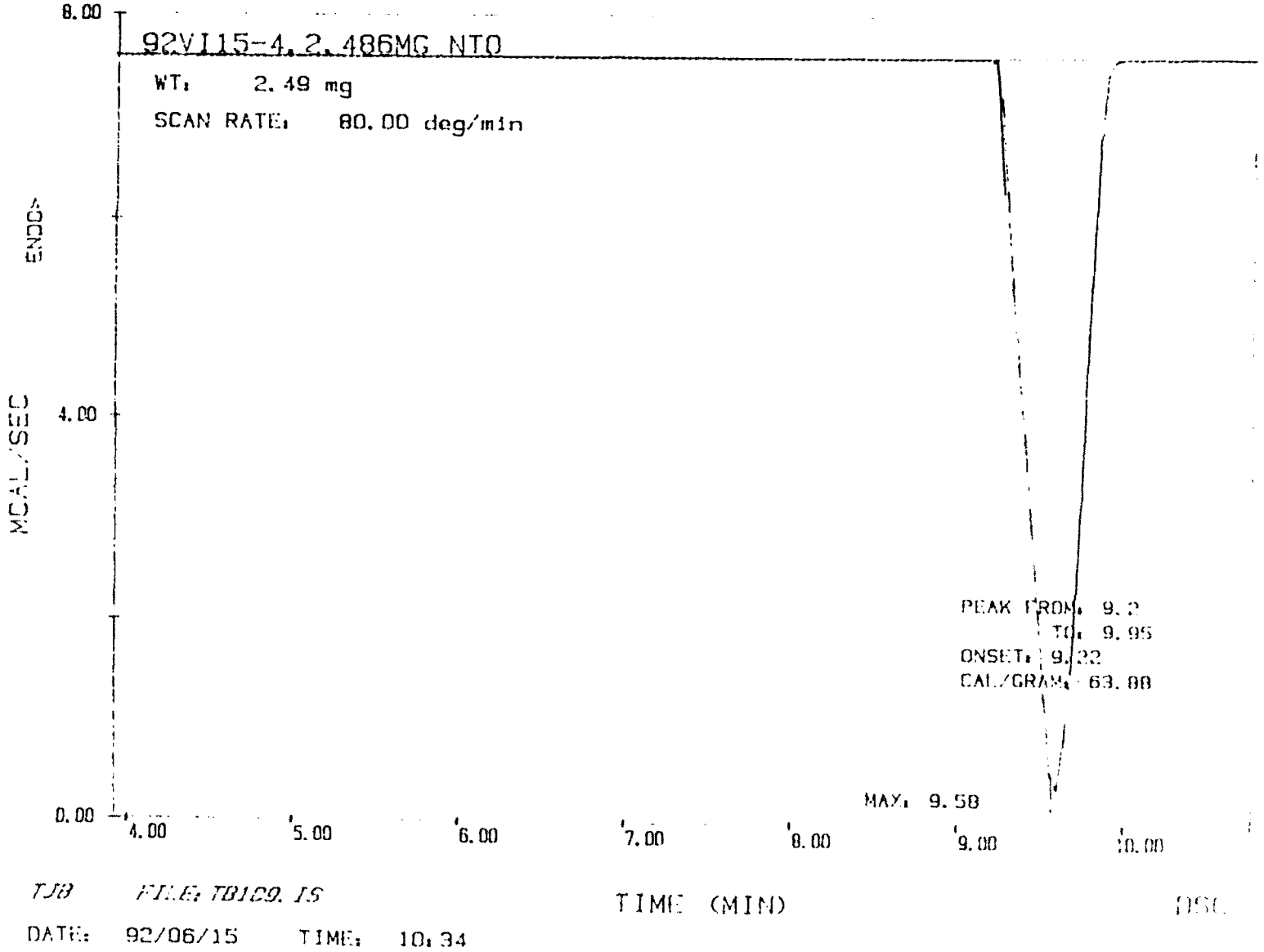

8.00
92V115-4.2.486MG NTO
WT: 2.49 mg
SCAN RATE: 80.00 deg/min
ENDO
MCAL/SEC
4.00
0.00
PEAK FROM: 9.2
TO: 9.95
ONSET: 9.22
CAL/GRAM: 63.88
MAX: 9.58
4.00
5.00
6.00
7.00
8.00
9.00
10.00
TJB FILE: TB1C9.IS
TIME (MIN)
DSC
DATE: 92/06/15 TIME: 10:34

Table 2. Effect of 1% Antioxidant on the Thermal Decomposition of NTO in Stainless Steel Capsules at 239 °C.

sample	onset	max	n	nb ref #
NTO	9.3(2)[a]	9.29(25)	4	92VI15-4,7,8,9
Trolox/NTO	7.0(4)	7.1(4)	3	92VI16-1,3,4
Trolox	none		1	92VI16-6
dithiothreitol/NTO	4.67(7)	4.72(5)	3	92VI16-11,12
dithiothreitol	none		1	92VI16-9
Alcoa 1401/NTO[b]	8.26(26)	8.33(30)	4	92VI16-14,VII17-1,4,5
$NaNO_3$/NTO	0		3	92VII18-3,4,5
NaOH/NTO[c]	0.87	0.97	1	92VII18-6

[a]parenthetical values are the errors in the last digit(s) [b]10% aluminum by weight [c]1:1 mol ratio

Table 3. Effect of DTT on the Thermal Decomposition of RDX at 210 °C (second exotherm)

sample	onset	max	n	nb ref #
RDX	31	31	2	92VII1-11 92VII2-1
DTT/RDX	12.8(6)	12.9(6)	4	92VII1-6,7,8,10

Thermal Decomposition of TNAZ: A 0.873 mg sample of TNAZ was heated 10 °C/min from 50 to 300 °C and lost 50% of its weight (92VII6-14). Heat evolution occurs as low as 190 °C. An exotherm of 2.8 kcal/mol has an onset at 238 °C and a maximum at 260 °C (92VII6-14). Heating DTT doped TNAZ isothermally at 210 °C clearly increases the amount of heat evolved (Table 4). Evaporation of acetone required longer heating than with RDX or NTO. This is presumably due to the stronger solvation. Since DTT is low melting loss of DTT before DSC heating may contribute to the variability of the data. A preliminary experiment with Trolox suggests that it also promotes TNAZ decomposition.

Table 4. Effect of 1% Antioxidants on the Thermal Decomposition of TNAZ at 210 °C

sample	cal/g	% wt loss	n	nb ref #
TNAZ	34(53)	14(8)	3	92VII7-6,VII8-2,4
DTT/TNAZ	712(111)	22(7)	4	92VII7-5,VII8-8,10,VII9-1
	592(284)	21(6)	5	92VII7-5,VII8-5,8,10,VII9-1
Trolox/TNAZ	182		1	92VII8-9
	500	27	1	92VII9-5

Conclusions

Decomposition of NTO occurs only with strong oxidants. For example no reaction occurs with iodine,[3] On theother hand, oxidation with acidic ammonium ceric nitrate clearly causes produces carbon dioxide and nitrogen gas. The incomplete oxidation in this case may be due to precipitation of intermediates. Our preliminary experiments indicate that a mole of each gas is formed per mole of NTO oxidized. Oxidation with $KMnO_4$ produced greater than one mole of carbon dioxide per mole of NTO in agreement with results obtained with infrared spectroscopy by Perry Fan, a graduate student at our Memphis laboratory. Surprisingly no nitrogen was detected after $KMnO_4$ oxidation (note this experiment is yet to be repeated). We can only speculate that N_2 is not a free intermediate of oxidation. It may be that the greater oxidizing power of $KMnO_4$ results in soluble nitrogen products. Scheme 2 is consistent with these results although the chemistry is much more complex. Carbon dioxide can be produced by the oxidation of carbon monoxide formed by the fragmentation of NTO. In addition, CN from the $NCNO_2$ fragment can hydrolyze to carbonic acid which would decompose to carbon dioxide in the acidic so ..tion. This additional carbon dioxide formation appears to occur only during $KMnO_4$ oxidation where strongly coordinating manganese species are available for catalysis.

The purpose of the oxidation studies was to generate didehydro-NTO as a possible common intermediate with NTO thermal decomposition. The oxidation results in the decomposition of NTO to gases, and while we have not observed didehydro-NTO directly, our results are consistent with didehydro-NTO formation during the thermal decomposition of NTO where gases are known.

Independent of the capsule type or isothermal temperature, NTO decomposes sooner in the presence of DTT and other inhibitors. This effect is the opposite expected for the inhibition of a free radical reaction. Some inhibitors had no effect at all, *i.e.*, octanethiol, phenol, and C_{60}. This suggests that either the decomposition involves no radicals, the radicals are not free to migrate to the inhibitor, or the radicals do not react appreciably with the inhibitors. The second option seems the most plausible.

Inspection of the oxidation potentials of the inhibitors suggests that the decomposition is accelerated the most by highly reducing species. Thus the decomposition is probably initiated by electron or hydrogen atom transfer. The only inhibitor that cannot donate a hydrogen atom is C_{60}, and this shows no noticeable effect. Electron or hydrogen atom transfer would generate a radical. This is reasonable since our studies at Memphis State show that NTO is easily reduced. Furthermore, nitro compounds are well known to accept electrons and form radicals that undergo fragmentation.[4] It may be speculated that one pathway for the initiation of thermal decomposition of undoped NTO is an electron transfer between neighboring NTO. However, we found at Memphis State that NTO is highly resistant to

consistent with the fact that the N-H bond is relatively weak, and nitro groups are known to abstract hydrogen atoms.[1]

Although few studies with free-radical inhibitors were completed with RDX and TNAZ it is clear, even in these cases, that thermal decomposition is affected. The acceleration by inhibitors of heat loss appears to be a general effect for nitro group containing explosives. More studies need to be done to better understand this effect.

Finally it may be worth considering energetic compounds or formulations that rely on redox chemistry. At higher temperatures the barrier to electron transfer will be more easily overcome and mechanisms involving electron transfer can become dominant when atomic/molecular diffusion becomes rate limiting.

---

1. Menapace, J. A.; Marlin, J. E.; Bruss, D. R.; Dascher, R. V. *J. Phys. Chem.* **1991**, **95**, 5509.

2. Shackelford, S. A., "Condensed Phase Kinetic Deuterium Isotope Effects in High Energy Phenomena: Mechanistic Investigations and Relationships" Air Force Systems Command Technical Report FJSRL-TR-89-0010, December 1989.

3. Fan , P.; Burkey, T. J. unpublished results.

4. Lowry, T. H. and Richardson, K. S. "Mechanism and Theory in Organic Chemistry" third ed. 1987, 409 pp.

EVALUATION OF THE SBR AND GRE METHODS FOR COMPUTING THE TIME DOMAIN ELECTROMAGNETIC SCATTERING FROM LARGE OPEN-ENDED WAVEGUIDE CAVITIES

Robert J. Burkholder
Postdoctoral Research Associate
Department of Electrical Engineering

The Ohio State University
1320 Kinnear Road
Columbus, Ohio 43212

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Wright-Patterson Air Force Base, Dayton, Ohio

September 1992

EVALUATION OF THE SBR AND GRE METHODS FOR COMPUTING THE TIME DOMAIN ELECTROMAGNETIC SCATTERING FROM LARGE OPEN-ENDED WAVEGUIDE CAVITIES

Robert J. Burkholder
Postdoctoral Research Associate
Department of Electrical Engineering
The Ohio State University

Abstract

Two approximate but versatile high-frequency asymptotic ray shooting methods are evaluated in terms of their ability to accurately and efficiently predict the time domain electromagnetic scattering from large open-ended waveguide cavities: (1) the shooting and bouncing ray (SBR) method and (2) the generalized ray expansion (GRE) method. A hybrid modal solution for simple rectangular and circular open-ended waveguide cavities is used as a reference solution. It is found that the SBR method does not predict the time dispersive effects of cavities because it does not include the fields diffracted into the cavity by the rim at the open end. However, for very large cavities (on the order of 20 wavelengths wide or larger), the diffracted fields are much weaker so the dispersive effects are less significant and the accuracy of the SBR method in the time domain improves. The GRE method predicts the time dispersive effects because the diffracted fields are intrinsically included, but it may require a much larger number of rays to be tracked and is not as easily implemented as the SBR method. In general, the SBR method is more efficient than the GRE method unless more than approximately 100 incidence angles are to be computed for a given cavity geometry. Approximate expressions are included for estimating the number of rays required for a given cavity geometry for both methods.

EVALUATION OF THE SBR AND GRE METHODS FOR COMPUTING THE TIME DOMAIN ELECTROMAGNETIC SCATTERING FROM LARGE OPEN-ENDED WAVEGUIDE CAVITIES

Robert J. Burkholder

INTRODUCTION

The electromagnetic (EM) scattering from electrically large open cavities can be a very significant contributor to the overall radar cross section (RCS) of objects containing such cavities. In particular, the EM scattering from engine inlets and exhausts is often the most prominent feature contributing to the RCS of modern jet aircraft (Figure 1). The time domain EM scattering from aircraft, i.e., the reflected radar return from a target as a function of time, is often used for radar target identification. Therefore, it is important to understand and accurately predict the time domain behavior of EM fields scattered by large cavities, such as the jet inlet and exhaust. However, because of the size and complexity of these types of cavities, it is not currently possible to obtain an exact solution either in closed form or by using a numerical solution. Approximate high-frequency asymptotic ray methods for cavity RCS prediction solutions exist, and it is of interest here to apply two of these methods and evaluate them in terms of their ability to accurately and efficiently predict the time domain response. This is achieved by analyzing two classes of canonical open-ended waveguide cavity geometries for which there exists a waveguide modal reference solution. The canonical geometries used here are perfectly conducting open-ended rectangular and circular waveguides terminated by a flat plate (short circuit). Although these cavities have nowhere near the complexity of an actual engine cavity as in Figure 1, they are useful for evaluating the ability of the ray methods to predict the important time domain cavity scattering effects.

METHODOLOGY

Two versatile ray-based high-frequency asymptotic methods have been developed for calculating the RCS of electrically large arbitrarily shaped cavities: the shooting and bouncing ray (SBR) method [2, 3, 4], and the generalized ray expansion (GRE) method [3, 4, 5].

In the SBR method, a dense grid of parallel ray-tubes is launched into the cavity, repre-

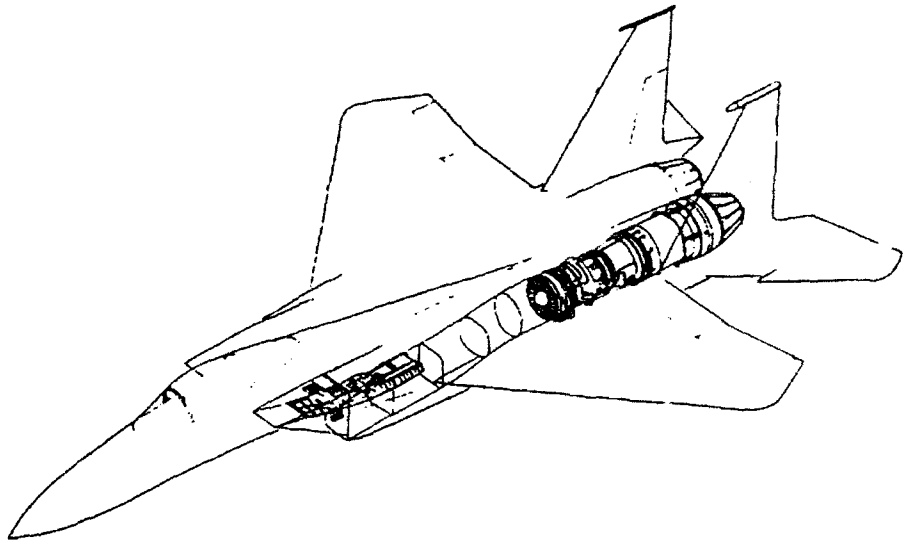


Figure 1: F-15 jet inlet duct and engine compartment [1]

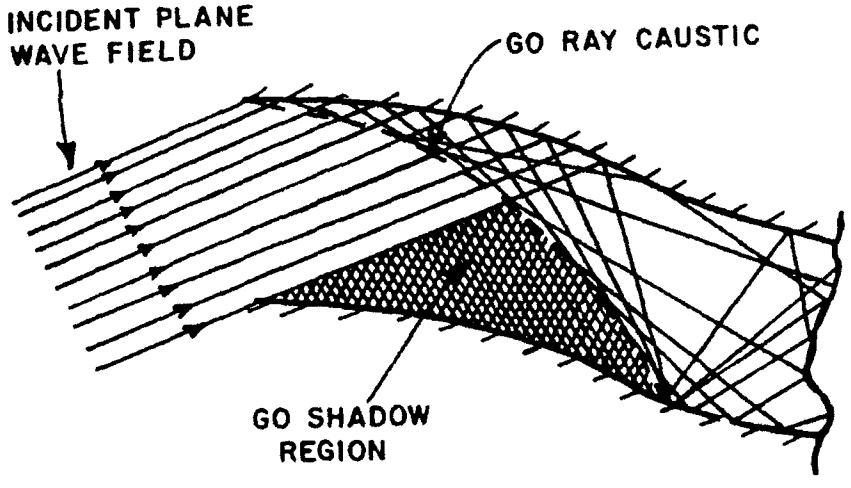


Figure 2: Ray launching in the shooting and bouncing ray (SBR) method, shown in 2-D.

senting the geometrical optics (GO) portion of the incident (illuminating) plane wave which is intercepted by the open end (Figure 2). The individual ray-tubes are tracked inside the cavity via multiple reflections (or "bounces") from the inner cavity walls using the laws of GO [6], until they exit again through the open end (Figure 3). The scattered field is obtained by performing an aperture integration over the equivalent Huygen's sources [7] defined by the projections (or "footprints") of the ray-tubes in the open end as they exit the cavity [3, 8]. The SBR method can more specifically be referred to as the geometrical optics/aperture integration (GO/AI) method [3, 9].

The SBR method is very versatile and easy to apply, but it tracks only the GO field inside the cavity, ignoring the fields diffracted into the cavity by the rim at the open end. As Figure 2 shows, this leaves GO shadow regions which in reality contain the diffracted field. The GRE method intrinsically includes these diffracted fields by launching rays in all directions into the cavity. In the GRE method, the aperture at the open end is broken up into a small number of subapertures, and ray-tubes are launched radially into the cavity from the phase centers of the subapertures (Figure 4). Each ray-tube is weighted by the far-field radiation pattern of its subaperture when the subaperture is excited by an external field incident on the open end of the cavity. The subaperture far-field radiation pattern is obtained by integrating the equivalent currents defined by the incident field over the subaperture. Once launched in this manner, the ray-tubes are treated in exactly the same manner as the SBR ray-tubes, i.e., they are tracked inside the cavity via multiple GO bounces and aperture integration is used to obtain the scattered field. This method can more specifically be referred to as the GRE/AI method.

In the SBR method, a new grid of ray-tubes must be tracked for each plane wave incidence angle. An advantage of the GRE method is that only the amplitude weighting of the ray-tubes is dependent on the incident field, and not the initial launch directions. Therefore, a sufficiently large number of ray-tubes can be launched, tracked inside the cavity, and stored; later the rays can be weighted and summed to give the cavity scattered field for any number of incidence angles. However, a much larger number of GRE ray-tubes is generally required

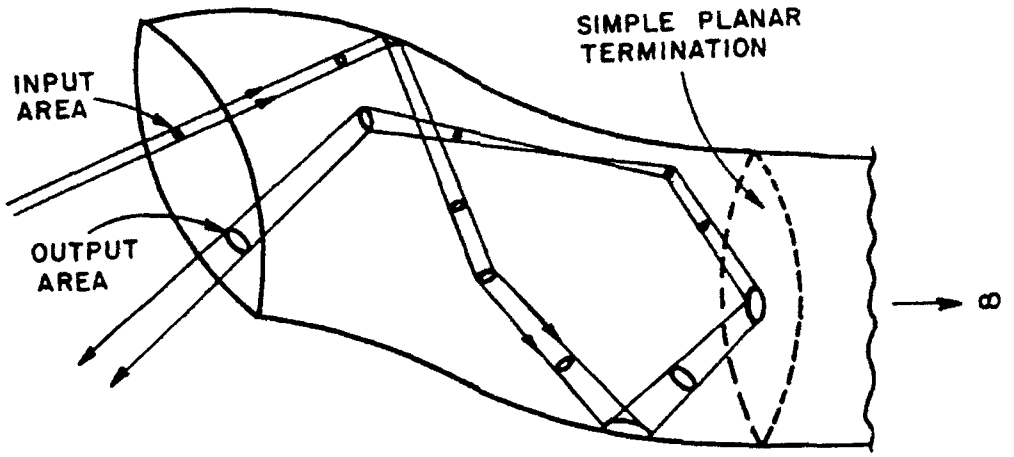


Figure 3: GO ray-tube tracked inside a waveguide cavity until it exits through the open end.

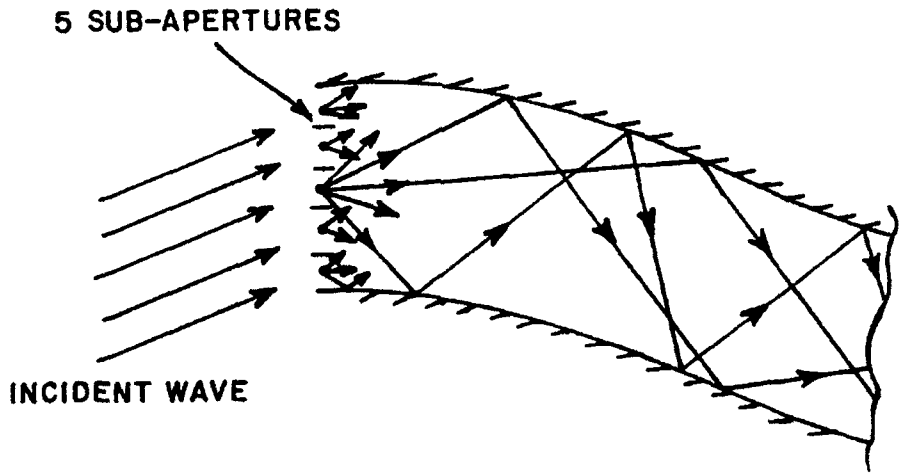


Figure 4: Ray launching in the generalized ray expansion (GRE) method, shown in 2-D.

compared with the number of ray-tubes used in SBR for a single incidence angle. Therefore, SBR is more efficient than GRE unless a sufficiently large number of incidence angles is necessary (typically, on the order of 100 or more incidence angles are necessary before GRE surpasses SBR in terms of ray efficiency for a given cavity). On the other hand, GRE is expected to be more accurate because it includes diffraction effects.

An estimate of the total number of ray-tubes in the GRE/AI method required to cover all incidence angles within 60° of the waveguide axis, including both elevation and azimuth angles, is very roughly given by:

$$\begin{aligned} \text{No. of rays} &\approx 56 \times \left(\frac{V}{\lambda_{min}^3}\right) \qquad (1) \\ V &= \text{Volume of cavity,} \\ \lambda_{min} &= \text{smallest wavelength of interest.} \end{aligned}$$

An estimate of the number of ray-tubes in the SBR (GO/AI) method required for *a single incidence angle* is very roughly given by:

$$\begin{aligned} \text{No. of rays} &\approx 25 \times \left(\frac{A}{\lambda_{min}^2}\right) \qquad (2) \\ A &= \text{Area of aperture (open end).} \end{aligned}$$

Covering all incidence angles within 60° of the waveguide axis, with 1° spacing, would take approximately 10,000 incidence angles. So, 10,000 times the result in (2) is the total estimated number of rays required by SBR to cover the same angular sector (at 1° increments) as the GRE rays given by (1).

These two ray methods are both very versatile in that they can handle relatively arbitrarily shaped metallic cavities which may have material treatments. But they are approximate, so to determine their relative accuracies the hybrid asymptotic modal method [3, 9] is used here as a reference solution. This method is highly accurate and has been validated with the method of moments and measurements. However, because it uses natural waveguide eigenmodes, it is restricted to cavities which can be made up of sections of uniform waveguides for which the natural modes are easily found. In the numerical results, straight open-ended waveguide cavities with rectangular or circular cross-sections will be analyzed.

Figure 5 shows a graphical illustration of the power flow inside an open-ended, semi-infinite, parallel plate waveguide illuminated by an incident plane wave [10]. This figure clearly shows the difference between the three methods, and the difference between using one and five subapertures in the GRE method. The SBR, or GO, result shows very discrete regions of shadow inside the waveguide, while the reference modal solution is much more continuous. However, this is a rather small waveguide (only 3 wavelengths across); as the width of the waveguide increases, the modal result will look more and more like the SBR result because the fields diffracted into the waveguide by the edges at the open end will become less significant. Of course, the GRE results reproduce the modal results much better for this small waveguide because diffraction effects are intrinsically included. The difference between the two GRE results is that the 5 subaperture result begins to agree with the modal solution much closer to the open end than the 1 subaperture result. This property of the GRE method makes it necessary to use more subapertures for shallow cavities than for deep ones [3].

For more details on the implementation of the three methods discussed above, consult the published references.

NUMERICAL RESULTS

The open-ended waveguide cavities analyzed here are axially straight with square, rectangular or circular cross-sections and have a flat plate termination normal to the waveguide axis. They are perfectly conducting and only the scattering from the cavity interior is considered. All external scattering effects are ignored, such as the direct scattering from the rim at the open end. The cavities are illuminated by a vertically polarized EM plane wave incident in the horizontal plane.

In the GRE results presented here, the apertures are subdivided into a number equally sized rectangular subapertures which best approximate the original aperture. Therefore, the square and rectangular apertures are easily gridded into a relatively small number of subapertures, while the circular apertures require a much larger number of subapertures to adequately reproduce the circle. A more sophisticated subaperture gridding scheme could

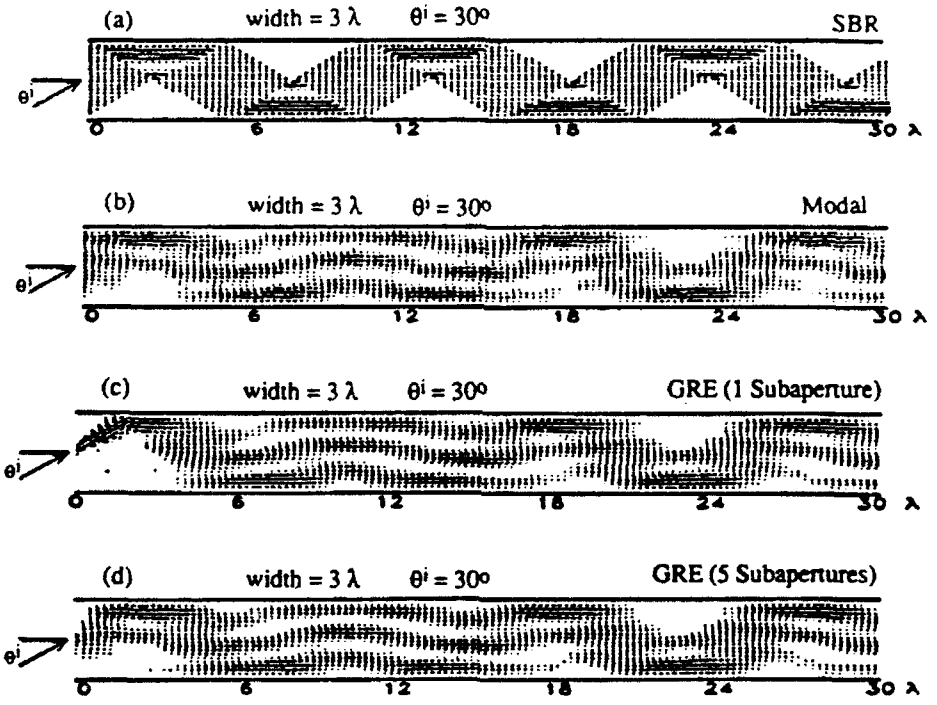


Figure 5: Power flow inside an open-ended, semi-infinite, parallel plate waveguide illuminated by an incident EM plane wave polarized perpendicular to the plane of the page.

have been used to reduce the number of subapertures for the circular waveguides, but it has been found that using fewer subapertures degrades the accuracy of the GRE method for cavities which have curved walls at the open end.

In practice, the rays of the GRE method are launched in a cone whose half-angle is at least as large as the maximum incidence angle measured with respect to the waveguide cavity axis, and then tracked and saved so that any incidence angle within this cone can be accommodated. However, since only one incidence angle at a time is of interest in the numerical results presented here, only ray-tubes lying within a certain cone *with respect to the incidence direction* are used in the computations. (It is assumed that ray-tubes launched outside of this cone will be weakly excited and not contribute significantly.) The cone half-angle depends on the electrical size of the cavity, and is indicated in the text for each case. Therefore, the number of rays indicated in the GRE results is not the total number of rays necessary for a wide range of incidence angles, but the number of rays used for that single incidence angle. An estimate of the total number of rays which would be necessary to accommodate all incidence angles out to 60° from the waveguide axis will also be presented. The number of rays indicated in the SBR results reflects the number of rays which are tracked for that single incidence angle. An estimate will also be given of the total number of rays required in the SBR method for all incidence angles in the same 60° angular sector, spaced at 1° increments. Note that this angular sector includes all angles, in both elevation and azimuth, which are within 60° of the waveguide axis.

Currently, all three methods used here to calculate the cavity scattered field have only been formulated in the frequency domain. Computer codes based on the three methods have been modified to generate RCS as a function of frequency in an efficient manner. The time domain scattering is obtained by inverse Fourier transforming the frequency domain RCS. The RCS vs. frequency is found in the 8 to 12 GHz band at 128 equally spaced steps. The frequency domain data is weighted with a Kaisser-Bessel window (alpha=2) before transforming, to eliminate unwanted sidelobe contamination of the time domain response. The result is a "band-limited" time domain impulse response. Such a time domain response

will exhibit a series of peaks, or "impulses" which are modulated by the midband (10 GHz) carrier frequency. The resolution of the impulses increases with the bandwidth, and the true impulse response is obtained as the bandwidth approaches infinity. For the relatively narrow bandwidth used here, the impulses may overlap and not be easily resolved for some cases.

Figure 6 shows the time domain impulse response of an open-ended waveguide cavity with a square cross-section, found using the three methods discussed here. At midband this cavity is 6.67 wavelengths across and 26.67 wavelengths long. The SBR method predicts well the main impulse but misses the side impulses which are "dispersed" in time. This "time dispersive" property is characteristic of cavities in general, and distinguishes cavities from simple scatterers such as corners and edges. The dispersion is due to the fields diffracted into the cavity by the edges at the open end which propagate inside the cavity at steeper or shallower angles (with respect to the waveguide axis) than the GO fields. Therefore, these diffracted fields undergo more or fewer bounces and traverse the length of the cavity later or earlier, respectively, than the GO fields. Notice that the GRE solution predicts the primary time dispersive effects because the diffracted fields are intrinsically included. However, the amplitudes of the peaks are not in perfect agreement because of the approximations associated with both the GRE ray launching scheme and ray shooting methods in general. (The pragmatic approach to ray tracking in arbitrarily shaped cavities is to "shoot" a given ray into the cavity, numerically search for the first intersection of the ray with the cavity walls, then compute the direction of the reflected ray and search for the next intersection, etc., until the ray exits the cavity. At each reflection, some error is introduced which degrades the accuracy of all subsequent reflections, so considerable error can accumulate if a ray undergoes many bounces inside the cavity.)

For the GRE result of Figure 6, only rays within 17.5° of the incidence direction are included in this computation, for reasons explained earlier in this section. It is estimated that a total of 116,000 rays would be required to accommodate all incidence angles within 60° of the waveguide axis. Using the SBR method for all incidence angles within this same angular sector, spaced at 1° increments, is it estimated that a total of 2,640,000 rays would

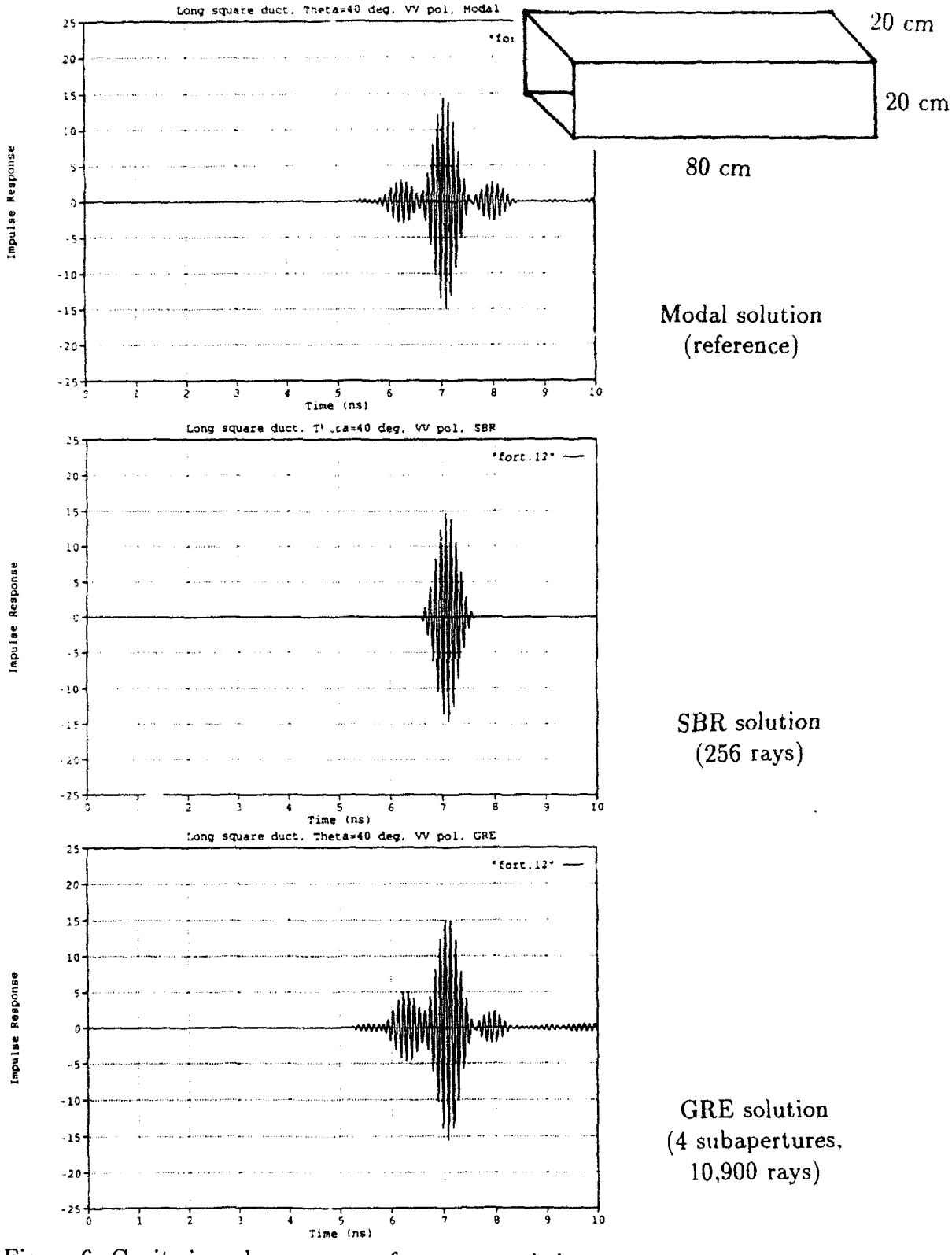


Figure 6: Cavity impulse response of an open-ended square waveguide cavity, illuminated 40° off-axis.

have to be tracked.

Figure 7 shows the impulse response for a circular waveguide cavity of the same dimensions as the rectangular cavity of Figure 6. Again, the dispersive effects of the cavity are clearly apparent in the modal and GRE results, while the SBR results shows a single non-dispersed impulse. In the GRE method, rays within 17.5° of the incidence direction are included in this computation. It is estimated that a total of 1,448,000 rays would be required to accommodate all incidence angles within 60° of the waveguide axis. This large number of rays is mainly due to the large number of subapertures required. Using the SBR method over this same angular sector (1° increments), is it estimated that a total of 15,264,000 rays would have to be tracked.

Notice that more rays are necessary to obtain the SBR result of Figure 7 than the previous case in Figure 6. This is because the GO rays diverge (de-focus) when they reflect from the curved cavity walls, making their cross-sectional areas enlarge. When a ray-tube exits the cavity it has an unknown cross-sectional shape because it has been distorted by reflections from the cavity walls. To perform the aperture integration of the ray-tube exactly, this shape needs to be known. However, if the cross-sectional *area* of the ray-tube is electrically small enough, the exact shape is not important, so any convenient shape can be assumed. This area can be easily found by conserving power flux in the ray-tube. Therefore, in practice it is only necessary to make the ray-tubes small enough so that they are less than approximately one square wavelength in cross-section when they exit the cavity. So, when the cavity walls cause divergence of the ray-tubes, more, smaller, ray-tubes are needed to maintain electrically small ray-tube cross-sections. This is also true of the GRE method.

In Figures 8 and 9, the time scale is shifted by 30 nanoseconds (ns) because the inverse Fourier transform routine which is used here has a periodic 30 ns window centered around zero. Therefore, impulses which occur later than 15 ns in time are automatically shifted back to this ± 15 ns window. The actual times for the impulse responses of Figures 8 and 9 are exactly 30 ns greater than those shown.

Figure 8 shows the impulse response for a much larger rectangular cavity. The dimensions

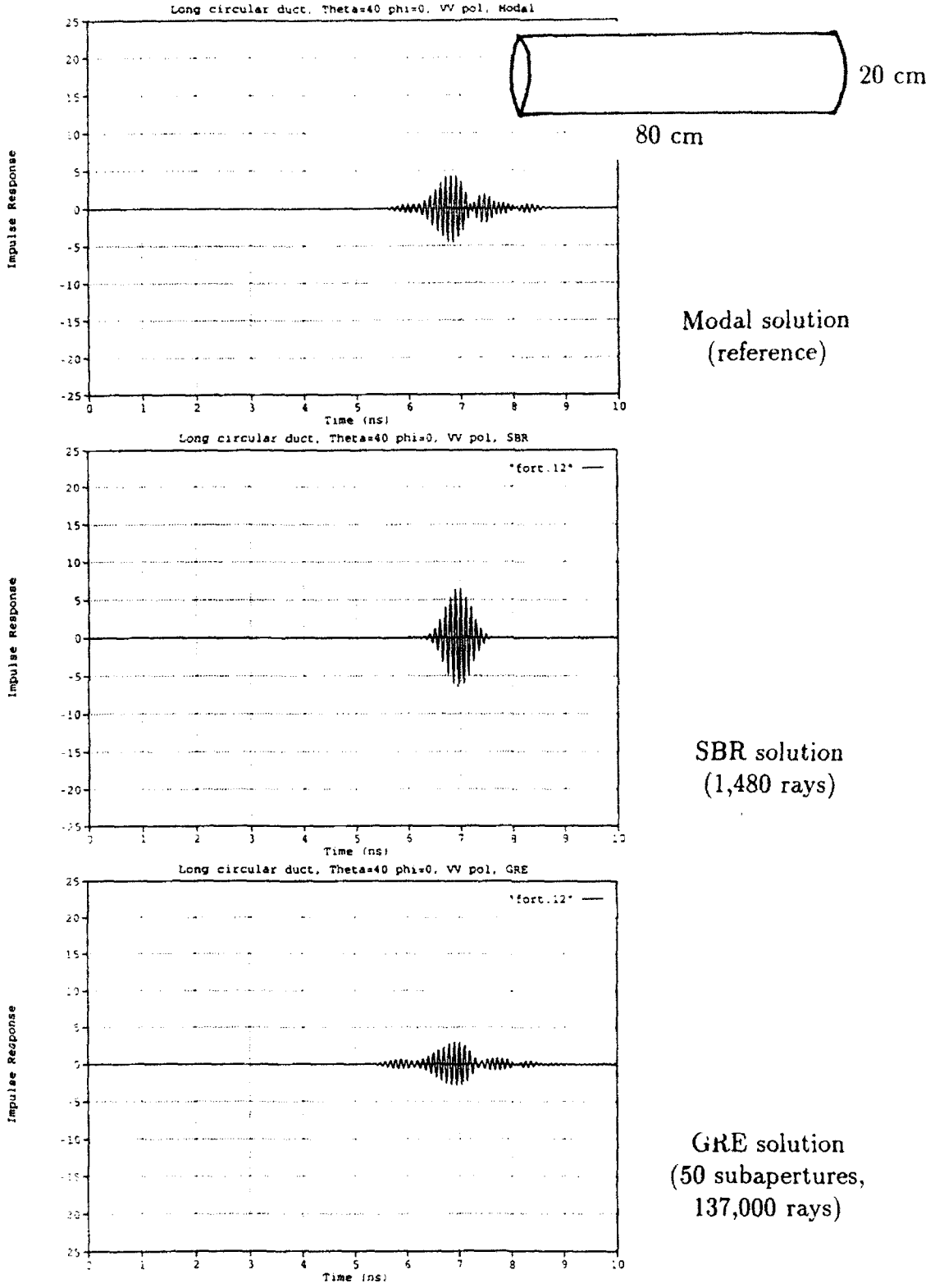


Figure 7: Cavity impulse response of an open-ended circular waveguide cavity, illuminated 40° off-axis.

of this cavity are comparable to an F-15 engine inlet duct. At midband the cavity is 30.48 wavelengths high, 20.32 wavelengths wide, and 152.40 wavelengths long. For this case the dispersion effect is still noticeable in the modal and GRE results, but is much less significant than in the smaller cavities of Figures 6 and 7. As before, the SBR method predicts the main peak but not the time dispersed peaks. In the GRE method, only rays within 5.6° of the incidence direction are included in this computation. It is estimated that a total of 5,674,000 rays would be required to accommodate all incidence angles within 60° of the waveguide axis. Using the SBR method over this same angular sector (1° increments), is it estimated that a total of 36,787,000 rays would have to be tracked.

Figure 9 shows the impulse response of a very large open-ended circular waveguide cavity. Again, the dimensions of this cavity are comparable to an F-15 engine inlet duct. At midband the cavity is 27.09 wavelengths in diameter and 152.40 wavelengths long. Inside circular cavities the GO fields tend to diverge as they reflect from the curved walls, so the SBR method gives a more spread out time domain impulse response compared with a rectangular waveguide cavity which has planar walls. For this case the SBR method gives a result which is somewhat spread out in time and agrees very well with the modal reference solution; only the less significant time dispersed peaks are missed, as expected. The GRE result predicts the peaks in the right location, but the amplitudes of the peaks are somewhat in error. In the GRE method, rays within 4.2° of the incidence direction are included in this computation. It is estimated that a total of 192,910,000 rays would be required to accommodate all incidence angles within 60° of the waveguide axis. This large number of rays is mainly due to the large number of subapertures required. Using the SBR method over this same angular sector (1° increments), is it estimated that a total of 283,614,000 rays would have to be tracked.

CONCLUSIONS AND FUTURE WORK

It has been shown that open cavities exhibit a dispersive effect on the EM scattering of a plane wave in the time domain. This is due to the fields diffracted into the cavity by the rim at the open end, which undergo multiple bounces inside the cavity before re-radiating out through the open end. This time dispersive effect differentiates open cavities from simple

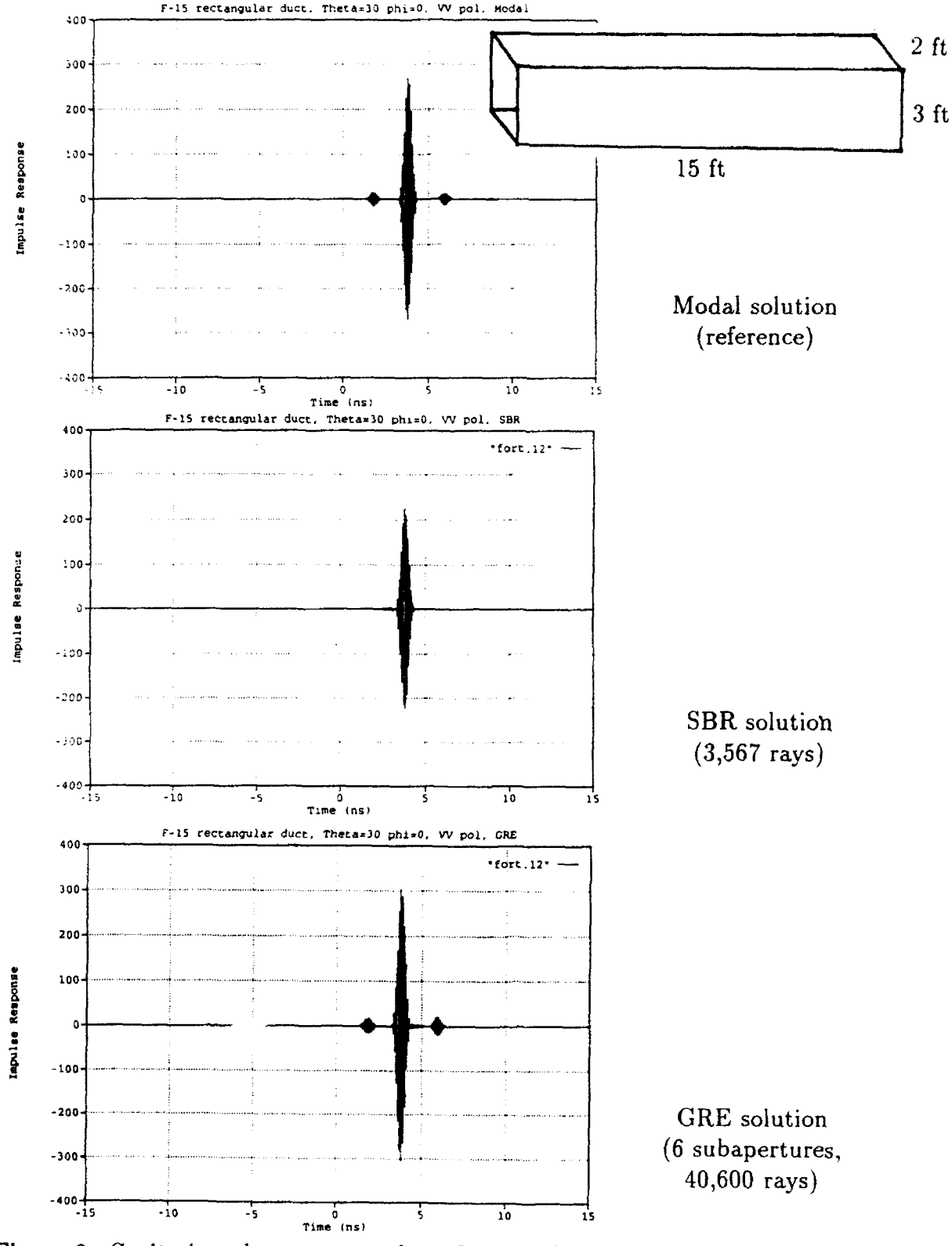


Figure 8: Cavity impulse response of an F-15 sized open-ended rectangular waveguide cavity, illuminated 30° off-axis.

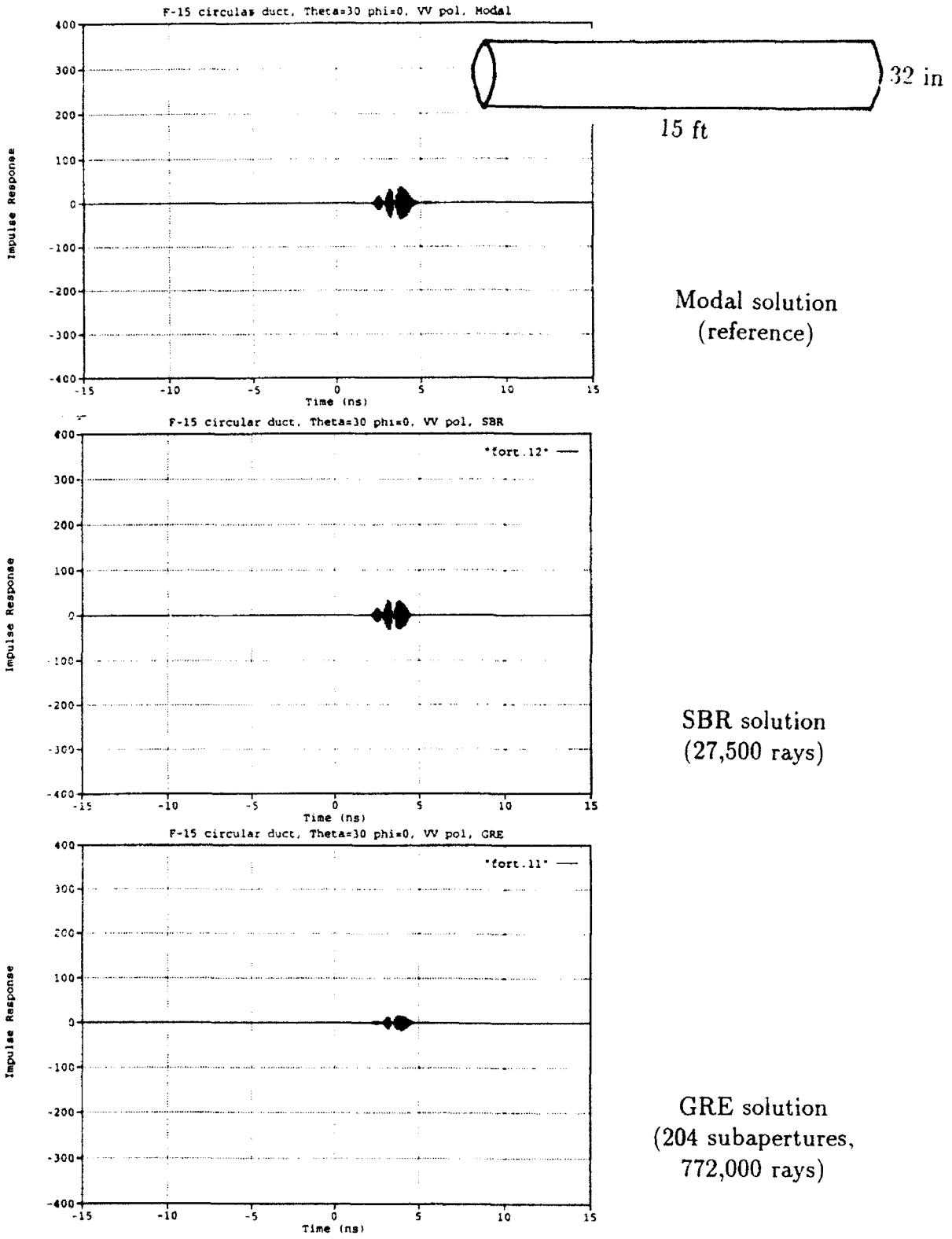


Figure 9: Cavity impulse response of an F-15 sized open-ended circular waveguide cavity, illuminated 30° off-axis.

scatterers, which would return a single impulse in the time domain.

The SBR method is very versatile and easy to implement, but it does not predict cavity time dispersion effects because the fields diffracted into the cavity by the open end are not included. However, it has been shown that for very large cavities the diffracted field is much weaker than the GO field, so the SBR method gives very good agreement with the reference solution for such cases.

The GRE method is versatile and predicts cavity time dispersion effects because the fields diffracted into the cavity by the open end are intrinsically included. However, this method is not as straightforward to implement as the SBR method because the aperture at the open end of the cavity must be gridded up into subapertures. Also, this method is sometimes not as accurate as the SBR method for very large cavities because additional approximations are introduced by the subaperture gridding. Finally, for cavities as large as the ones shown in Figures 8 and 9, the GRE method uses so many rays that it is not generally practical to store them for later use. (A single ray requires 64 bytes of memory.)

In conclusion, as a general rule of thumb when analyzing the EM scattering from cavities in the time domain, it is suggested that the GRE method should be used for cavities which are 10 wavelengths or less in width and the SBR method should be used for cavities which are 20 wavelengths or more in width. In the range from 10 to 20 wavelengths, either method could be used depending on the degree of accuracy required and the number of incidence angles. If the rays are stored and re-used in the GRE method, the SBR method becomes less efficient in ray tracing than the GRE method if on the order of 100 or more incidence angles are to be computed for a given cavity geometry. To cover all incidence angles within the full 60° angular sector (with 1° spacing), as in equation (1), the SBR method would in general require far more rays than the GRE method.

Both of the ray methods evaluated here are very versatile and relatively easy to implement for realistically large, complex open cavities, such as jet engine inlets. However, it is clear that the accuracy of both methods is limited by their respective approximations. In the SBR method, the true fields inside the cavity are approximated by only the GO fields. In

the GRE method, the fields radiated into the cavity (by the equivalent sources defined by the incident field in the aperture) are approximated by the fields radiated from finite sized subapertures in the presence of the cavity walls. In both methods, the ray tracking method is approximate because it ignores any diffraction effects from the inside of the cavity, such as creeping waves and diffraction from discontinuities inside the cavity. (This does not refer to the fields diffracted into the cavity by the rim at the open end, which the GRE ray launching method intrinsically includes.)

Also, both methods use an approximate aperture integration scheme based on the requirement that the cross-sectional areas of the ray-tubes are electrically small when they exit the cavity, as discussed in the numerical results section. As part of a future research effort, it would be extremely helpful if a more efficient field tracking scheme could be developed which overcame the "one square wavelength" limitation on the cross-sectional size of the ray-tubes. For example, the ray-tubes could be replaced with some sort of "basis field propagators" which would have well defined propagation and reflection properties and for which the fields in a cross-section would be known after multiple reflections. This could dramatically reduce the number of ray-tubes, or "basis field propagators," and could also increase the accuracy.

Currently, the GRE method has only been applied to relatively simple cavities. In the future it should be applied and evaluated for realistically complex inlets, such as the F-15 engine duct of Figure 1, using a CAD model of the geometry (this has already been done with the SBR method). Also for future work, the complexities of the inlet termination (i.e., the engine face), and other discontinuous structures inside the cavity which cause additional diffraction, need to be taken into account.

References

[1] R.E. Ball, *The Fundamentals of Aircraft Combat Survivability Analysis and Design*, New York: AIAA Education Series, 1985.

[2] H. Ling, R. Chou, and S.W. Lee, "Shooting and Bouncing Rays: Calculating RCS of an Arbitrary Cavity," *IEEE Trans. on Antennas and Propagation*, Vol. AP-37, No. 2, pp. 194-205, February 1989.

[3] R.J. Burkholder, "High-Frequency Asymptotic Methods for Analyzing the EM Scattering by Open-Ended Waveguide Cavities," Ph.D. Dissertation, The Ohio State University, June 1989.

[4] R.J. Burkholder, R. Chou and P.H. Pathak, "Two Ray Shooting Methods for Computing the EM Scattering by Large Open-Ended Cavities," *Computer Physics Communications,* Vol. 68, pp. 353-365, 1991, North-Holland Science Publishers.

[5] R.J. Burkholder and P.H. Pathak, "High-Frequency Electromagnetic Scattering by Open-Ended Waveguide Cavities," *Radio Science,* Vol. 26, pp. 211-218, January-February 1991.

[6] G.A. Deschamps, "Ray techniques in electromagnetics," *Proc. of the IEEE,* Vol. 60, pp. 1022-1035, September 1972.

[7] R.F. Harrington, *Time Harmonic Electromagnetic Fields,* McGraw Hill, New York, 1966.

[8] S.W. Lee, H. Ling and R. Chou, "Ray-Tube Integration in Shooting and Bouncing Ray Method," *Microwave and Optical Tech. Letters,* Vol. 1, No. 8, October 1988.

[9] P.H. Pathak and R.J. Burkholder, "Modal, Ray and Beam Techniques for Analyzing the EM Scattering by Open-Ended Waveguide Cavities," *IEEE Trans. on Antennas and Propagation,* Vol. AP-37, pp. 635-647, May 1989.

[10] R. Chou, T.T. Chia, and R. Lee, "The Energy Flow Inside a Waveguide Cavity Using the SBR and GRE Methods," presented at the IEEE Antennas and Propagation International Symposium, Chicago, Illinois, July 1992.

ONE-DIMENSIONAL WAVE MECHANICS MODEL FOR TERMINAL BALLISTICS

E. Eugene Callens, Jr.
Associate Professor
J. Scott McMurtry
Graduate Student
Department of Mechanical and
Industrial Engineering

Louisiana Tech University
Ruston, LA 71272

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

August 1992

ONE-DIMENSIONAL WAVE MECHANICS MODEL FOR TERMINAL BALLISTICS

E. Eugene Callens, Jr.
Associate Professor
J. Scott McMurtry
Graduate Student
Department of Mechanical and
Industrial Engineering
Louisiana Tech University

Abstract

A quasi-steady wave mechanics model for penetration of structural targets by high strength cylindrical penetrators is presented. The model is based on the postulated physical behavior of one-dimensional penetrator response, axisymmetric target response, negligible thermal effects in the interface region, and steady state conditions at the interface between wave reflections. The developed equations constitute a comprehensive model for the terminal ballistics process including the initial transient, the developed penetration regime, and the terminal transient. The model predicts crater depth, penetrator velocity, and penetrator length as a function of time as well as the final crater diameter as a function of crater depth. It also gives the penetrator stress history and the stress history in the target bow region. The equations are easily programmed and readily solved on a personal computer. Typical run times are 15 seconds on a 25-MHz 386 PC with a math coprocessor. Comparisons with experimental data for several different metal penetrators into metal targets are presented for penetrator length to diameter ratios from 3 to 20. The predictions are shown to be within the experimental uncertainties for the entire velocity regime of interest from 500 to 4,500 m/s. Two significant characteristics of the method are that it contains no empiricisms and it permits the quantification of physical property effects on crater depth and diameter. These effects include penetrator strength, density, length to diameter ratio, and impact velocity as well as target strength and density.

ONE-DIMENSIONAL WAVE MECHANICS MODEL FOR TERMINAL BALLISTICS

E. Eugene Callens, Jr.
J. Scott McMurtry

INTRODUCTION

When a target material is impacted by a projectile at an impact velocity sufficiently high to cause penetration, the stress levels at the penetrator/target interface exceed the ultimate compressive strength of the target. If the target material is ductile, the material will move away from the high pressure interface in a continuous plastic flow. However, if the target material is brittle, the material in the high pressure region will be pulverized into small discrete particles which move away as a densely packed particulate flow. In either case, the rate of flow away from the interface, which determines the penetration rate and the crater diameter, is a function of the difference between the interface pressure and the compressive strength of the material.

The accurate prediction of the time resolved interface pressure requires the utilization of a realistic model of the quasi-steady wave mechanics that characterize terminal ballistic phenomenology. The total time for the developed penetration phase, exclusive of the initial and terminal transients, is typically less than 200 microseconds. The corresponding deceleration levels are typically more than one million times greater than the standard value of gravitational acceleration. Thus, the classical differential equations for rigid body solid mechanics cannot adequately describe the time resolved response of either the penetrator or target material.

Additionally, the exterior surface of the penetrator is a constant pressure boundary from which expansion waves emanate in response to incident compression waves from the high pressure interface. The interaction of these release waves with the forward propagating compression waves results in a standing wave in front of the penetrator. The target flow response to this free surface type penetration is different from the flow of a continuous medium around an immersed body.

The purpose of the current work is to develop a quasi-steady wave mechanics model for this free surface type penetration of structural targets by high strength penetrators. The formulation includes initial and terminal transients as well as the developed erosion regime.

QUASI-STEADY WAVE MECHANICS MODEL

The proposed quasi-steady wave mechanics model is based on the following postulated physical behavior:

1. One-dimensional penetrator response
2. Two-dimensional, axisymmetric target response
3. Negligible thermal effects in the interface region
4. Steady state conditions at the interface between wave reflections

The significant consequences of the postulation of negligible thermal effects in the interface region are temperature independent material properties and the uncoupling of the energy equation. The governing equations are therefore the conservation of momentum, the conservation of mass, and a velocity component kinematic relationship.

The penetrator/target interface momentum equation is

$$S_P' + \rho_P' U_{PF}^2 = \rho_T' U_{TF}^2 + S_T' \qquad (1)$$

where

S_P' = Stress in compressed penetrator
S_T' = Stress in compressed target
ρ_P' = Density of compressed penetrator
ρ_T' = Density of compressed target
U_{PF} = Penetrator flow velocity
U_{TF} = Target flow velocity

The one-dimensional wave continuity equation in the penetrator is

$$\rho_P = \rho_P \left(\frac{V_{WP}}{V_{WP} - U_{PP}} \right) \tag{2}$$

and for the target

$$\rho_T = \rho_T \left(\frac{\bar{V}_{WT}}{\bar{V}_{WT} - \bar{U}_{PT}} \right) \tag{3}$$

where

ρ_P = Penetrator initial density

ρ_T = Target initial density

V_{WP} = Penetrator wave velocity

$\bar{V}_{WT}$ = Mean target wave velocity in the bow region

U_{PP} = Penetrator particle velocity

$\bar{U}_{PT}$ = Mean target particle velocity

The particle velocity is the change in velocity associated with the passage of the disturbance wave. The kinematic relationship between the various velocity components is given by

$$V_P = U_{PF} + U_{PP} + U_{TF} + \bar{U}_{PT} \tag{4}$$

where V_P = velocity of the free end of the penetrator.

The velocities in the penetrator and target before and after the disturbance waves are shown schematically in Fig. 1 relative to the interface.

The wave velocities are given by

$$V_{WP} = C_{WP} + S_{IP} U_{PP} \tag{5}$$

$$\bar{V}_{WT} = C_{WT} + S_{IT} \bar{U}_{PT} \tag{6}$$

where

C_{WP} = Zero pressure wave velocity in the penetrator

C_{WT} = Zero pressure wave velocity in the target

S_{IP} = Hugoniot constant for the penetrator material

S_{IT} = Hugoniot constant for the target material

The stress in the compressed penetrator is given by

$$S_P' = \rho_P U_{PP} V_{WP} \quad \text{for} \quad U_{PF} = 0 \tag{7}$$

$$S_P' = \rho_P U_{PPM} V_{WP} \quad \text{for} \quad U_{PF} > 0 \tag{8}$$

$$U_{PPM} = \frac{C_{SP}\sigma_{SP}}{\rho_P V_{WP}} \tag{9}$$

$$C_{SP} = \nu_P^{-1} \tag{10}$$

where

U_{PPM} = Maximum penetrator particle velocity

C_{SP} = Penetrator free surface maximum stress constant

ν_P = Poisson's ratio for the penetrator material

σ_{SP} = Penetrator compressive stress

From these relationships it is observed that either the penetrator particle velocity or flow velocity is known for any condition. If the interface pressure is below the maximum allowable penetrator stress, the flow velocity is zero, and the particle velocity is the variable. If the interface pressure is above the maximum penetrator stress, the particle velocity is at its known maximum value and the flow velocity is the variable.

The stress in the compressed target is given by

$$S_T' = \rho_T \bar{U}_{PT} \bar{V}_{WT} \quad \text{for} \quad U_{TF} = 0 \tag{11}$$

$$S_T' = \rho_T \bar{U}_{PTM} \bar{V}_{WT} \quad \text{for} \quad U_{TF} > 0 \tag{12}$$

$$\overline{U}_{PTM} = \frac{C_{ST}' \sigma_{ST}}{\rho_T \overline{V}_{WT}} \tag{13}$$

$$C_{ST}' = C_{ST}^2 = \nu_T^{-2} \tag{14}$$

where

$\overline{U}_{PTM}$	=	Maximum mean target particle velocity
C_{ST}'	=	Target confined surface maximum stress constant
C_{ST}	=	Target free surface maximum stress constant
ν_T	=	Poisson's ratio for the target material
σ_{ST}	=	Target compressive stress

As in the case of the penetrator, either the target particle velocity or flow velocity is known for any condition. The penetrator deceleration is determined by the wave mechanics. In this model the compression wave that originates at the penetrator/target interface upon impact is tracked in the penetrator as it travels back and forth between the interface and the free end. The compression wave reflects from the free end as an expansion wave and the penetrator free-end velocity decreases by twice the particle velocity with each reflection. Also, the particle velocity changes sign with each wave reflection.

When the expansion wave reflects from the interface, the sum of the penetrator and target variable velocities decrease by twice the particle velocity. The new values of these velocities are determined from the interface momentum equation and the velocity component kinematic relationship. Again, the particle velocity changes sign with each wave reflection.

The penetrator traveling wave is tracked until the terminal expansion velocity exceeds the developed penetration velocity. When this occurs the developed penetration phase is complete and the remainder of the penetration process is a terminal transient cylindrical expansion of the interface pressure as presented below. The change in penetrator length due to the developed penetration phase is

$$\Delta L_{PD} = -\int (U_{PF} \pm U_{PP})\,dt \tag{15}$$

The change in crater depth due to the developed penetration phase is

$$\Delta L_{CD} = \int (U_{TF} + \overline{U}_{PT})\,dt \tag{16}$$

The final crater diameter at each penetration location is given by

$$\frac{D_C}{D_P} = \sqrt{\frac{\rho_P \overline{U}_{PF}^2 + C_{ST}\sigma_{ST}}{C_{ST}\sigma_{ST}}} \tag{17}$$

where D_P = penetrator initial diameter.

INITIAL TRANSIENT MODEL

The initial impact shock transient model is a cylindrical expansion of the interface force from the impact Hugoniot shock condition to the initial developed penetration condition. The impact Hugoniot shock stress is given by

$$\sigma_{IS} = \rho_T U_{PTS} V_{WTS} \tag{18}$$

where the target shock particle velocity is determined from the simultaneous solution of the momentum equation

$$\rho_P U_{PPS} V_{WPS} = \rho_T U_{PTS} V_{WTS} \tag{19}$$

and the velocity component kinematic relationship

$$V_{PI} = U_{PPS} + U_{PTS} \tag{20}$$

where

U_{PPS} = Penetrator shock particle velocity

U_{PTS} = Target shock particle velocity

V_{WPS} = Penetrator shock wave velocity

V_{WTS} = Target shock wave velocity

The initial developed penetration stress is given by

$$\sigma_{ID} = \left(\rho_T U_{TF}^2 + S_T\right)_{ID} \tag{21}$$

and evaluated from the solution of Eqns. 1-14 where the penetrator velocity is the initial impact velocity.

The crater length due to the initial transient expansion between these conditions is given by

$$L_{CI} = R_{PR}\left(\sqrt{\frac{\sigma_{IS}}{\sigma_{ID}}} - 1\right) \tag{22}$$

where the reference radius for the cylindrical expansion is

$$R_{PR} = \frac{V_{WTL}}{V_{WTT}} R_P \tag{23}$$

and

V_{WTL} = Target longitudinal wave velocity

V_{WTT} = Target transverse wave velocity

R_P = Penetrator radius

The time for the expansion is determined from

$$\Delta t_{TT} = \int_{R_{PR}}^{R_{PR}+L_{CI}} \frac{dz}{U_{IF}} \tag{24}$$

The interface velocity during the transient is given by a quadratic expansion

$$U_{IF} = AZ^2 + BZ + C \tag{25}$$

where the constants are evaluated from the boundary conditions

$$U_{IF} = U_{PTS} \quad \text{at} \quad z = R_{PR}$$

$$U_{IF} = \left(U_{TF} + \overline{U}_{PT}\right)_{ID} \quad \text{at} \quad z = R_{PR} + L_{CI} \tag{26}$$

$$\frac{dU_{IF}}{dz} = 0 \quad \text{at} \quad z = R_{PR} + L_{CI}$$

The length of penetrator loss during the initial transient expansion is obtained from a form of the velocity kinematic relationship

$$\Delta L_{PI} = -\left(V_{PI}\Delta t_{IT} - L_{CI}\right) \tag{27}$$

The release waves from the penetrator free surface create tension regions at the interface during the transient process. These tensile waves follow the compression waves into the penetrator and attenuate the deceleration effect. Because of this influence, it is assumed that the penetrator free-end remains at the initial impact velocity during this transient time and that a compression wave is generated at the interface at the beginning of the developed penetration phase.

TERMINAL TRANSIENT MODEL

The terminal transient model is a cylindrical expansion of the interface force from the developed penetration condition to the target free surface condition. The expansion is initiated when the terminal expansion velocity exceeds the developed penetration velocity. The developed penetration stress is

$$\sigma_{DP} = \rho_T' U_{TF}^2 + S_T' \tag{28}$$

and evaluated from the solution of Eqns. 1-14 when the tensile wave reflects from the interface. The target free surface stress is

$$\sigma_T = C_{ST}\sigma_{ST} \tag{29}$$

The crater length due to the terminal transient expansion between these conditions is given by

$$L_{CT} = R_{PR}\left[\sqrt{\frac{\sigma_{DP}}{\sigma_T}} - 1\right] \tag{30}$$

Where the reference radius for the cylindrical expansion is given by Eqn. (23).

The time for the expansion is determined from

$$\Delta t_{TT} = \int_{R_{PR}}^{R_{PR}+L_{CT}} \frac{dz}{U_{IF}} \tag{31}$$

The interface velocity during the transient is given by a quadratic expansion as in Eqn. (25) where the constants are evaluated from the boundary conditions

$$U_{IF} = U_{TF} + \bar{U}_{PT} \quad \text{at} \quad z = R_{PR}$$

$$U_{IF} = \frac{C_{ST}\sigma_{ST}}{\rho_T \bar{V}_{WT}} \quad \text{at} \quad z = R_{PR} + L_{CT} \tag{32}$$

$$\frac{dU_{IF}}{dz} = 0 \quad \text{at} \quad z = R_{PR} + L_{CT}$$

Since the interface expansion velocity exceeds the developed penetration velocity, the penetrator must decelerate at the expanding interface rate.

This means that the penetrator particle velocity decreases below the maximum value and the penetrator ceases to erode. The final penetrator length is therefore the value at the end of the developed penetration phase.

COMPARISON WITH EXPERIMENTAL DATA

The above equations constitute a comprehensive model for the complete terminal ballistics process including the initial transient, the developed penetration regime, and the final transient. The model predicts crater depth, penetrator velocity, and penetrator length as a function of time as well as the final crater diameter as a function of crater depth.

The equations have been programmed using the True BASIC language and the program is readily solved on a personal computer. Typical run times are 15 seconds on a 25-MHz 386 PC with a math coprocessor.

Comparisons with experimental data are presented in Figs. 2-9. Figs. 2 and 3 show the ratio of penetration depth to original penetrator length (P/L) for tungsten alloy penetrators into rolled homogeneous armor (RHA) targets. In Fig. 2, the penetrators have a nominal original length to diameter ratio (L/D) of 20 and impact velocities range from 1900 to 4500 m/s. The one-dimensional wave mechanics code predicts the experimental results (Ref. 1) to within the uncertainty of the data.

Fig. 3 presents comparisons of the calculated P/L results with data from Ref. 2 for L/D = 3, 6, and 12. Impact velocities range from 600 to 1600 m/s. The predictions show the experimentally observed L/D effects. Both the experiment and calculations show a change in slope at an impact velocity of approximately 800 m/s. This corresponds to the velocity at which the RHA target reaches the maximum confined surface stress and begins to flow.

These same comparisons are illustrated in Fig. 4 for tungsten into tungsten at an L/D of 10.4. The experimental data are from Ref. 3 and cover an impact velocity range from 600 to 3700 m/s. The significance of the velocity regime is that it covers the range from low velocity where target strength effects are dominant to hypervelocity where strength effects are

small and a hydrodynamic condition exists. The predictions are within the data uncertainty over the entire range. There is a change in slope at approximately 900 m/s in both the data and predictions, corresponding to the velocity where the target begins to flow.

Additional results from Ref. 3 for a tungsten alloy penetrator into a steel target are presented in Fig. 5. The comparisons show the same trends and excellent agreement between the experiment and predictions as for tungsten into tungsten.

Fig. 6 presents P/L data from Ref. 2 for RHA into RHA for L/D = 3, 6, and 12. The change in slope occurs at an impact velocity of approximately 1200 m/s for these materials and strengths in both the data and calculations.

Comparisons are made with data from Ref. 4 for hardened RHA into annealed RHA for an L/D of 5 in Fig. 7. Additional steel into steel data from Ref. 3 are shown in Fig. 8 for an L/D of 10. The comparisons are similar to those in previous figures.

Fig. 9 shows a cold-worked copper into annealed copper from Ref. 5 for an L/D of 20. The annealed copper is subject to significant strain hardening effects and it is essential to use the compressive strength at the large values of strain experienced during penetration. This corresponds to 425 MPa for the target material in this case.

The model has significant value for parametric studies of the influence of penetrator and target variables as is observed in Figs. 10 and 11. Fig. 10 illustrates the effect of penetrator L/D on penetration depth for an RHA penetrator with 1 GPa compressive strength into a 1 GPa RHA target. Results are presented for values of L/D of 1, 2, 3, 6, 15, and 30. The differences in penetration for the L/D of 15 and 30 are seen to be very small. However, the effect of L/D increases significantly as the value decreases below 15. These effects are due entirely to the initial and terminal transients and account for the observed increase in penetration associated with segmented penetrators.

The effect of target compressive strength on penetration depth is illustrated in Fig. 11 for an RHA penetrator with 1 GPa compressive strength into concrete. Results are presented for target strengths of 20, 40, 60, and 150 MPa. It is observed that for the weaker targets, there is a maximum penetration corresponding to the impact velocity where the penetrator reaches its maximum allowable stress and begins to flow. For the 20 MPa concrete this occurs at an impact velocity of approximately 1500 m/s. At an impact velocity near 3500 m/s, the influence of target strength is negligible. However, above this velocity the strength curves begin to diverge because of the influence of the terminal transient.

It is concluded that a mathematical model of the terminal ballistics process based on postulated physical mechanisms with no empiricisms has been developed. It is easy to apply and is readily programmed for a personal computer. Computational run times are very short and the results accurately predict experimental results for the materials investigated. A major contribution of the method is that it quantifies physical property effects such as strength, density, geometry, and impact velocity on crater depth and diameter.

REFERENCES

1. Silsby, G. F., "Penetration of Semi-infinite Steel Targets by Tungsten Long Rods 1.3 to 4.5 Km/s," Proc. Eight Int. Symp. Ballistics, Oct. 1984.
2. Tate, A., K. E. B. Green, P. G. Chamberlain, and R. G. Baker, "Model Scale Experiments of Long Rod Penetration," Proc. Fourth Int. Symp. Ballistics, Oct. 1978.
3. Hohler, V., and A. J. Stilp, "Penetration of Steel and High Density Rods in Semi-infinite Steel Targets," Int. Symp. on Ballistics, 1977.
4. Wilson, L. L., Joseph C. Foster, Jr., S. E. Jones, and Peter P. Gillis, "Experimental Rod Impact Results," Int. J. Impact Eng., Vol. 8, No. 1, pp. 15-25, 1989.
5. Christman, D. R., A. B. Wenzel, and J. V. Gehring, "Penetration Mechanisms of High-Velocity Rods," Proc. of the Seventh Hypervelocity Impact Symp., Tampa, FL, Vol. V and VI, Experiments, Nov. 17-19, 1964.

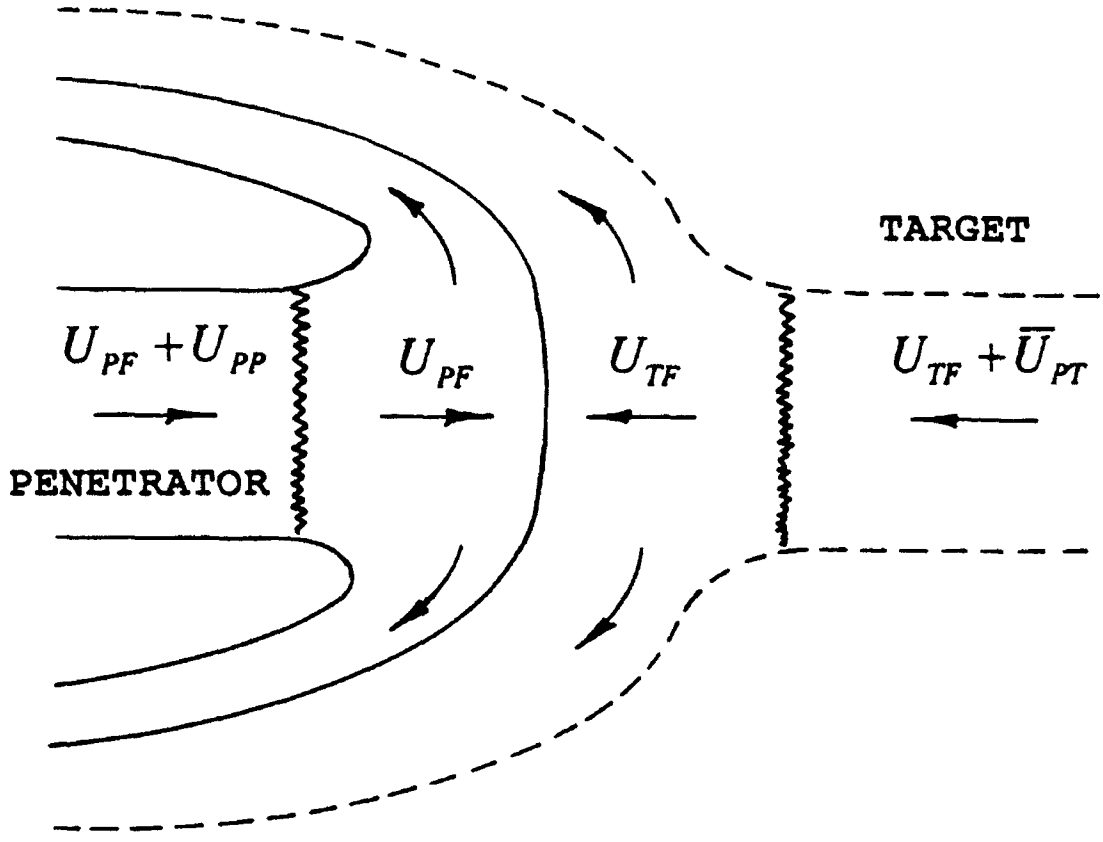


FIG. 1. FLOW AND PARTICLE VELOCITIES RELATIVE TO THE INTERFACE

FIG. 2. TUNGSTEN INTO RHA, L/D = 20

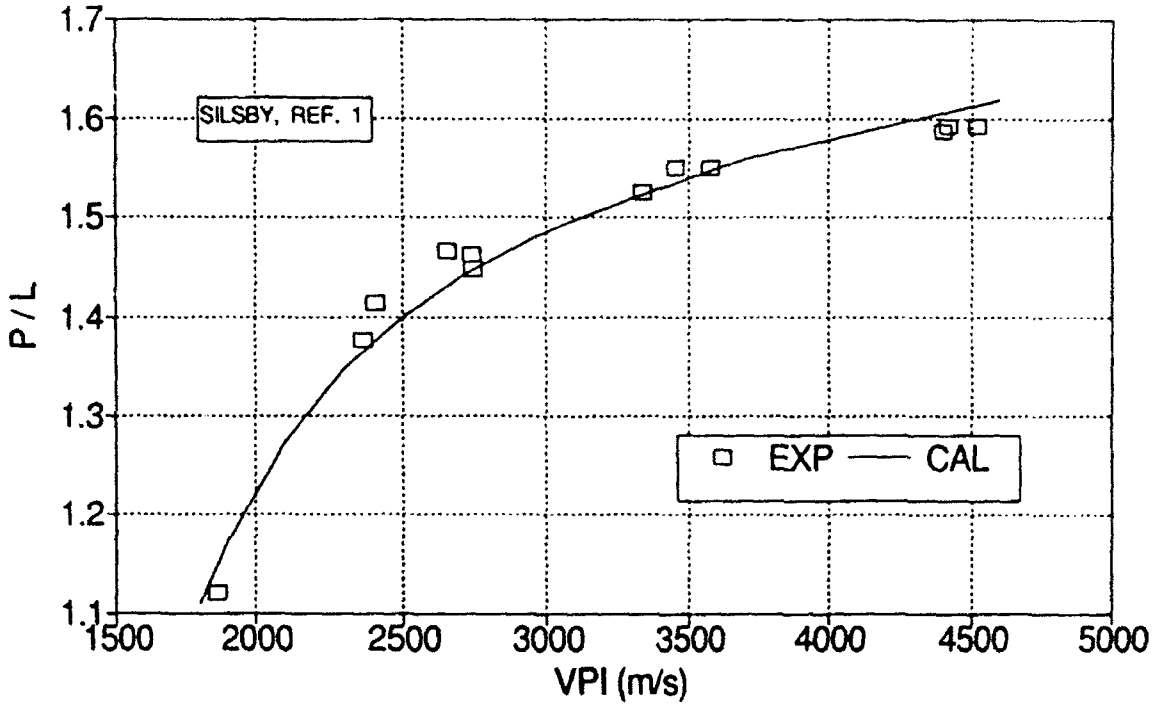


FIG. 3. TUNGSTEN INTO RHA, L/D = 3, 6, 12

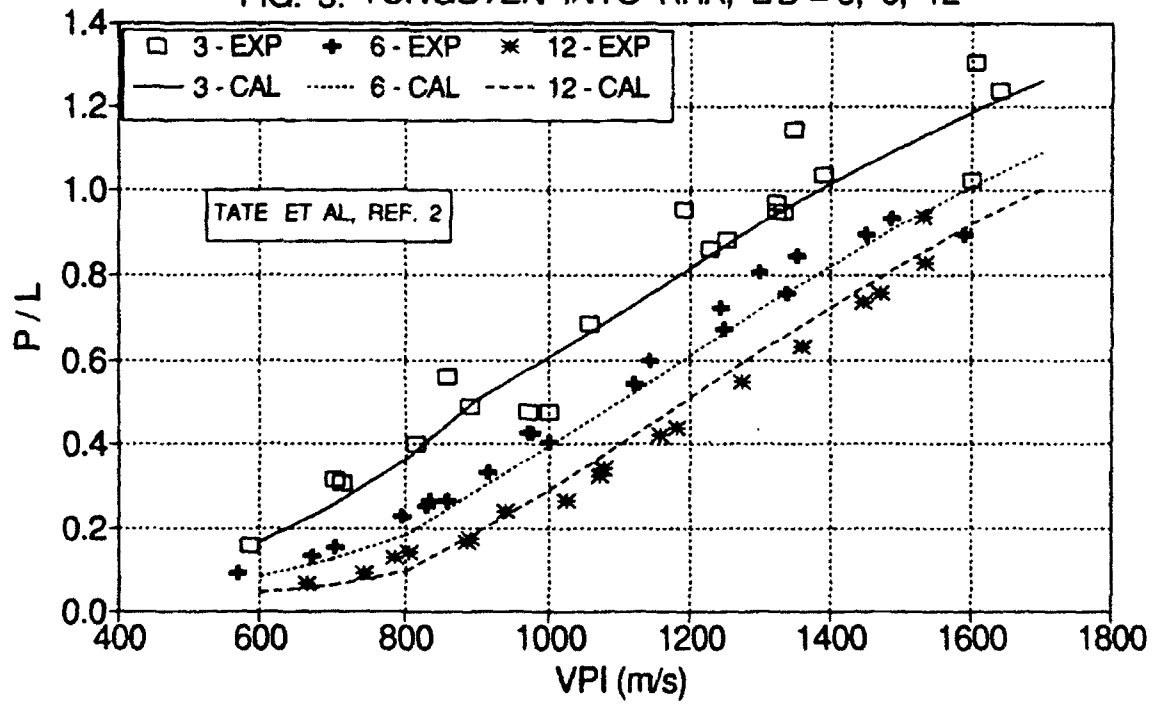


FIG. 4. TUNGSTEN INTO TUNGSTEN, L/D = 10.4

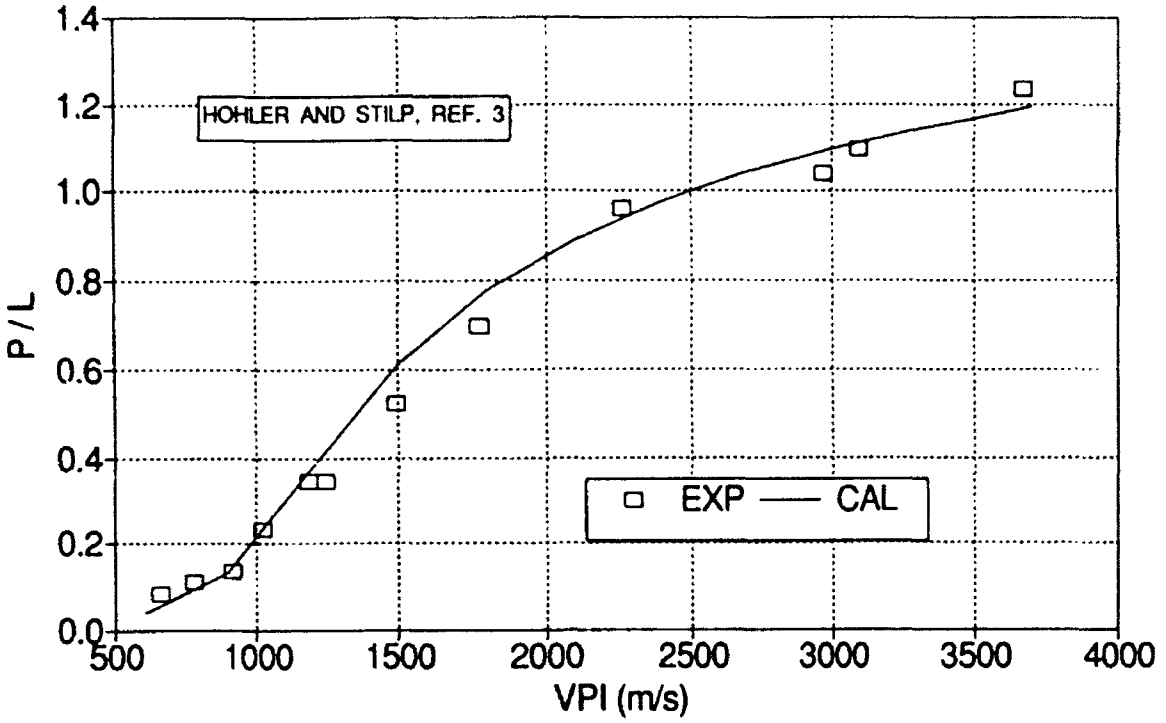


FIG. 5. D17 TUNGSTEN INTO HZB20 STEEL, L/D = 10.4

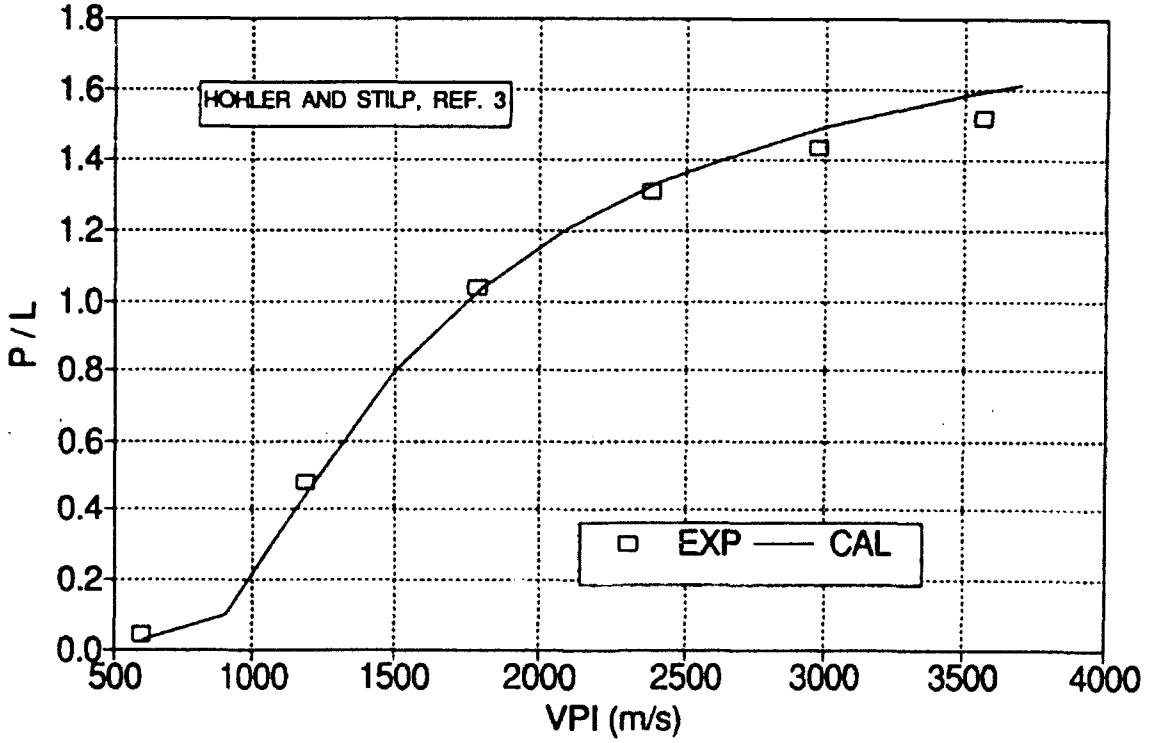


FIG. 6. RHA INTO RHA, L/D = 3, 6, 12

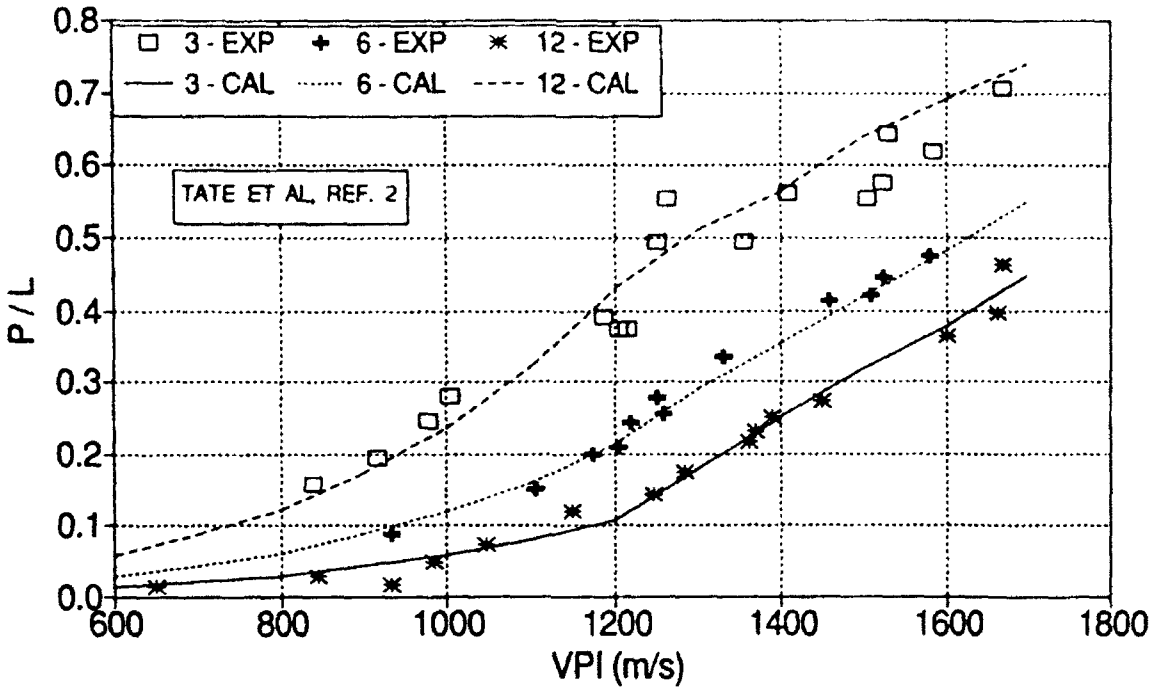


FIG. 7. HARDENED RHA INTO ANNEALED RHA, L/D = 5

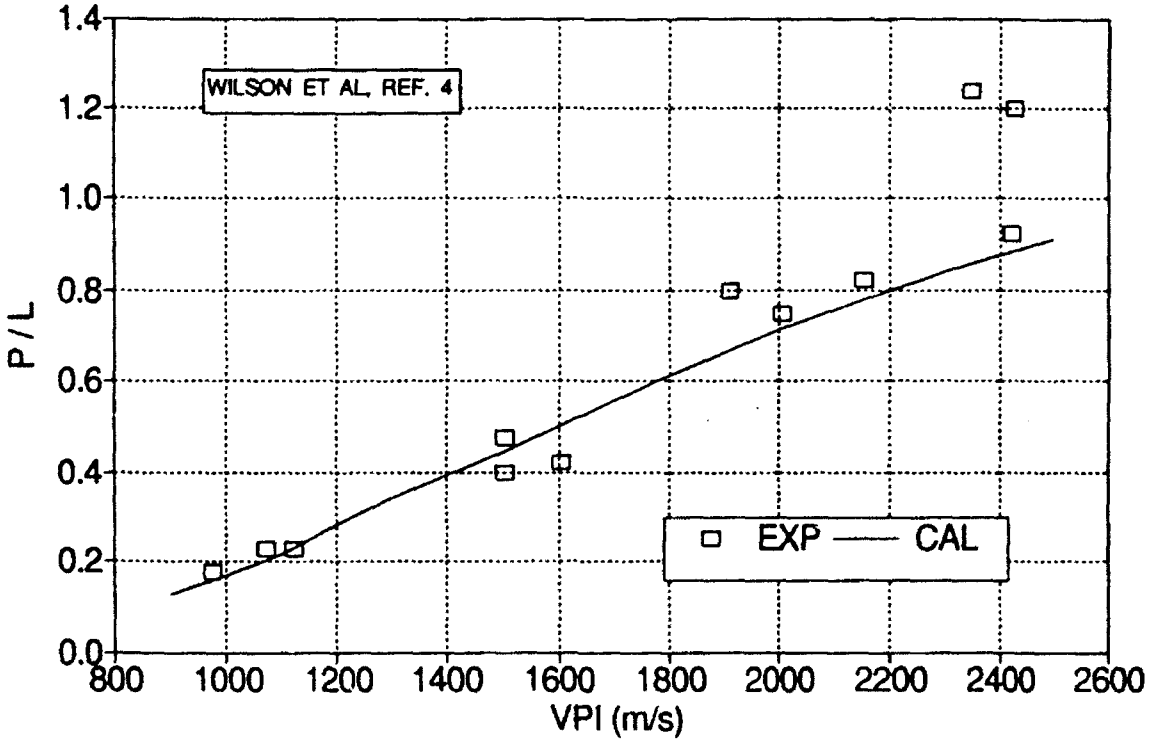


FIG. 8. C110W2 STEEL INTO HZB20 STEEL, L/D = 10

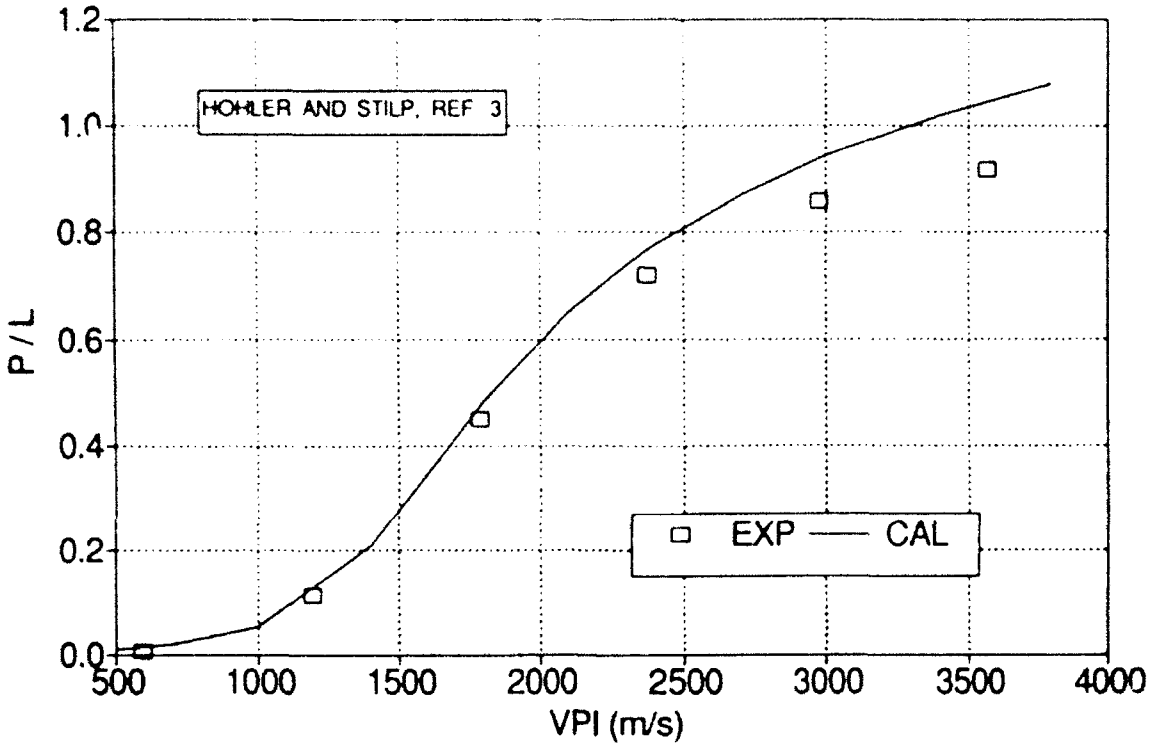


FIG. 9. COPPER INTO ANNEALED COPPER, L/D = 20

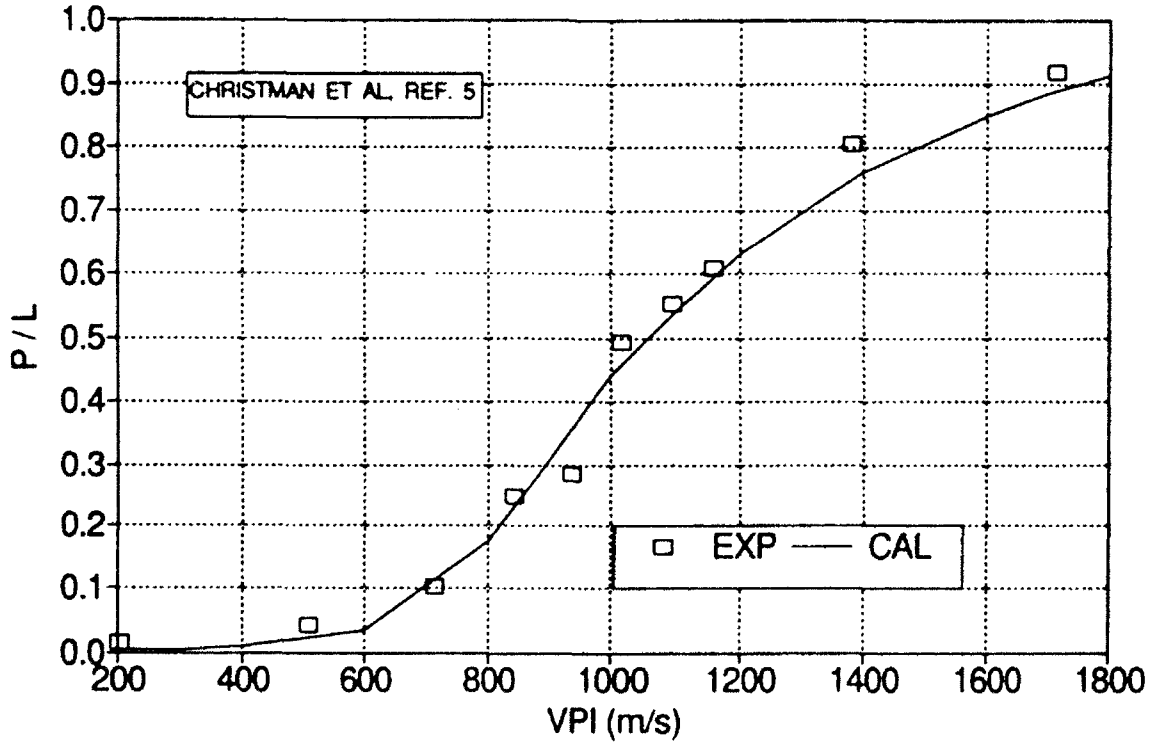


FIG. 10. EFFECT OF PENETRATOR L/D ON PENETRATION DEPTH, L/D = 1, 2, 3, 6, 15, 30

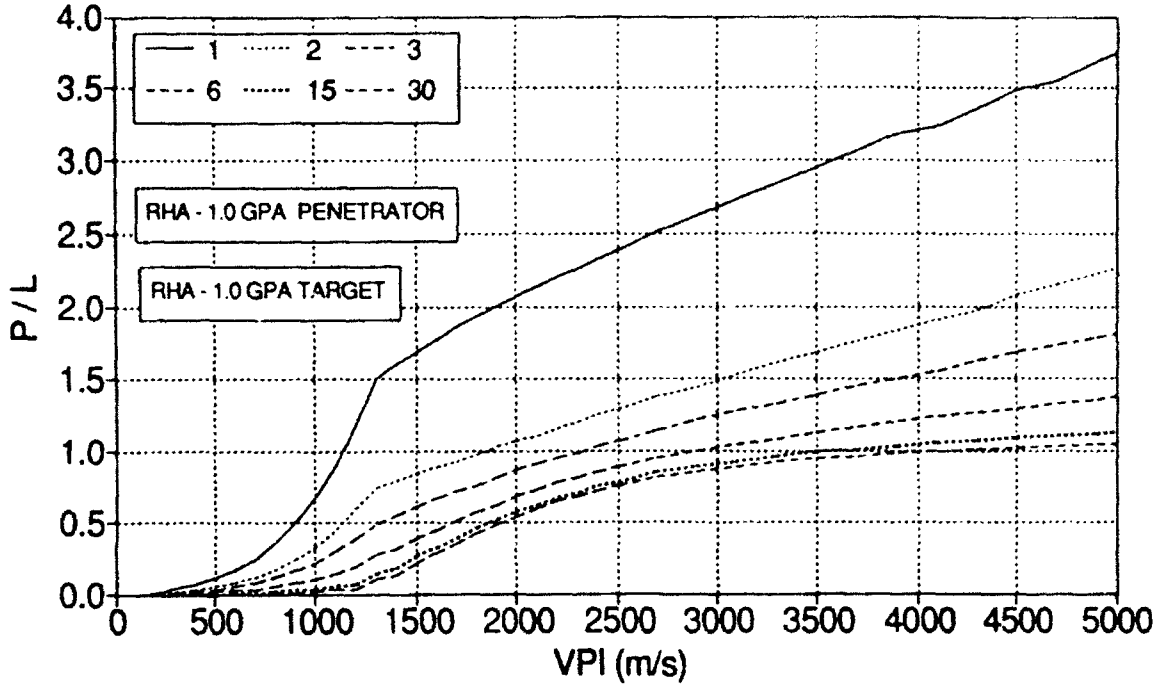
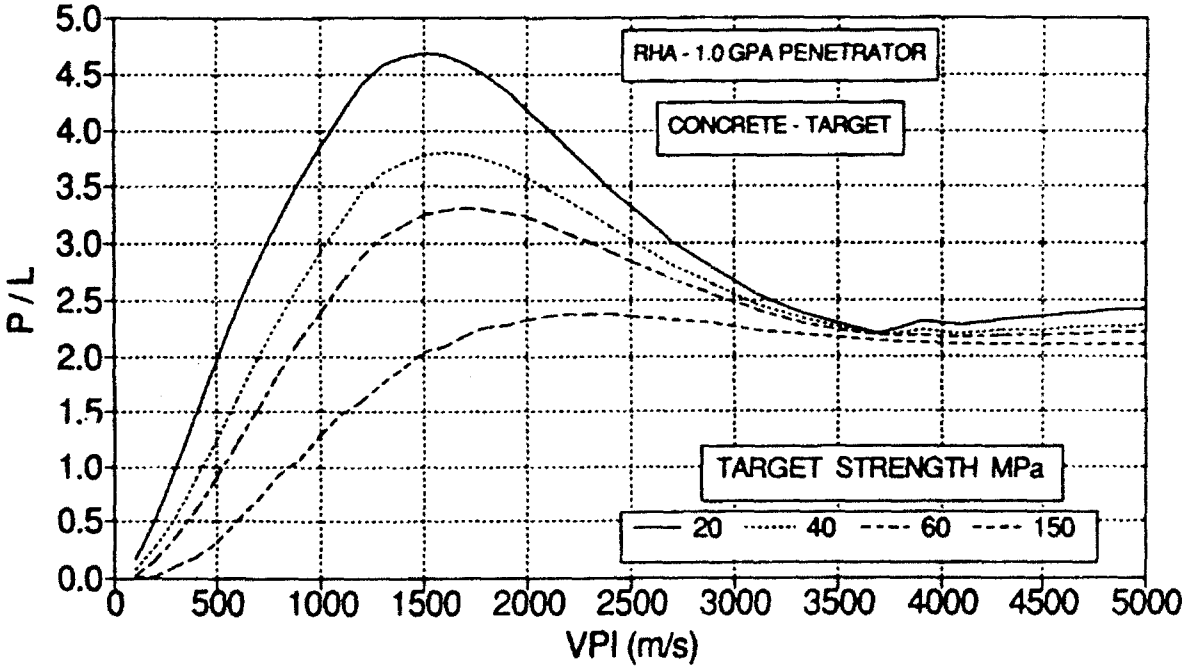


FIG. 11. EFFECT OF TARGET COMPRESSIVE STRENGTH ON PENETRATION DEPTH



AN INVESTIGATION OF THE USE OF EMBEDDED FIBER OPTIC SENSORS IN COMPOSITE MATERIALS

Gregory P. Carman
Visiting Assistant Professor
Department of Engineering Science and Mechanics

Virginia Polytechnic Institute and State University
227 Norris Hall
Blacksburg, VA 24061

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

AN INVESTIGATION OF THE USE OF EMBEDDED FIBER OPTIC SENSORS IN COMPOSITE MATERIALS

Gregory P. Carman
Visiting Assistant Professor
Department of Engineering Science and Mechanics
Virginia Polytechnic Institute and State University

Abstract

The use of embedded optical sensors in aircraft for control of structural components and health monitoring of critical members has evolved in the last decade. The emergence of the "Smart" structures and materials field provides a crucial product to United States industries for giving them the technological edge required in today's market place. However, the mechanics of embedded sensors essential to this evolution process has not been adequately addressed in the literature. It is the purpose of this report to present first an understanding of the detrimental effect embedded sensors have on the performance of a composite material and second to investigate the accuracy and usefulness of an embedded extrinsic Fabry Perot Fiber Interferometric (EFPI) strain sensor near large stress gradients. The first task is evaluated by experimentally measuring the degradation in an unidirectional graphite-epoxy's transverse strength due to the presence of an optical fiber. This parametric study which varies fiber size, fiber coating, and adhesion properties, demonstrates that the transverse strength can be degraded as much as 50%. These results also suggest that an optimum coating thickness exists for polyimide coated optical fibers which minimizes their effect on composite performance. The experimental results of the second task demonstrate that an EFPI sensor provides accurate strain data without a calibration procedure. The test data also suggests that the EFPI may provide a valuable tool in studying stiffness reduction which occurs in a composite around anomalous regions such as holes. The EFPI sensor embedded in the graphite epoxy layups failed prior to failure of the composite and their fatigue life was shorter than the composite. The sensors considered in this study are crudely manufactured, however, with proper construction the strength and fatigue life of an EFPI should rival and surpass that of a composite.

AN INVESTIGATION OF THE USE OF EMBEDDED FIBER OPTIC SENSORS IN COMPOSITE MATERIALS

Gregory P. Carman

1.0 INTRODUCTION

The work performed for the Air Force Office of Scientific Research under the summer faculty research program by this author was divided into 3 separate research thrusts. The first task was to develop a detailed literature review of the significant problems which have been and are being addressed by the mechanics community in investigating embedded optical fibers. This overview was lengthy and is not presented in this report, however, the reader is referred to an article published by the author on this subject [1] and the notes provided by the author to the Fatigue and Fracture group at Wright Patterson Air Force Base WPAFB. Additional research performed by the author during the 12 week appointment at WPAFB involved experimental work studying the response of embedded optical sensors.

In the first of the two experimental tasks, the potential detrimental effect of an embedded optical sensors on the performance of a composite material was investigated. The degradation of a composite's transverse strength due to the presence of an embedded optical fiber was used to characterize the severity of this effect. This particular mechanical property is widely recognized [2] as being one of the most sensitive to the influence of an embedded inclusion such as an optical fiber. The experiments employed in this study included varying the optical fiber diameters, varying the physical properties of the coating, and varying the adhesion between the optical fiber and the host material. The tests performed established a fundamental basis to determine if specific combinations of these quantities (i.e. fiber size, coating, and adhesion) can effectively minimize the obtrusive behavior indicative of the embedded devices.

In regard to optimizing these physical quantities, a theoretical model has recently been published by this author [3] which indicates that an optimal coating exists and is a function of fiber volume fraction, fiber size, coating thickness, and constituent material properties. This theoretical analysis and recent experimental results have suggested that appropriate coatings can effectively alter the transverse strength by 50 to 90 % [4]. Experimental results suggest that the theoretical predictions provide an adequate estimation of an optimal coating for optical fibers. A more detailed analysis of the data generated on this experimental task are provided by Carman et al. [5].

The second set of experimental tests conducted by the author at WPAFB employed the use of embedded extrinsic Fabry-Perot Interferometer (EFPI) sensors to monitor strain concentrations near simulated damage events in composite materials. The EFPI sensors were embedded near a circular hole drilled into the composite which was than subjected to cyclic loading. Externally adhered resistance strain gauges applied to the composite provide baseline data to evaluate the performance and accuracy of the embedded optical sensor. The progression of damage in the

composite was closely watched during these tests to evaluate the sensor's ability to detect and monitor the evolution of the damaged region as well as its ability to accurately perform under cyclic loading. A more complete description of damage in the composite tested in this study is provided by Paul et al. [6].

2.0 EXPERIMENTAL SETUP TASK 1

2.1 Specimen fabrication

The prepreg utilized in the present study is an AS4/3501-6 graphite fiber epoxy matrix manufactured by Hercules. Panels were laid up in a $(90_4,90_{of},90_4)$ laminate to determine the transverse tensile strength of the specimens. The 90_{of} nomenclature refers to an optical fiber placed at the midplane of the panel parallel with the structural fibers. A total of 10 specimens were cut from each panel manufactured during this investigation. Each panel contains an optical fiber with a different physical attribute (e.g. fiber diameter). A detailed discussion of these physical attributes is provided in the following subsection. The autoclave utilized to cure the panels is capable of curing 6 panels (i.e. 60 specimens) for each run.

The total number of panels required for this test was 31. Therefore, to ensure that the curing process did not influence the results, each autoclave run included a control panel. The control panels did not have embedded optical fibers in them and thus represent baseline data for strength values. In this report, the strength values obtained from panels containing optical fibers are compared with the strength values obtained from control specimens that were manufactured with those panels. To further ensure that the mechanical properties did not vary from panel to panel, all prepreg came from the same roll.

The cut composite test specimens had dimensions of 2 x 10 inches with a nominal thickness of .042 inches. The specimens were end tabbed with a gauge length between the tabs of approximately 6 inches. After cutting and tabbing each specimen, a 1/4 inch hole was drilled through the geometric center of the specimen. This was done to ensure that failure occurred in the gauge length of the specimen and not in the grip region. Prior to drilling the hole, the optical fiber was precisely located with a x-ray image of the specimen. The hole was drilled in the specimen such that the optical fiber traversed the holes major diameter and the center of the hole was within 200 microns of its intended location. The presence of the hole also ensured that failure occurred at the optical fiber as opposed to an anomalous region in the composite. For an illustration depicting the specimen see Figure 1.

2.2 Embedded optical fibers

Each of the panels fabricated for the transverse tension tests contained an optical fiber with a different physical attribute. The physical attributes which distinguishes one fiber from another can be categorized into three basic areas: optical coating, fiber diameter, or adhesion characteristic. The coatings chosen for this investigation were based upon published work in the Smart materials area [7 & 8]. Researchers have suggested that acrylate coated fiber have a

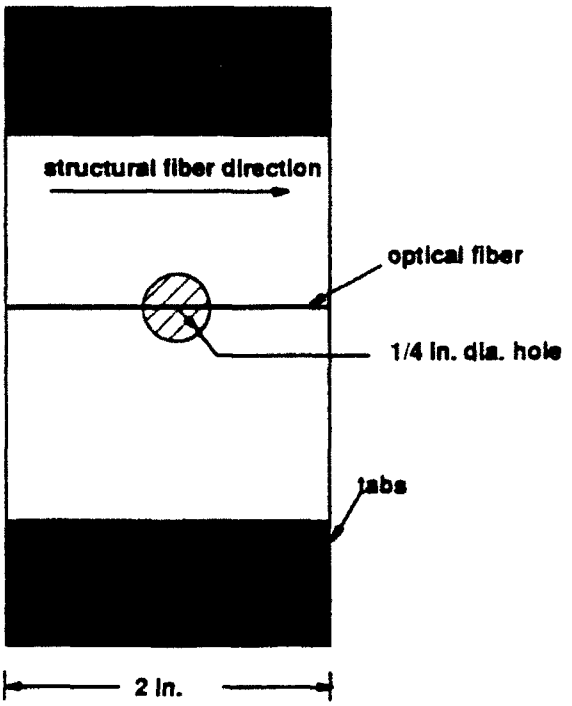


Figure 1: Illustration of a transverse tension specimen with an embedded optical fiber.

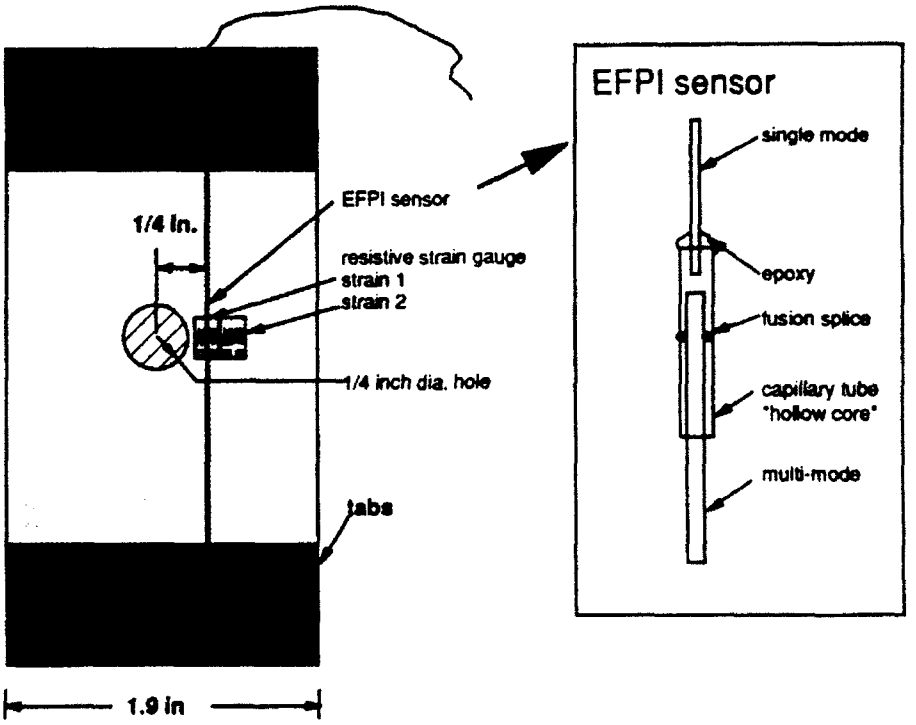


Figure 2: Illustration of a test specimen containing an embedded EFPI optical fiber sensor.

tendency to degrade during the manufacturing process and may not be compatible with epoxy systems, on the other hand, certain polyimide coated fibers can withstand processing temperatures in excess of 400 degree C and apparently are compatible with epoxy matrices. Therefore, polyimide coated optical fibers with different thickness coatings were exclusively chosen for the present study. There was one exception to this coating type, one panel contained an optical fiber coated with an extremely compliant coating with an additional exterior polyimide coating. In the results section this optical fiber is represented by ** . In addition to polyimide coated fibers, several researchers are utilizing bare or stripped optical fibers in heath monitoring systems. These systems include damage detectors and optical strain sensors. Therefore, the tests discussed in this report include a number of uncoated optical fibers embedded in the graphite epoxy composite.

The effect optical fiber size has on the strength, while qualitatively discussed in the literature, has not been quantitatively addressed in many instances. Several authors have suggested that a minimum optical fiber diameter needs to be achieved (i.e. 100 microns) to fully eliminate or minimize the obtrusive behavior of optical fibers. They suggest that if the fiber size is below 100 microns there is virtually no degradation in certain mechanical properties [2 & 9]. This should be true if the stress concentration caused by the presence of the optical fiber is smaller than the stress concentration caused by other anomalies statistically present in the composite.

In addition to the effect fiber coatings and size have on composites performance, the question of adhesion between the optic/coating and the host on strengths naturally arises. As shown by Case et al. [4], adhesion between fibers and matrix significantly alters the transverse strength of a composite both beneficially and detrimentally depending on the specific composite. Therefore, several of the coated and uncoated fibers embedded in the graphite epoxy composite were treated with Freecoat and vacuum grease to ensure no adhesion would exist between the optical fiber and the host. However, the fibers which were not treated with Freecoat or vacuum grease were thoroughly cleaned with acetone prior to embedding.

Following the guide lines described above for varying the physical attributes of the embedded optics, the following optical fibers were embedded in the AS4/3501-6 composite systems. First, polyimide coated fiber ranging from 85 micron diameter to 292 micron diameter were embedded in the composite systems. These fibers included a wide range of coating thickness values. While a large number of different diameter polyimide coated fibers were embedded, the authors point out that the optical fibers were not all from the same manufacturer. This implies that the coatings may be differ in both adhesion characteristics and mechanical properties. However, the vast majority of optical fiber manufactures suggested that the stiffness of the coating was approximately 0.5 msi. A wide range of stripped uncoated optical fibers were embedded in the transverse tension specimens also. These ranged in diameters from 40 microns up to 237 microns. To address the issue of adhesion, several of the optical fibers were coated with Freecoat and vacuum grease to assure that no bonding took place. This included uncoated optical fibers

ranging in size from 80 microns to 237 microns and coated optical fibers from 85 microns up to 292 microns.

3.0 Experimental Setup Task II

An embedded extrinsic Fabry-Perot fiber optic strain sensor was evaluated for accuracy and reliability in the second portion of this experimental study. For details on the manufacturing process and the implementation of the EFPI sensor the reader is referred to [10]. The sensors were embedded in AS4/3501-6 prepreg in two different panel layups, a cross-ply $(0_2,90_4,0_2,0_{of})_s$ and a quasi-isotropic $(0,45,-45,90_2,45,-45,0,0_{of})_s$. In both lamination schemes, the EFPI sensor is place midplane between the zero degree plies parallel to the adjacent structural fibers. These particular layups were chosen to investigate the performance of the EFPI sensor in the presence of two different damage phenomena, that is matrix cracking and ply delamination. That is, these damage phenomena are normally observed in the cross-ply and quasi-isotropic laminates when subjected to fatigue loading.

A brief discussion on problems encountered during the curing process of the panels are presented in the results section. Each fabricated panel provided 9 test specimens to evaluate fatigue performance. Five of the 9 specimens from each panel were manufactured with EFPI sensors while the other 4 were used as control specimens. Therefore, a total of 5 cross-ply and 5 quasi-isotropic specimens with EFPI sensors were manufactured for this study. The control specimens provide baseline data which indicates if the presence of the EFPI seriously affects the performance of the composite.

To study the stress redistribution around a local damaged region in a composite material with an embedded EFPI sensor, a 1/4 inch diameter hole was drilled in the center of each specimen to generate a local stress gradient near the embedded sensor. A hole was chosen due to its relative simplistic stress redistribution profile and its prevalent use in composite evaluation. The hole was precisely positioned relative to the sensor with the use of x-ray images of the specimen. X-ray images proved to be an invaluable aid for locating the optical sensor and in detecting sensor failure which occurred during the manufacturing process. The resolution of the x-ray image was such that the hollow core fiber, the single mode fiber, the multi mode fiber, and the air gap could all easily be identified. Once the gauge portion of the EFPI sensor was located, the hole was drilled a lateral distance of 1/4 inch away from the optical sensor. This location allowed the optical sensor to be situated within the region of stress redistribution (see Figure 2).

After drilling the specimens, two resistance strain gauges Micromeasurements CEA-13-125UN-350 were adhered to the surface of the test coupon to monitor surface strains. With the use of x-ray images, the resistance strain gauges were placed on the composite directly over the EFPI sensor (strain 1) and directly adjacent to the optical sensor (strain 2). Strain gauge 2 was placed on the side of the optical sensor furthest from the hole (see Figure 2). The gauge lengths of the optical sensors manufactured for this test were approximately 0.125 inches, therefore, the chosen

resistance strain gauges had a comparable gauge length . Strain gauge 2 should measure a smaller strain than the strain gauge 1 or the EFPI sensor. On the other hand, the strain gauge 1 should be comparable with the strain obtained from the optical sensor. The data obtained from the resistance gauges are utilized to validate and study trends in the data obtained from the embedded EFPI sensor.

Once the specimens were prepared, they were tested in a servo-hydraulic test frame in load control under either static tension, quasi static tension, or tension-tension cyclic tests. The tension tests were conducted by ramping the specimen to failure load. The quasi-static tension tests is defined in this report as cyclic loading at 1 hertz (tension-tension) at R=.1 for less than 10 cycles at a prescribed load level. The choice of this loading facilitated data collection procedures with the EFPI sensor. Cyclic loading occurred at 5 hertz and R=.1. The average tensile strength for the cross-ply specimens was 126 ksi while for the quasi-isotropic it was 60 ksi. In the test accomplished during this study, there was no indication that the EFPI sensor seriously degraded the static or fatigue properties of the laminate. This should be expected for the given layups, provided that the sensor did not interact with the damage region in an adverse fashion. Nonetheless, a limited sample size was utilized to investigate this phenomenon and the results should not be considered conclusive, since this was not the intended purpose of the investigation.

4.0 RESULTS

4.1 Transverse Strength

The results of the transverse strength tests are provided in Table 1. The first column corresponds to the optical fiber's core diameter (silica glass), the second column to the cladding diameter (silica glass), and the third column to the coating diameter. The coatings on all of the optical fibers considered in this report is polyimide except for the **134/152 micron coating diameter fiber. This particular optical fiber had a compliant coating (unknown) of 134 microns coated with an additional coating of polyimide 152 microns. The fourth column in Table 1 indicates if vacuum grease was applied to the optical fiber to alter adhesion. The fifth column provides the mean value of the transverse strength measured for that particular panel. The sixth column gives the strength of the control panel which was manufactured with the test panel. The reader should realize that the table is not complete at the writing of this report. Several tests are still being conducted and will be completed in the near future. The principle purpose of this report is to present the major findings of the research performed at WPAFB. For a more detailed and complete discussion of these results the reader is referred to a Carman et al. [5].

In one of the autoclave runs (i.e. control strength 1.19/1.58 ksi), the strengths of every panel manufactured including the control were extremely low. The reason for these low strength values is presently being investigated, however, it does not appear to be related to fiber volume fraction or void volume fraction. The measured volume fraction values were consistent with plies manufactured in other autoclave runs. In the first three rows of Table 1, data obtained from the anomalous autoclave run on uncoated optical fibers is presented. The panel's strength values

TABLE 1: Transverse Strength Results

Core dia. (microns)	Cladding dia. (microns)	Coating dia. (microns)	Grease	Strength (ksi)	Controll (ksi)
40	-	-	no	2.50	1.19/1.58
68	-	-	no	1.54	1.19/1.58
80	-	-	no	1.61	1.19/1.58
140	-	-	no	3.97	4.49
237	-	-	no	2.18	4.49
80	-	-	yes		
140	-	-	yes		
237	-	-	yes	1.38	4.49
50	70	85	no	4.49	3.82
80	-	90	no	4.58	4.64
80	-	** 134/152	no	2.82	4.49
101	112	126	no	3.56	3.82
101	140	172	no	3.62	3.82
125	-	133	no	3.66	4.64
150	165	185	no	4.21	4.64
199	220	238	no	3.46	4.64
203	237	292	no	4.05	4.64
???	???	245	no	2.17	1.19/1.58
50	70	85	yes	4.36	3.82
101	140	172	yes	4.00	3.82
203	237	292	yes		

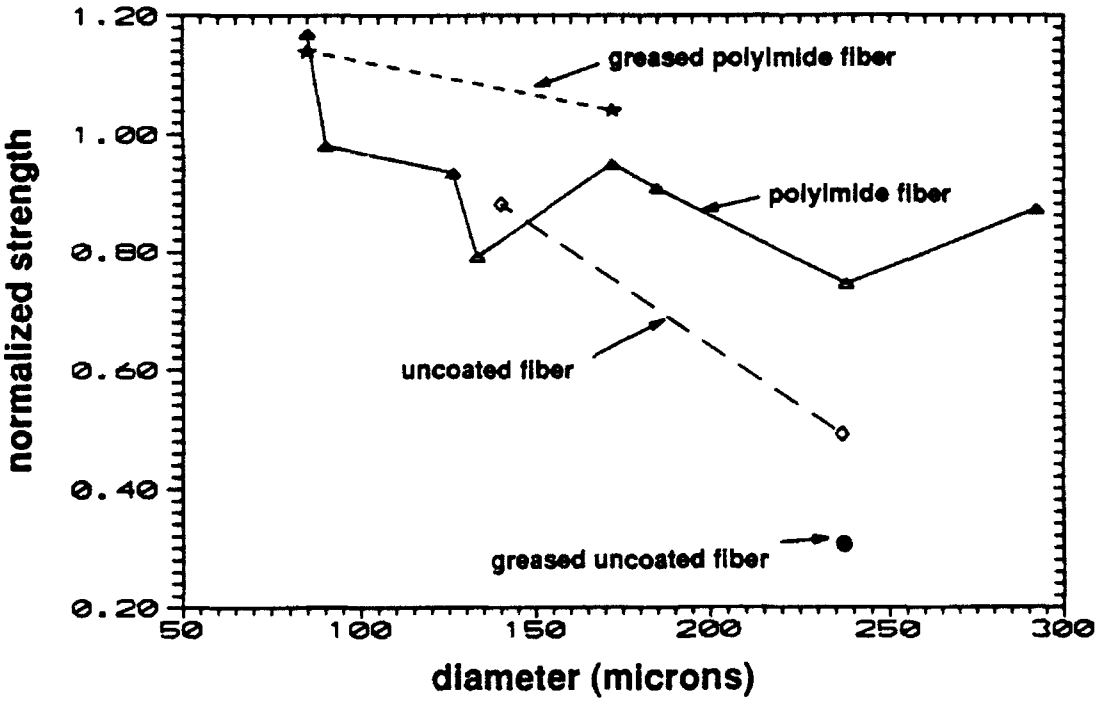


Figure 3: Depiction of the effect optical fiber size and adhesion has on transverse strength.

increase as the optical fiber diameter decreases. However, all of the measured strength values are larger than that obtained from the control panel. Therefore, these results are inconclusive and may actually be misleading. The results from this autoclave run are not considered in subsequent discussions.

In Figure 3, a comparison of normalized strength versus optical fiber diameter is presented for the data described in Table 1. The strength of each panel is normalized to the strength of the control panel it was manufactured with. The outer optical fiber diameter is defined in this graph by the outer most dimension of the fiber, including coating. Four curves are provided in Figure 3 representing polyimide coated fibers, greased polyimide coated fibers, uncoated optical fibers, and greased uncoated optical fibers. The curves for the greased uncoated fiber curve and the greased polyimide coated fiber are incomplete due to ongoing testing. However, results are presented in Carman et al. [5].

For the data presented in Figure 3 concerning the uncoated optical fiber, it appears as though strength is extremely dependent upon fiber size. That is, a 140 micron uncoated fiber decreases the strength by 12% but a 237 micron uncoated fiber decreases the strength by 52%. However, when analytically studying this problem the stress concentration calculated for a 237 micron diameter fiber is not significantly larger than the stress concentration for a 140 micron diameter fiber, that is only a 5% change. This theoretical analysis is based on a concentric cylinders model and a relatively small optical volume fraction which is consistent with the panels experimentally investigated in these tests. Therefore, if the change in size of the uncoated optical fiber should not alter the strength as significantly as depicted in Figure 3, the question naturally arises at to what causes this strength reduction. The strength decrease may be the result of an anomalous manifestation which arises in the composite during the manufacturing process and is a function of optical fiber size, e.g. voids. That is, a larger optical fiber may create other stress risers and may be directly related to the evolution of a finite interphase region around the optical fiber.

The experimental results for polyimide coated fibers presented in Figure 3 suggests that strength decreases for larger optical fiber sizes. However, for some optical fiber sizes it increases. The polyimide coated fibers exhibit larger strength values than measured for uncoated optical fibers. For coated optical fibers smaller than 100 microns it appears as though the strength of the composite is unaffected, which agrees with other researchers results [2 & 9]. However, in the panels with polyimide coatings there exists a complex interaction between coating thickness and strength which is not readily apparent from the data presented in Figure 3. This is evident by the inconsistent trend in strength as a function of fiber diameter. That is, not continually decreasing as fiber diameter is increasing. A more detailed discussion describing this interaction is provided in a subsequent figure.

In regard to adhesion affects, the strength of composites with greased uncoated optical fibers is significantly smaller than any of the other values. However, altering adhesion between coated polyimide fibers with the composite appears to have no affect on the transverse strength. This tends to contradict the assumption that adhesion between

a polyimide coated optical fiber and a composite is stronger than between an uncoated optical fiber and a composite [11]. If this were true one would expect the opposite trend to that depicted in Figure 3. That is, vacuum grease would have no affect on the uncoated fiber and would significantly effect the transverse strength of a polyimide coated fiber. A more detailed discussion of this result will be provided by Carman et al. [5] once all tests are completed and the results of scanning electron microscopy are in.

In Figure 4, a comparison of the normalized strength for polyimide coated fibers versus the normalized coating thickness is presented. This figure provides an indication of the effect coating thickness has on transverse strength. The volume fraction or normalized coating thickness depicted in the figure is a ratio of the radius of the silica glass to the radius of the polyimide coating squared. This effectively gives the local optical fiber volume fraction, such that when the value is one no coating is on the fiber, however when the value is zero there is no silica glass. This value is utilized for comparison purposes consistent with the analysis of Carman et al. The triangles represent the experimental data from each specimen and the solid line represents the mean of the test data from each panel. While there is scatter, the mean of the data suggests that the strength initially increases with coating thickness and than decreases with coating thickness. Such that a maximum or optimal strength occurs at approximately 68% to 75% volume fraction. Using the analysis of Carman et al. [3], assuming that the thermal stresses arise due to a 150 degree C cool down, and assuming that the strength of the host material 10 ksi (assumed strength of 3501-6 resin), an optimal coating thickness is predicted to be 70%. This corresponds reasonably well with the experimental data. However, once again we point out there appears to be considerable fluctuations in the data.

4.2 EFPI failures in manufacturing

The egress point of the EFPI was protected with teflon shrink tube during the manufacturing process to reduce the possibility of failure. While no failure occurred at this location, epoxy resin did wick down the interior of the teflon tube and coated the single mode portion of the wave guide external to the composite. The epoxy embrittled this portion of the optical fiber, that is, the coating is brittle and can cause the optical fiber to break if not handled with extreme care. Future tests should consider using a plugged teflon tube to prevent the epoxy from wicking into the tube.

Of the 5 EFPI sensors embedded in the cross-ply panel, only one displayed any visible signs of failure following the manufacturing process as evidence by x-ray imaging. This sensor, failed in the region near the hollow core/single mode junction (see Figure 5). This junction is a potential site for stress concentrations if any stresses are induced transverse to the optical fiber during the curing process. It appears that resin flowed in this direction (i.e. parallel to the 90 degree fibers) and precipitated this failure. This type of failure has been reported by other authors Iesko et al.[12] who also attributed it to resin flow during the curing process. This type of failure can be eliminated if proper care is undertaken during the manufacture of the sensor itself, for example, if this region is coated with a

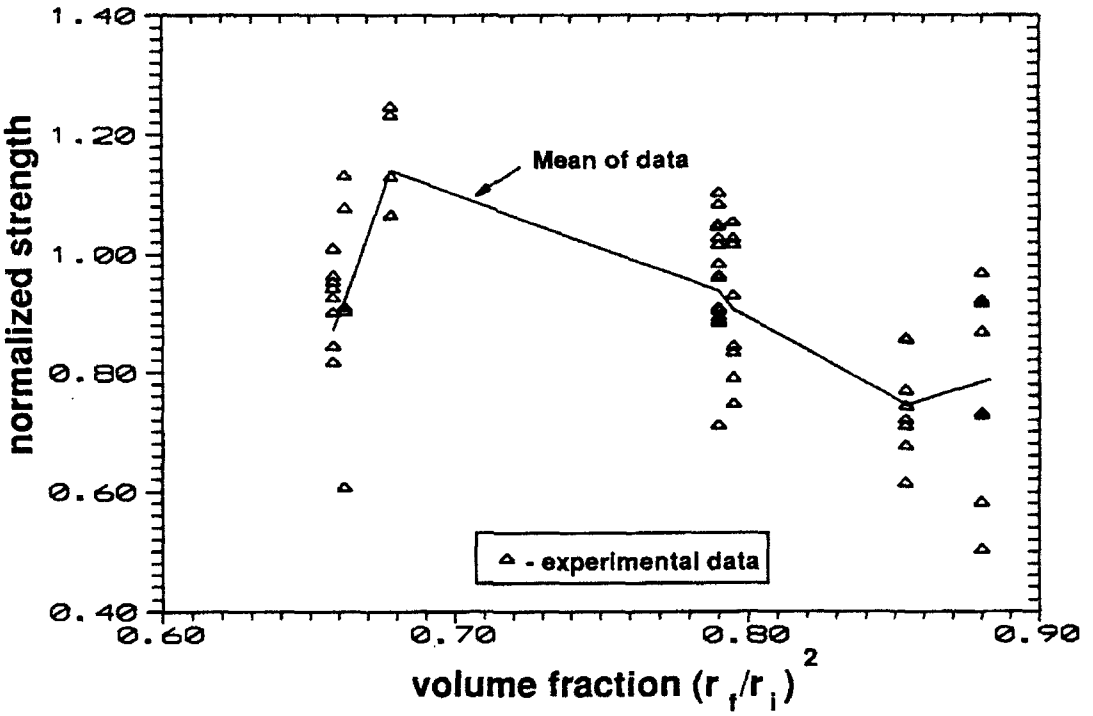


Figure 4: Normalized transverse strength data as a function of coating thickness for the polyimide coated fibers.

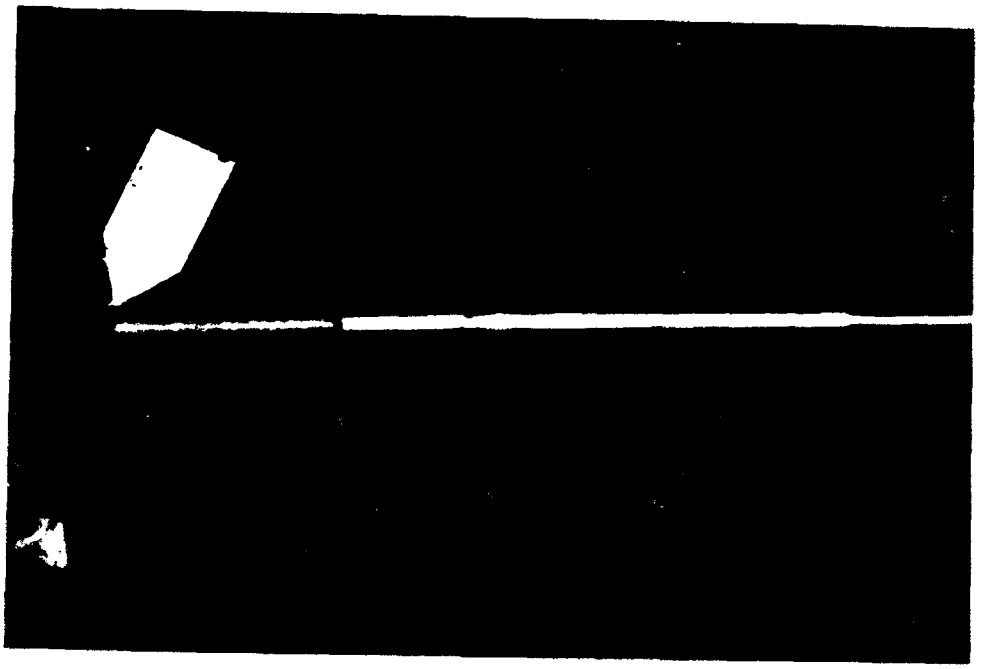


Figure 5: X-ray image of an EFPI sensor embedded in a composite material. The EFPI sensor is fractured at the single mode/hollow core junction.

polyimide to reduce stress concentrations.

The cross-ply panel discussed above was manufactured 4 weeks earlier than the quasi-isotropic panel. Therefore, the EFPI sensors were exposed to room environments for an additional 4 weeks as compared to the cross-ply specimens. During this time period the sensors were not placed in a desiccator. Following the construction and cure of the quasi-isotropic panels, review of x-ray images indicated that 4 of the 5 EFPI sensors had failed. All of these failures occurred at the fusion splice between the hollow core fiber and the multi-mode optical fiber (see Figure 6). This failure is attributed to residual stresses generated by thermal gradients set up during the fusion splice and moisture absorption of the silica based fiber. This conclusion is based on available literature [13 & 14] and the fact that none of the cross-ply specimens manufactured 4 weeks earlier displayed this failure pattern. Sensor failure occurred during either a resin flow cycle or the cool down cycle in the manufacture of the quasi-isotropic laminate. This failure can be eliminated if proper care is undertaken during the manufacturing process of the EFPI sensors or the sensors are properly desiccated.

4.3 EFPI static testing

Of the 5 sensors which survived the manufacturing process (1 quasi-isotropic and 4 cross-ply), 3 EFPI sensors provided recognizable output. Of the remaining 2, no discernable data could be obtained. This may be attributed to a break in the exposed single mode fiber that became coated with epoxy and embrittled this region. Of the 3 operable sensors, the EFPI sensor embedded in the quasi-isotropic laminate gave data which differed from external strain gauges by a factor of two. This may have been caused by epoxy wicking into the hollow core portion of the sensor region and altering the gauge length. No data is provided in this report on this sensor. Of the remaining 2 sensors (both cross-ply), one was tested quasi-statically.

The cross-ply specimen was initially loaded at an applied load of 1000 lbs and was incremented in 500 lb increments up to 9500 lbs. In Figure 7, a comparison of the strain data obtained from the external resistance strain gauges and the EFPI sensor is provided. This represents the first loading cycle on the specimen, that is an applied load of 1000 lbs. The EFPI sensor is positioned in the composite such that the data obtained from it should correlate with the data obtained from resistance strain gauge 1. As one can see the discrepancies between the two are minimal. The gauge length of these EFPI sensor may have been inaccurately measured by as much as 10% (see Carman et al. [15]) and can easily account for this error. The reader should also be aware that strain 2 provides a lower strain reading for the same stress level, this is to be expected because it is located further from the hole.

The specimen was loaded at successively higher stress levels until 9500 lbs load was applied which corresponds to 85% ultimate failure strength of the cross-ply specimen. At this point the sensor became inoperable. The precise reason for failure of the sensor could not be ascertained from x-ray images. This is being investigated in more detail

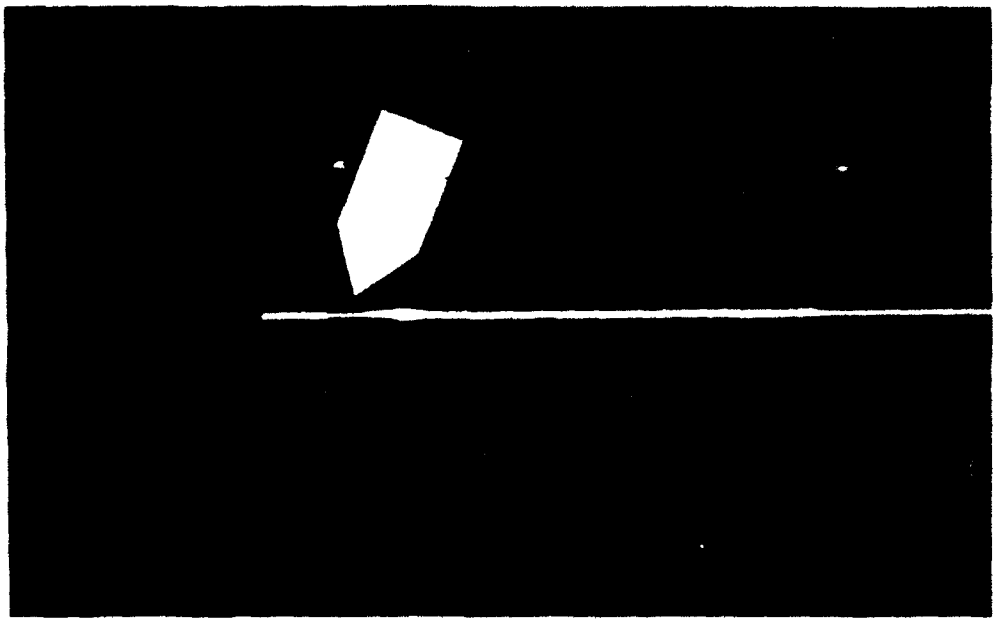


Figure 6: X-ray image of an EFPI sensor embedded in a composite material. The optical sensor is fractured at the fusion splice in the hollow core.

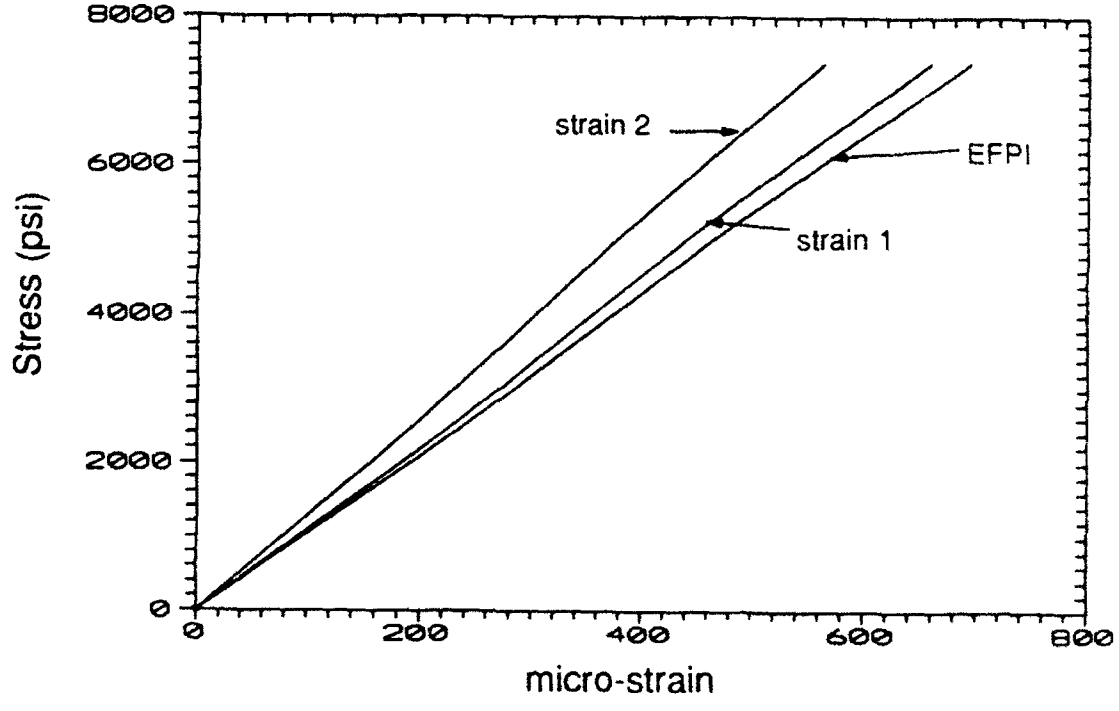


Figure 7: Comparison between the strain data obtained from the EFPI sensor and the resistance strain gauges for the initial quasi-static loading (1000 lbs).

and is discussed by Paul et al. A comparison of the strain data obtained from all the strain gauges on and in the specimen for the final loading profile is provided in Figure 8. As one can see the discrepancies between the EFPI sensor and resistance strain gauge 1 is more severe than in the first set of data provided in Figure 7.

In Figure 9 a comparison of the local stiffness values calculated from data obtained from the three gauges as a function of the applied stress level is presented. Local stiffness is defined to be slope of the stress-strain curved as determined by each respective gauge. As one can see, the dat gathered from each gauge suggests that as the load increases the local stiffness decreases. That is, damage evolution resulting in stress redistribution effects cause local regions to become more compliant. However, the EFPI sensor embedded in the middle plies suggests a more severe degradation in local stiffness than do the resistance strain gauges attached to the external plies.

The apparent difference in the rate local stiffness changes as a function of the applied which exists between the EFPI sensor and strain gauge 1 may be attributed to one of two (or a combination) damage phenomena. First, local damage may be evolving such that the center plies (i.e. the ones with the EFPI) are supporting more of the stress and thus exhibit a larger strain. On the other hand, the EFPI sensor itself may becoming damaged in a manner that its gauge length is changing. This latter statement includes the possibility that adhesion affects between the EFPI sensor and the composite play a role. However, no distinct conclusion can be drawn concerning which phenomenon resulted in the apparent disagreement in stiffness values. A more detailed review of damage in the local area is given by Paul et al.[6].

4.4 EFPI fatigue testing

The remaining cross-ply specimen containing an operable EFPI sensor was subjected to cyclic loading at 5 hertz at R=0.1. The initial peak load applied to the composite was 3000 lbs for 100,000 cycles followed by 4500 lbs for 100,000 cycles. The final load was 6000 lbs for 5000 cycles at which time the EFPI sensor displayed anomalous output.

A comparison between the initial data gathered from the EFPI sensor and resistance strain gauge 1 was similar to that depicted in Figure 7, at least for the first 100 cycles. The discrepancies between strain gauge 1 and the EFPI sensor were less than 5%. However, as the test proceeded this discrepancy became larger as described in the preceding section. In Figure 10 a comparison of the local stiffness change as a function of cycles is presented for the resistance strain gauges and the EFPI sensor. In the first few hundred cycles the local stiffness of the composite decreased by 10% as measured with the EFPI sensor. However, the two resistance strain gauges suggest an increase in stiffness which is not plausible. Apparently the adhesive utilized to mount the strain gauges onto the composite began to deteriorate even after a few hundred few cycles. Nonetheless, the EFPI sensor detected a measurable decrease in local stiffness change as a function of remaining life [16]. This decrease is expected, that is as damage

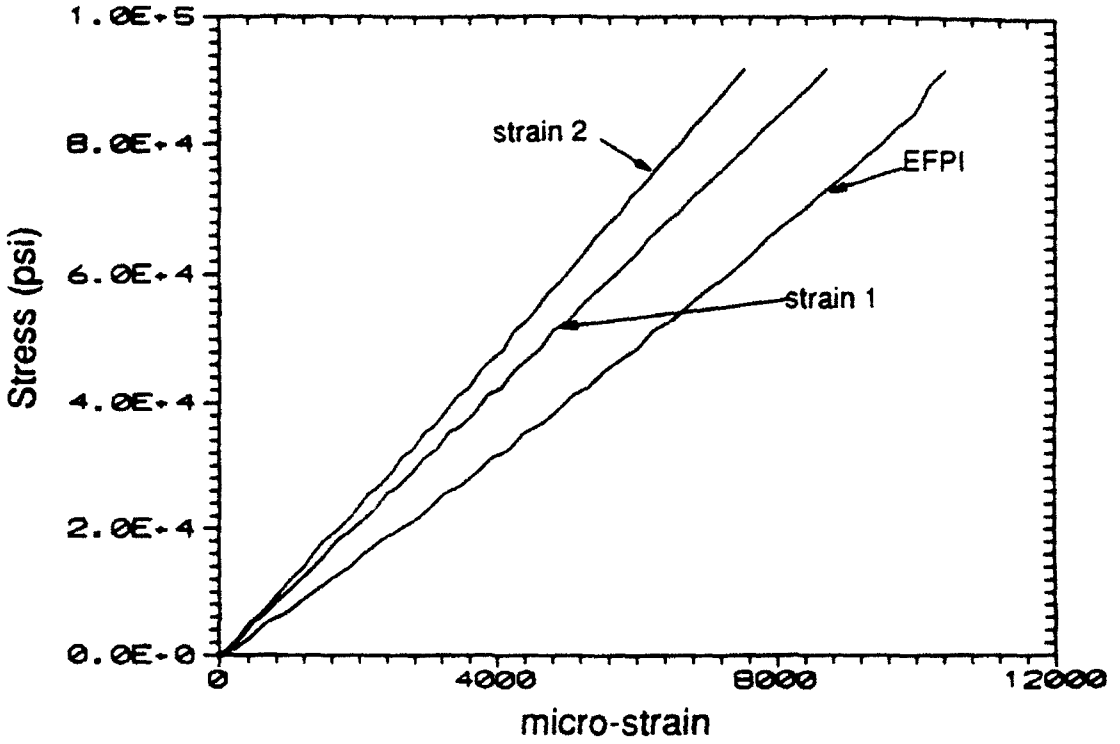


Figure 8: Comparison between the strain data obtained from the three sensors for the final quasi-static loading (9500 lbs).

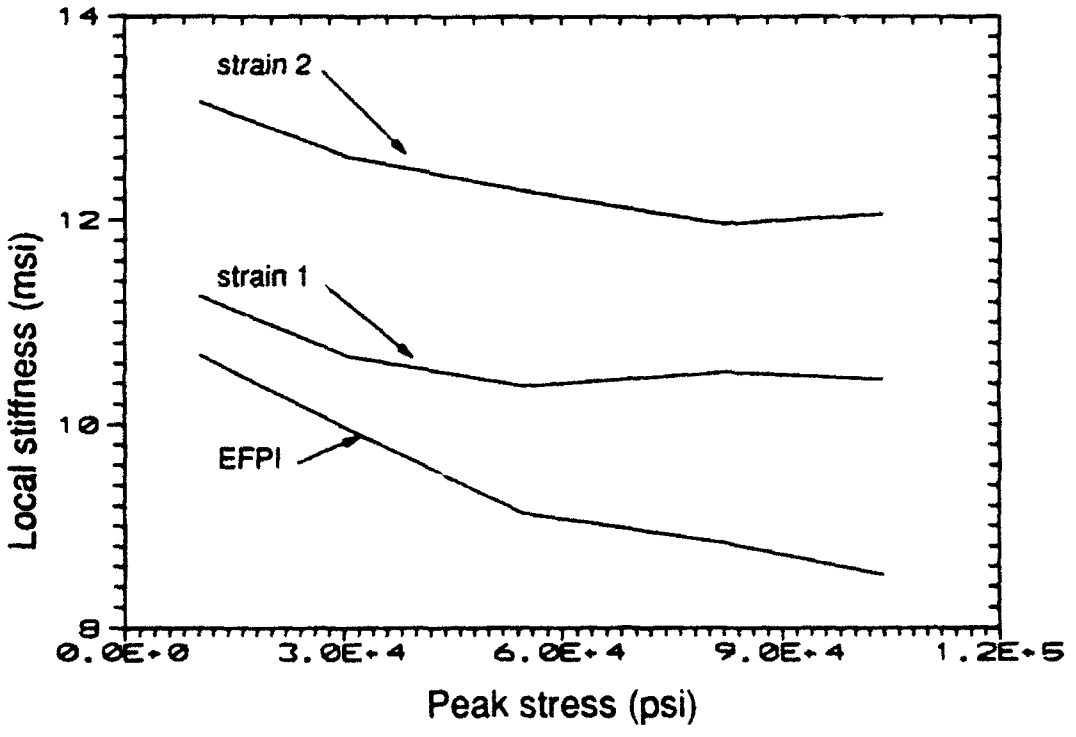


Figure 9: Comparison of local stiffness values calculated from the three strain gauges as a function of applied peak stress.

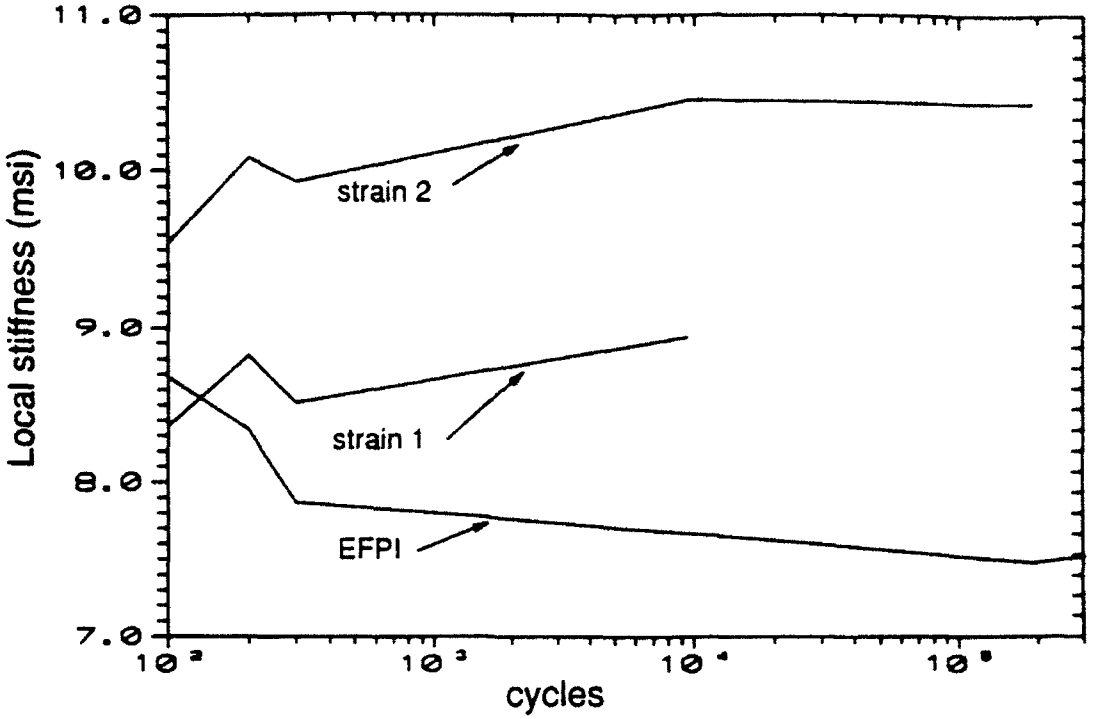


Figure 10: Comparison of local stiffness values obtained from the three strain gauges as a function of the number of cycles applied to the specimen.

progresses in the laminate in the form of cross-ply cracking, the local stiffness decreases and can be related to the residual strength of the specimen. After 10,000 cycles, resistance strain gauge 1 became inoperable and after 200,000 cycles resistance strain gauge 2 became inoperable. Both of these failures are associated with the gauges fatigue life. The EFPI sensor did not begin to show signs of degradation until after 300,000 cycles were applied to the composite.

At 305,000 cycles, the fringe pattern generated by the output from the EFPI sensor underwent a severe reduction in intensity (almost a factor of 3). Furthermore, another fringe pattern appeared to be superimposed on the original fringe pattern. This second patten was stronger in intensity and had significantly fewer fringes associated with it. The secondary fringes may have arisen due to a break in the single mode section of the optical fiber embedded in the composite. This would cause the output intensity of the EFPI sensor to degrade or even disappear entirely. Furthermore, if the fracture surface was "clean" a pseudo EFPI sensor may have been created. This pseudo sensor would have a smaller gauge length which would result in fewer fringes and would have a larger intensity associated with it, both seen in this experiment. At this writing the phenomenon is still being evaluated with a detailed description provided by Paul et al. [6].

While the fatigue life of this particular EFPI sensor is less than 305,000 cycles at the given applied strain, the

fatigue life of an EFPI sensor could be significantly improved if appropriate care was taken during the manufacturing process of the sensor itself. That is, these sensors are presently being crudely constructed as compared to the construction of high quality resistance strain gauges. Therefore, the fatigue lives and strengths of the EFPI sensors reported in this document should be considered a lower bound.

5.0 CONCLUSIONS

It was shown that embedded uncoated optical fibers can degrade the transverse strength of a laminate by as much as 50% and appears to be dependent upon fiber size. The use of polyimide coated optical fibers degraded the transverse strength of the laminate by less than 20%. It was suggested, both theoretically and experimentally, that an optimal polyimide coating thickness exists which will minimize the degradation of this composite property. The lack of adhesion between the optical fiber and the composite appeared to effect the uncoated fibers more significantly than the coated ones. This implies that uncoated optical fibers may exhibit better adhesion than coated ones.

In the second task, only 30% of the EFPI sensors were operable following the manufacturing process. Of these sensors, all had a strain to failure and a fatigue life smaller than the composite. This poor performance is attributed to the crude construction process utilized in the manufacture of these specific EFPI sensors. If manufactured properly this low survivability rate could be increased drastically. This could easily be accomplished with some quality control during the fabrication of these sensors. With such quality control it should be possible to manufacture EFPI sensors so that their strain to failures rivals that of an coated optical fiber which surpasses a composite material.

The initial data obtained from the operable EFPI sensors provided good correlation with the external strain gauges. This correlation was performed without a calibration procedure following the embeddment process. As larger loads were applied to the composite or a larger number of cycles, the stiffness reduction calculated from the data obtained from the EFPI sensor was more severe than that calculated from the resistance strain gauges. This may be attributed to local stress redistribution effects inherent to the composite or local damage associated with the EFPI sensor (e.g. adhesion problems). Nonetheless, the EFPI sensor outlasted the resistance strain gauges and measured a stiffness degradation. Future work should determine if EFPI sensors are a more viable fatigue gauge than classic resistance strain gauges.

6.0 ACKNOWLEDGEMENTS

The author of this report would like to acknowledge the assistance of Clare Paul in the Fatigue and Fracture Group at Wright Laboratory. All work presented in this document is a direct result of his hard and unrelenting effort to see this endeavor completed. Furthermore, this work could not of have been accomplished without the diligent work and timely response of the Flight Dynamics Directorate, the Structures Division, and the Structural Concepts Branch at Wright Lab and the Fiber & Electro Optics Research Center at Va. Tech.

7.0 REFERENCES

1. Carman, G.P. & Sendeckyi, G.P., "A Review of the Mechanics of Embedded Optical Sensors," to be submitted.

2. Roberts, S.S. & Davidson, R., "Mechanical Properties of Composite Materials Containing Embedded Fibre Optic Sensors," SPIE, V.1588, Fiber Optic Smart Structures & Skins IV, pp. 326-341, 1991.

3. Carman, G.P., Averill, R., Reifsnider, K.L., & Reddy, J.N., "Optimization of Fiber Coatings to Minimize Micromechanical Stress Concentrations in Composites," J. Comp. Mat., accepted Oct. 1992.

4. Case, S.W., Carman G.P. & Reifsnider, K.L., "Optimization of Transverse Performance Through Application of Appropriate Coatings," Ad. Comp. Lett., to appear Sept. 1992.

5. Carman, G.P., Paul, C., & Sendeckyi, G.P. "Transverse Strength of Composites Containing Optical Fibers," Smart Struct. & Mat Conf., Alb. N.M. Feb. 1993.

6. Paul, C., Carman G.P.,& Sendeckyi, G.P. "Monitoring the Fatigue behavior of Composites with Fabry-Perot Optical Sensors," Smart Struct. & Mat Conf., Alb. N.M. Feb. 1993.

7. Nath, D.K., Nelson, G.W., Griffin, S.E., Harrington, C.T., Yi He, Reinhart, L.J., Paine, D.C., Morse, T.F. "Polyamide Coated Embedded Optical Fiber Sensors," Spie, V. 1489, Structures Sensing & Control, 1991, pp. 17-32.

8. Urruti, E.H., Hawk, R.M., & Blaszyk, P.E., "Optical Fibers for Structural Sensing Applications," SPIE, V.986, Fiber Optic Smart Structures & Skins, 1988, pp.158-163.

9. Chang, C.C., "Low Velocity Impact of Laminated Graphite/Epoxy Panels with Embedded Optical Fibers," M.S. thesis, Univ. of MD., 1991.

10. Lesko, J.J., Carman, G.P., Fogg, B.R., Miller, W.V., Reifsnider, K.L., & Claus, R.O., "Embedded Fabry-Perot Optic Strain Sensors in the Macromodel Composite,", Optical Eng., V. 31 no.1, pp. 13-23, Jan 1992.

11. Roarke, D. of Ethyl Corporation personal communication.

12. Lesko, J.J., Case, S.W., Fogg, B.R., & Carman, G.P., "Embedded Fabry Perot Fiber Optic Strain Rosette Sensor for Internal Stress State Assessment," ASC 7th tech. conf. on Comp. Mat., Penn State Oct. 1992.

13. Skutnik, B., Munsey, B.D., & Brucker, C.T., "Coating Adhesion Effects On Fiber Strength and Fatigue Properties," Mat. Res. Soc. Symp. Proc., V. 88, 1987, pp. 27-34.

14. Wei, T., "Effects of Polymer Coatings on Strength and Fatigue Properties of Fused Silica Optical Fibers, "Advanced Ceramic Mat, V. 1, no. 3, 1986, pp.237-241.

15. Carman G.P., Case, S.W., Lesko, J.J., & Fogg, B.R., "Implementation and Evaluation of an Embedded Fabry-Perot Fiber Optic Strain Rosette Sensor for Internal Stress State Assessment," SPIE, Fiber Optic Smart Structures & Skins V, Boston, Sept. 1992.

16. Camponeschi, E.T. & Stinchcomb, W.W., "Stiffness Reduction as an Indicator of Damage in Graphite/Epoxy Laminates," ASTM Comp. Mat: testing & design 6th conf, STP 787, pp. 225-247, 1981

A STUDY OF FLIGHT DYNAMIC MODELING FOR NONLINEAR AERODYNAMIC PARAMETER ESTIMATION

Gary T. Chapman
Adjunct Professor
Department of Mechanical Engineering

University of California, Berkeley
6107 Etcheverry Hall
Berkeley , CA 94720

Final Report for:
Summer Research Program
Wright Laboratory
Armament Directorate

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D. C.

August 1992

A STUDY OF FLIGHT DYNAMIC MODELING FOR NONLINEAR AERODYNAMIC PARAMETER ESTIMATION

Gary T. Chapman
Adjunct Professor
Department of Mechanical Engineering
University of California, Berkeley

Abstract

A review of modeling of flight dynamic systems for nonlinear aerodynamic parameter estimation is presented. Modeling of the flight dynamic system is considered in a comprehensive manner: First, the modeling of a powered rigid body with controls is considered. Second, the aerodynamic modeling is considered in detail, including the effects of controls, propulsion and atmospheric winds. The assumptions involved in aerodynamic modeling are examined in detail. Third, the impact of elasticity of the modeling is considered including the effects on the aerodynamics. Work is required on aerodynamic modeling of high angle-of-attack and unsteady aerodynamics where an analysis of the assumptions shows a break down in conventional modeling. In addition work is required on handeling the effects of winds, elasticity and the propulsion system on aerodynamic parameter estimation. These latter issues of winds, elasticity and propulsion will also require more work on the measurements of these quantities, a topic not considered here.

A STUDY OF FLIGHT DYNAMIC MODELING FOR NONLINEAR AERODYNAMIC PARAMETER ESTIMATION

Gary T. Chapman

INTRODUCTION

The first theoretical basis for parameter estimation occurred with the first formal studies of flight mechanics and aerodynamic modeling by Lanchester and Bryan (Refs. 1 and 2). This early work concentrated on the study of linear systems. With the advent of modern control theory, as illustrated by Kalman's classical paper in 1960 (Ref. 3), work on estimation theory increased significantly, but it was mostly associated with state estimation. In the late 1960's and early 1970's with the advent of more powerful computers for both ground and flight and the demands of flight at high angles of attack, work on nonlinear estimation theory expanded greatly. One of the first conferences on nonlinear estimation theory occurred in the fall of 1970 (Ref. 4). Here again most of the effort had to do with state estimators, but some work on nonlinear aerodynamic parameter estimation was beginning to appear.

The determination of nonlinear aerodynamic information from ballistic ranges has been going on since the 1950's when average values of the drag coefficient or the wave lengths of the pitching motion were related to nonlinear polynomial expansions in amplitude (e.g., Refs. 5 and 6). The first algorithm for nonlinear aerodynamic parameter estimation in ballistic range testing was published in 1970 by Chapman and Kirk (Ref. 7). This work used polynomial representations of the aerodynamic functions. This approach has been extended to Splines by Sabot, Winchenbach and Chapman (Ref. 8) and to static aerodynamics with hysteresis (multi-valued functions) by Chapman (Ref. 9). Splines have also been used in parameter estimation procedures for flight testing (e.g., Ref. 10 and 11). Text books on parameter estimation for dynamic systems began to appear in the 1970's (e.g., Ref. 12), and since then work has continued (e.g., Ref. 13).

With the advent of nonlinear aerodynamic parameter estimation came the question as to what nonlinear representation is "best". For polynomial representations the questions become which terms are required or allowed and which can be estimated with statistical significance from the data. The first question has to do with fluid physics and the later with statistics. The first question has received very little attention in the literature on parameter estimation or model structure identification for flight and ballistic range testing. However, the study of aerodynamic mathematical modeling has existed as a separate subject for prescribing the types of wind tunnel testing required to determine a complete set of aerodynamic information, including unsteady aerodynamics. This work has been pioneered by M. Tobak and his colleagues at NASA-Ames Research Center (e.g., Ref. 14-17). The second question of identifying an aerodynamic model that is statistically significant has not received as much attention as nonlinear parameter estimation; however, some work has occurred in the last 10-15 years.

In reviewing the literature on the general subject of nonlinear aerodynamic model identification and parameter estimation, most attention seems to be directed towards the algorithm with little attention paid to the correctness of the model. When most algorithms are developed and used against simulated data that has been generated with the model employed in the algorithm, the resulting parameter estimates and comparisons with the simu-

lated observations are excellent, even in the presence of simulated measurement noise. However, when the same algorithm is applied to real measurements, the resulting comparisons with measurements are not nearly as good. This is illustrated in Fig. 1 with comparisons between measured and predicted longitudinal states. The difference between measured and predicted in these cases must be due to unmodeled effects. Incorrect or unmodeled thrust effects could account for the differences in the velocity, and unmodeled aerodynamics could account for the differences in the longitudinal and lateral state comparisons. The writers of Ref. 11 point out the aerodynamic modeling problem, which they illustrate by comparing the normal forces and pitching moments obtained from quasi-steady maneuvers and from large amplitude maneuvers. These are reproduced in Fig. 2. The quasi-steady maneuvers result in estimates that have hysteresis while the large amplitude maneuvers do not. This is a clear indication that the aerodynamic modeling is either incomplete or incorrect for the large amplitude maneuvers. This example would suggest that the major problems in model identification and parameter estimation may be insufficient or incorrect modeling.

This study will review the general topic of modeling for nonlinear aerodynamic parameter estimation. First, the overall area of mathematical modeling of flight dynamic systems will be presented, Second, the modeling of a powered, rigid body with controls will be considered. Third, the modeling of the aerodynamics will be considered in detail. Four, an elastic body will be considered with particular emphasis of the effects of the elasticity on the modeling.

FLIGHT DYNAMIC SYSTEMS MODELING

A vehicle in flight through the atmosphere can be modeled at several levels depending on whether one considers a rigid or an elastic body, a powered or an unpowered vehicle, or a vehicle that is in free flight or under controlled flight. In all cases the overall system is, in reality, an interaction between two or more systems. The simplest system is for an unpowered rigid body in free flight. This is the model normally employed in ballistic range data analysis. Here we have the interaction of the aerodynamic system and the inertial system of the rigid body. The overall system becomes more complex with the inclusion of elastic effects and a control system, and the addition of a propulsion system makes for an even more complex overall system. In the next section modeling of a powered rigid body with controls in flight will be considered. First, the inertial, control and propulsion systems will be described. Then the aerodynamic system and its interaction with these systems will be described in detail. A powered, aeroelastic body with controls will then be considered to illustrate the impact of the elasticity on the inertial terms and the aerodynamic model.

Powered, Rigid Body with Controls

This system is the interaction of four systems: an inertial system, a propulsion system, a control system and the aerodynamic system. The aerodynamic system will be considered in the next section. The following is a description of the modeling of the inertial, control and propulsion systems. The inertial system is modeled with Newton's laws applied to the motion of a rigid body:

$$m\,U' = -\,m\,(\omega \times U) + C_1 + F_a(t) + F_{th}(t) + F_g(R_0) \qquad (1a)$$

$$I\,\omega' = -\,\omega \times I\omega + C_2 + M_a(t) + M_{th}(t) + M_G(t) \qquad (1b)$$

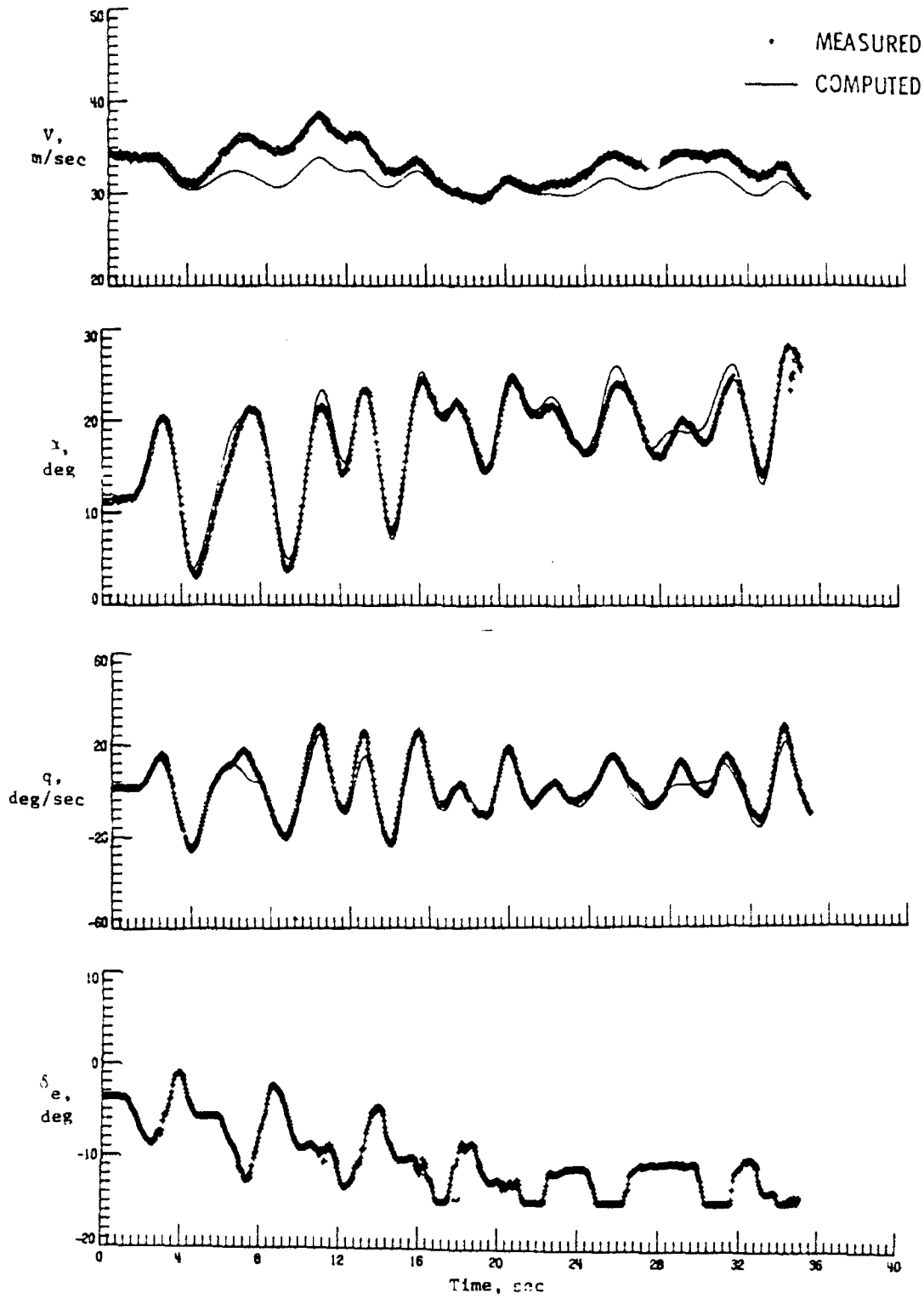


Fig. 1 Comparison of predicted and measured flight states, longitudinal motion states (Ref. 12).

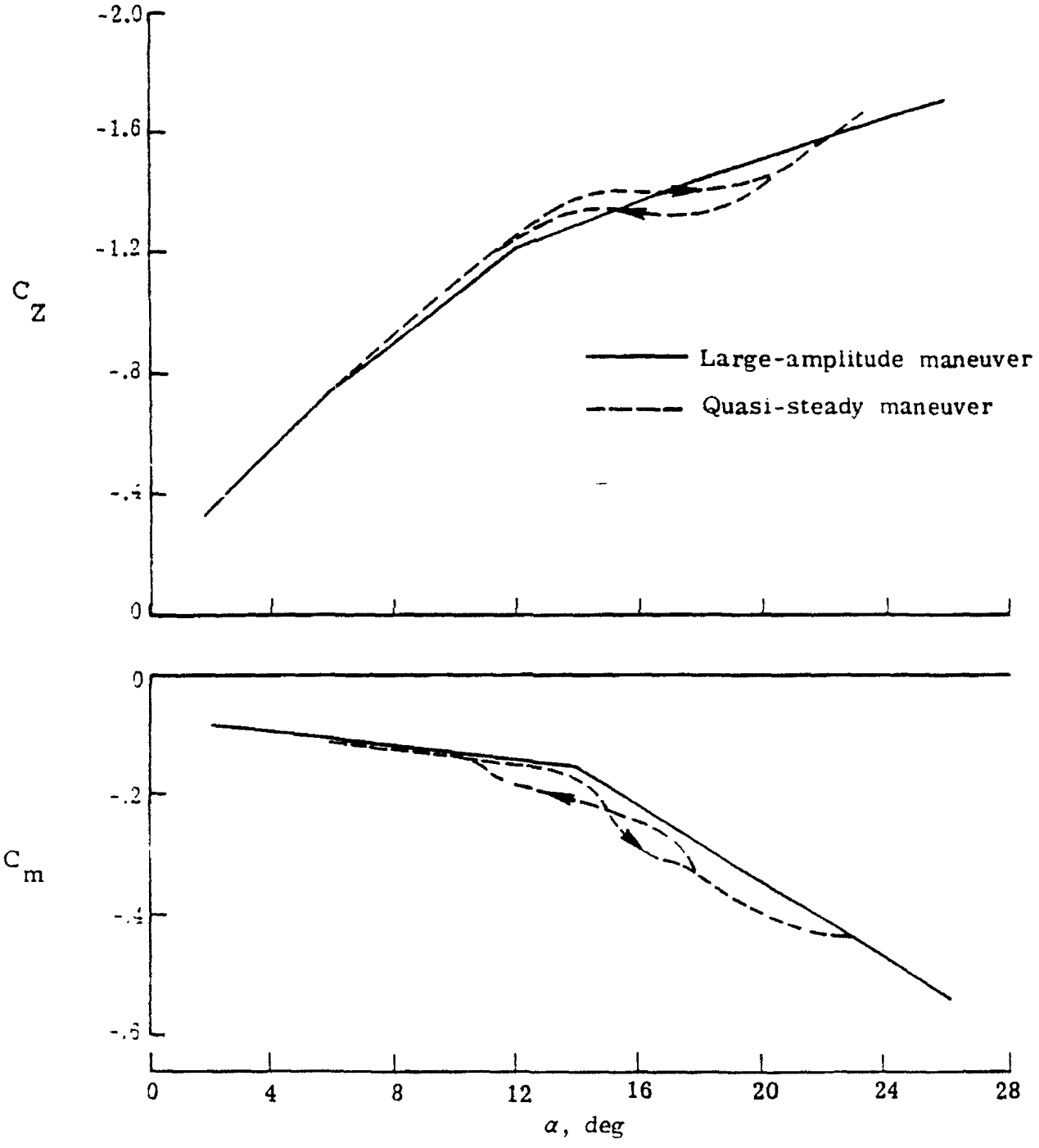


Fig. 2 **Comparison of normal force and pitching moment coefficients from quasi-steady maneuvers and large amplitude maneuvers (Ref. 12).**

Here $\mathbf{U}$ and ω are the linear and angular velocity vectors in the body coordinate system, the prime denotes partial differentiation with respect to time, t, m is the vehicle mass, $\mathbf{I}$ is the moment of inertia matrix of the vehicle (**changes to the mass and inertial properties due to fuel usage are small**), $\mathbf{F}_a(t)$ and $\mathbf{M}_a(t)$ are the aerodynamic force and moment vectors, $\mathbf{F}_{th}(t)$ and $\mathbf{M}_{th}(t)$ are the forces and moments due to the thrust of the propulsion system, $\mathbf{M}_G(t)$ is the moment introduced by gyroscopic effects of the propulsion system, and $\mathbf{F}_g(\mathbf{R}_0)$ is the force due to gravity which can be a function of $\mathbf{R}_0$, the location of the vehicle. The terms $\mathbf{C}_1$ and $\mathbf{C}_2$ are the effects due to rotating earth. For ballistic ranges the measurement coordinate system is fixed on a rotating earth, and the terms $\mathbf{C}_1$ and $\mathbf{C}_2$ must be included.

A full systems analysis of a control system would require another set of differential equations for the dynamics of that system including the actuators, the mass properties of the control surfaces, the aerodynamic moments on the control surfaces and the electronic portion of the control system. In parameter estimation work the control system is normally not modeled. It is simply replaced by a measurement of the control positions as functions of time, $\delta(t)$. This is an adequate and preferred approach for parameter estimation.

In parameter estimation the propulsion system, like the control system, is normally not modeled, but unlike the control system, it is very difficult to measure. Part of this difficulty arises from a long debated question among airframe and propulsion people, that is, what is accountable as drag and what is accountable as thrust. The thrust is most often treated as a known constant and the gyroscopic effects are often ignored. This is not adequate for obtaining good aerodynamic information, particularly in high angle-of-attack maneuvers where thrust levels and engine RPM may vary at a constant power setting. The engine RPM is normally measured, and it could be used to provide reasonably good estimates of the gyroscopic effects. As thrust vectoring is incorporated into newer aircraft the need for improved propulsion system modeling for parameter estimation will become imperative. Even a simple propulsion system model with a consistent accounting system for thrust and aerodynamics could provide some improvement over the constant thrust level approach. Thrust levels can also have a strong influence on the aerodynamic forces and moments. For example, the lift on a VSTOL aircraft, particularly those with over the wing blowing, is strongly influenced by the propulsion system. The extent of the impact of the propulsion system on aerodynamics can be seen in the amount of effort put into the simulation of inlet mass flow, exhaust jet and small scale propulsions systems in wind tunnel testing of small scale models and even entire, full scale, propulsion systems have been tested in the NASA-Ames 40 ft x 80 ft and 80 ft x 120 ft wind tunnels). The inclusion of the propulsion system effects in aerodynamic modeling will be considered in the next section.

Aerodynamic Modeling

Equation (1) as it stands is exact for a powered rigid body. However, in this form the terms $\mathbf{F}_a(t)$ and $\mathbf{M}_a(t)$ are unusable and must be modeled. For this modeling to be useful the aerodynamic force and moments need to be describable in terms of state variables, control positions, propulsion levels or other quantities which are derivable from flight information.

The fluid dynamic system around the vehicle is governed by Newton's laws applied to the air. This results in a set of partial differential equations called the Navier-Stokes equations (e.g , Ref. 18). These equations have

terms that depend on the thermodynamic properties, for example density. The solutions of these equations also depend on the boundary conditions at the vehicle surface, normally no flow through or no slip at the surface. The inclusion of a propulsion system requires different conditions at inlets and exhaust nozzles. These conditions are given by the time history of the vehicle motion, the control surface deflections, vehicle size, inlet mass flow and thrust levels. Conditions must also be specified far from the vehicle. These far field boundary conditions are where changes in the free stream thermodynamic properties enter the problem as well as the atmospheric winds. The Navier-Stokes equations are not appropriate modeling equations for the purposes of parameter estimation from flight. They represent an approach to determining aerodynamic information in their own right. However, whatever is done for modeling must be consistent with them, and they could prove useful in evaluating modeling ideas and for providing apriori parameter estimates, but that's a subject for another paper.

An approach to modeling a powered vehicle in controlled flight through a uniform atmosphere in the absence of winds that is consistent with the Navier-Stokes equations is to express the aerodynamic forces and moments as functionals of the time histories of the linear and rotational velocities of the vehicle, $\mathbf{U}(\xi)$ and $\omega(\xi)$ (both in a body axis coordinate system), the control surface deflections, $\delta(\xi)$, and the thrust levels, $\mathbf{T}(\xi)$. This modeling must also include functions of the size of the vehicle, L, and the thermodynamic and transport properties of the atmosphere, **P**. The inclusion of non uniform atmospheric properties and winds will be treated later in this section. With these ideas the aerodynamic forces and moments can be written as

$$\mathbf{F_a}(t) = \mathbf{F_a}[\mathbf{U}(\xi), \omega(\xi), \delta(\xi), \mathbf{T}(\xi); \mathbf{P}, L] \tag{2a}$$

$$\mathbf{M_a}(t) = \mathbf{M_a}[\mathbf{U}(\xi), \omega(\xi), \delta(\xi), \mathbf{T}(\xi); \mathbf{P}, L] \tag{2b}$$

where (ξ) represents the time history from 0 to time t.

There are at least two ways of reducing these functionals to functions, but before discussing them the various classes of aerodynamic behavior that can occur need to be considered. An understanding of these classes is essential to evaluating the assumptions that are utilized in reducing the functionals to usable functions. There are at least four different classes of aerodynamic behavior (Ref. 19) that are known to occur when a vehicle is held at some fixed velocity and orientation, for example as in a wind tunnel test. These are illustrated in Fig. 3. The first results in linear or nonlinear, single valued functions. This behavior is the one that most modeling procedures address. The second one is for behavior that results in static hysteresis and hence in multi-valued functions. An example is the multi-valued pitching moments, as a function of angle of attack, observed on a blunt-nosed cylinder with flare (Ref. 20). The third and fourth cases are for unsteady flows. The third is for periodic flows resulting from a Hopf bifurcation, for example the flow over a rigid cylinder in cross flow that results in a periodic oscillation in all of the aerodynamic terms (Ref. 21). This type of flow may also occur during the early stages of wing stall. In the fourth case the aerodynamic forces may be multiply periodic, quasi-periodic or even chaotic. The chaotic flow occurs with the onset of a strange attractor. An example of the latter type of flow is again the flow over a rigid cylinder in cross flow, but at Reynolds numbers where boundary-layer transition occurs near the separation point. This type of behavior may also occur during wing stall. Remembered that the behaviors described to this point are for conditions of steady flight at fixed angular orientations.

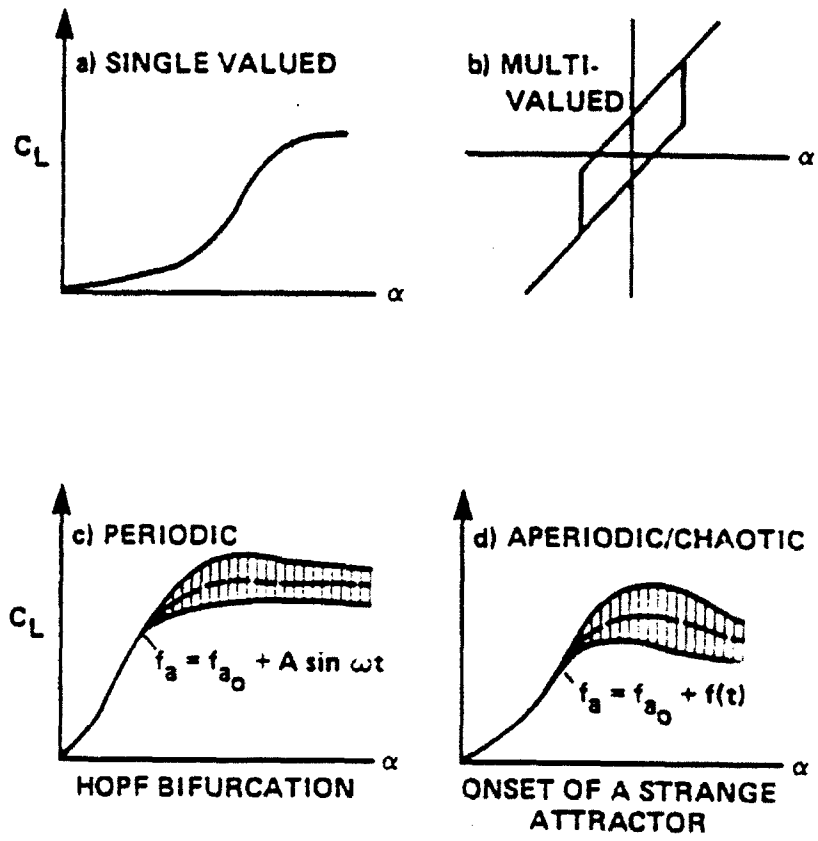


Fig. 3 **Basic types of aerodynamic forces with steady boundary conditions.**

When dynamic flight conditions occur the aerodynamic behavior becomes more complex (Ref. 22). In the first case considered previously, the dynamics of the vehicle normally lead to classical damping terms that are linear in rates (provided the rates are low) but can be nonlinear in angles. If the rates are high, more complex behavior can occur. When the aerodynamics are multi-valued, the effects of vehicle dynamics can be more extreme. When the aerodynamics are unsteady the effects of vehicle dynamics are much more complex; an example of this is the flow over an oscillating airfoil (e.g., Ref. 22). This is illustrated in Fig. 4 (Refs. 24-25). At the top of Fig. 4 are shown two flow fields over an oscillating airfoil at the same angle of attack, one during increasing and one during decreasing angles of attack as. As can been seen the flow structure is much different. The lower part of the Fig. shows the time averaged lift coefficient at several different oscillation frequencies. Also shown is the lift coefficient as a function of angle of attack for zero frequency. The lift under these dynamic conditions can not be captured by traditional modeling procedures, and it remains an unsolved modeling problem. See Refs 19 and 24 for a more detailed discussion of the different classes of aerodynamic behavior.

Two methods have been traditionally used to reduce the functionals, equations (2), to functions. These are the Taylor series expansion (e.g., Ref. 26) and the nonlinear indicial response approach (e.g., Ref. 14). The two methods lead to the same expressions for the functionals, but the indicial function approach can lead to more insight into the aerodynamics. Recently Truong and Tobak (Ref. 27) derived the indicial response approach directly from the Navier-Stokes equations under conditions that the flow over a vehicle at a constant speed and orientation in a uniform atmosphere was time invariant. Serious problems occur when these conditions yield unsteady flows. The implications of this will be explored later in this section. For purposes of this discussion the Taylor series expansion approach will be used. This will be illustrated using the moment vector, but the ideas apply to all of the aerodynamic forces and moments. Expansion of $\mathbf{U}(\xi)$ in a Taylor series about t yields $\mathbf{U}(\xi) = \mathbf{U}(t) + \mathbf{U}'(t)(\zeta - t) + \mathbf{U}''(t)(\zeta - t)^2/2 + \ldots$. Here ζ is a dummy time variable. Similar expansions are obtained for $\omega(\xi)$, $\delta(\xi)$ and $\mathbf{T}(\xi)$. With these expansions the expression for $\mathbf{M}_a(t)$ can be written as

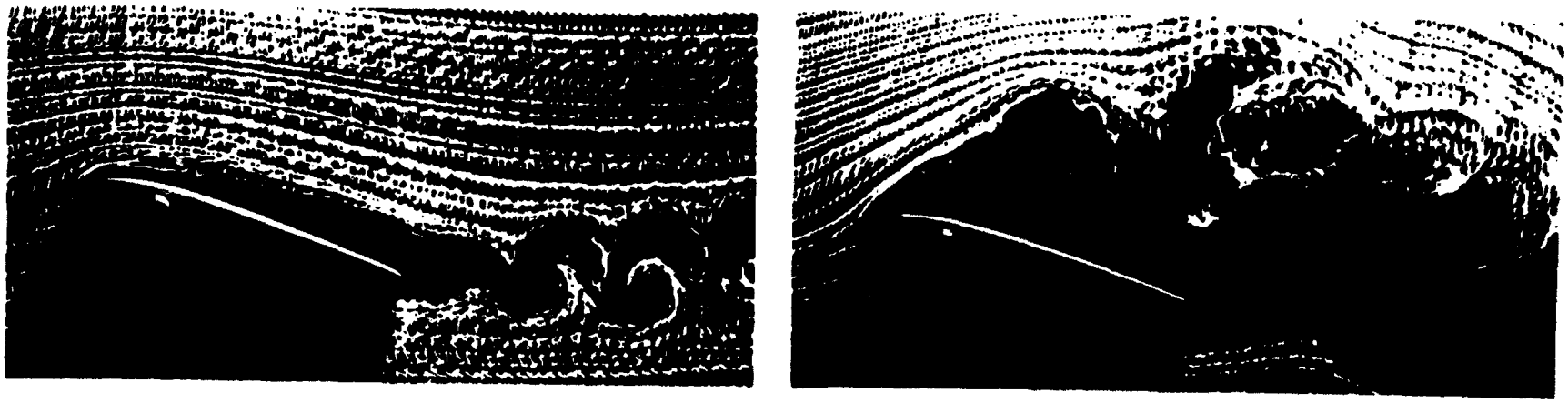
$$\mathbf{M}_a(t) = \mathbf{M}_a\{\mathbf{U}(t),\mathbf{U}'(t),\mathbf{U}''(t),\ldots,\omega(t),\omega'(t),\omega''(t),\ldots,\mathbf{T}(t),\mathbf{T}'(t),\mathbf{T}''(t),\ldots,\delta(t),\delta'(t),\delta''(t),\ldots;\ \mathbf{P},L\} \qquad (3)$$

It would appear that replacing a functional with a function of an infinite number of variables has not simplified the modeling situation. However, this form allows an order of magnitude analysis of the various terms. For periodic motions with frequency f involving the components of **U** and ω, one can readily show that the terms with increasing time derivatives are multiplied by increasing powers of frequency. Hence, if the reduced frequency, fL/V, is small compared to one, the higher order terms quickly become small (Ref. 12). If one further assumes that the rate of change with time is similarly small for δ and **T**, then the higher order terms in these quantities become small. With these assumptions Eq. (3) reduces to

$$\mathbf{M}_a(t) = \mathbf{M}\{\mathbf{U}(t),\mathbf{U}'(t),\omega(t),\delta(t),\delta'(t),\mathbf{T}(t),\mathbf{T}'(t);\ \mathbf{P},L\} \qquad (4)$$

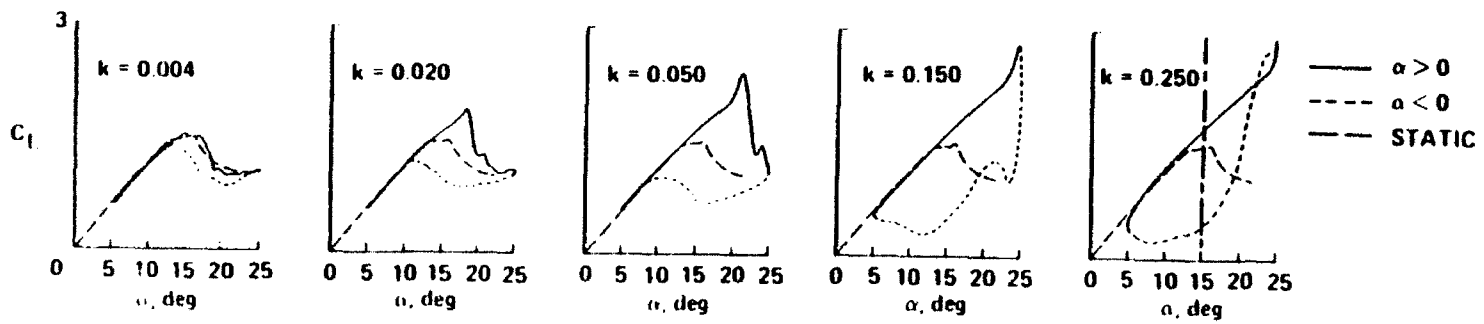
Note the subscript a has been dropped from the right side of this equation.

Up to this point the components of $\mathbf{U}(t)$ and $\omega(t)$ have been specified in body coordinates. The selection of a body coordinate system is not unique, but it is driven by the desire to have expressions for the aerodynamics that are both correct and expressed in the simplest way possible. The classical representation for nonlinear aerodynam-



$\alpha = 15°, \dot{\alpha} > 0, k = 0.25$ $\alpha = 10° + 10° \sin \omega t$ $\alpha = 15°, \dot{\alpha} < 0, k = 0.25$

a) McALISTER AND CARR



$\alpha = 15° + 10° \sin \omega t$

b) (McALISTER, CARR AND McCROSKEY)

Fig. 4 **Dynamic stall on an oscillating airfoil. Flow fields and lift coefficients.**

ics of an aircraft-like body is $U(V, v/V, w/V) = U(V,\alpha,\beta)$ and $\omega(p,q,r)$.

The expressions for forces and moments are normally not used in their dimensional forms, nor are the independent variables. Using dimensional analysis (e. g., Ref. 28), Eq. (4) can be written in the following form

$$M_a(t) = (QAL)C_M\{\underline{U},\underline{U}',\underline{\omega},\underline{\delta},\underline{\delta}',C_T,\underline{C}'_T;\underline{P}\} \tag{5}$$

where $Q = \rho V^2/2$ is the dynamic pressure, ρ is the air density, V is the magnitude of the resultant velocity, A is a reference area (a function of L), and $C_M\{\}$ is the vector of moment coefficients (C_l, C_m, C_n). The independent variables have been non-dimensionalized as follows $\underline{U} = \underline{U}(N_M,\alpha,\beta)$, $\underline{U}' = \underline{U}'L/V$, $\underline{\omega} = \omega L/V$, $\underline{\delta}' = \delta' L/V$, $C_T = T/QA$, $\underline{C}'_T = C'_T L/V$ and $\underline{P}$ is a set of non-dimensionalized thermodynamic parameters. Note the Mach number $N_M = V/c$ is represented in by the velocity vector $\underline{U}$.

The last step in this process is to expand the components of $\underline{C}_M$ about $\underline{\alpha}' = \underline{\beta}' = \underline{\delta}' = \underline{p} = \underline{q} = \underline{r} = 0$, $N_M = N_{M0}$ and $\underline{C}_T = \underline{C}_{T0}$. The variations with **P** have been assumed to be negligible, hence only the nominal value of $\mathbf{P} = \mathbf{P}_0$ need be retained. This is all consistent with the approximations made so far and in addition the $\underline{C}'_T$ term has been dropped. For the pitching moment, C_m, this yields

$$C_m = C_m(\alpha,\beta,\delta) + C_{m\underline{\alpha}'}(\alpha,\beta,\delta)\underline{\alpha}' + C_{m\underline{\beta}'}(\alpha,\beta,\delta)\underline{\beta}' + C_{m\underline{\delta}'}(\alpha,\beta,\delta)\underline{\delta}' + C_{m\underline{p}}(\alpha,\beta,\delta)\underline{p} + C_{m\underline{q}}(\alpha,\beta,\delta)\underline{q}$$
$$+ C_{m\underline{r}}(\alpha,\beta,\delta)\underline{r} + C_{mN_M}(\alpha,\beta,\delta)(N_M - N_{M0}) + C_{mC_T}(\alpha,\beta,\delta)(\underline{C}_T - \underline{C}_{T0}) \tag{6}$$

Similar expressions are obtained for C_l, C_n and the aerodynamic forces. Nonlinear dependence on α, β and δ has been retained in all coefficients and all rate terms are linear. For simplicity the dependence on N_{M0}, C_{T0} and $\underline{P}_0$ are not shown . This is a classical representation with the exception of the inclusion of the controls rate term $C_{m\underline{\delta}'}()$, the Mach number term $C_{mN_M}()$ and the thrust term $C_{mC_T}()$. These three terms could be as important as some of the other rate terms, particularly in the transonic speed range. Ballistic range parameter estimation procedures often include a dimensional form of the Mach number term, $C_{mV}()$.

Another approximation used extensively is that of small vertical and lateral accelerations of the center of gravity. If we write the angular rotation rates $\underline{q}$ and $\underline{r}$ as $\underline{q} = \underline{\alpha}' + v'_E L/V^2$ and $\underline{r} = \underline{\beta}' + w'_E L/V^2$ (note, v_E and w_E are in earth fixed coordinates) then for small lateral and vertical accelerations the second term in both of these expressions is small compared to the first and hence $\underline{q} \cong \underline{\alpha}'$ and $\underline{r} \cong \underline{\beta}'$. With these approximations Eq. (7) can be written as

$$C_m = C_m(\alpha,\beta,\delta) + \{C_{m\underline{q}}(\alpha,\beta,\delta) + C_{m\underline{\alpha}'}(\alpha,\beta,\delta)\}\underline{q} + \{C_{m\underline{r}}(\alpha,\beta,\delta) + C_{m\underline{\beta}'}(\alpha,\beta,\delta)\}\underline{r} + C_{m\underline{p}}(\alpha,\beta,\delta)\underline{p}$$
$$+ C_{m\underline{\delta}'}(\alpha,\beta,\delta)\underline{\delta}' + C_{mN_M}(\alpha,\beta,\delta)(N_M - N_{M0}) + C_{mC_T}(\alpha,\beta,\delta)(C_T - C_{T0}) \tag{7}$$

This is a classical representation for trajectories with small vertical and lateral motion of the center of gravity, with the exception of the last two terms. With the combining of the terms $\{C_{m\underline{q}}() + C_{m\underline{\alpha}'}()\}$ and $\{C_{m\underline{r}}() + C_{m\underline{\beta}'}()\}$ it has become customary in the literature to further truncate these, retaining only the first term in each expression. This is normally done in the interest of simplifying the nomenclature. This it does, but it is beginning to lead to improper understanding of the terms and it hides the approximation which allows them to be combined.

Before going on to other issues the assumptions and/or approximations used to arrive at Eq. (7) need to be reviewed and evaluated in terms of the classes of aerodynamic behavior described earlier in this section. **These assumption/approximations are: (1) The vehicle is a rigid body. (2) The atmospheric properties are uniform**

and constant and there are no atmospheric winds. (3) The existence of steady fluid dynamics/aerodynamics under conditions of flight at constant velocity and orientation is assumed. (4) The higher order time derivatives in the functional expansions are negligible (dropped between Eqs. (3) and (4)). (5) All of the rate terms in α', β', δ', p, q and r, are expanded about zero and only first order terms in these rates are retained, (between Eq. (5) and (6)). And (6) small vertical and lateral accelerations are assumed leading to the combining of rate terms.

The rigid body assumption (1) will be discussed in the next section on the flight of a powered, elastic body with controls. The inclusion of non uniform and time varying atmospheric properties, $P(R_0,t)$, and wind velocity, $U_{aw}(R_0,t)$ (assumption (2)) is complex. In a strict sense, the aerodynamic forces and moments depend on the entire spatial and time variation of P and U_{aw} up until time t. A full treatment of this problem is neither possible nor appropriate here. However, this problem can be simplified. This simplification can be done because the time variation in atmospheric properties and winds at a point in space are normally small compared to the changes in P and U_{aw} due to the motion of the vehicle through a spatially non uniform, but time invariant, atmosphere. Therefore the explicit dependence on time can be neglected. Next we expand the atmospheric properties and winds in a Taylor series in space about the point, R_0, the position of the vehicle at time, t.

$$P(R_0) = P(R_0(t)) + \text{Grad } P(R_0(t))(R - R_0(t)) + ...$$

$$U_{aw}(R_0) = U_{aw}(R_0(t)) + \text{Grad } U_{aw}(R_0(t))(R - R_0(t)) + ...$$

The expressions $P(R_0(t))$, $U_{aw}(R_0(t))$, $\text{Grad}P(R_0(t))$ and $\text{Grad } U_{aw}(R_0(t))$ are now treated in the same manner as the other time history variables in Eq. (2) resulting in

$$M_a(t) = M[U(\xi),\omega(\xi),\delta(\xi),T(\xi),P(\xi), U_{aw}(\xi),\text{Grad}P(\xi),\text{Grad}U_{aw}(\xi); L] \qquad (8)$$

Note, the dependence on $R_0(t)$ has been replaced by $R_0(\xi)$ and then R_0 dropped. A similar expression can be written for $F_a(t)$ If the steps from Eq. (2) to (4) are employed along with the dropping of the time derivatives of the two Grad terms, Eq. (4) becomes

$$M_a(t) = M[U(t),U'(t),\omega(t),\delta(t),\delta'(t),T(t),T'(t),P(t),P'(t),U_{aw}(t),U'_{aw}(t),\text{Grad}P(t),\text{Grad}U_{aw}(t); L] \qquad (9)$$

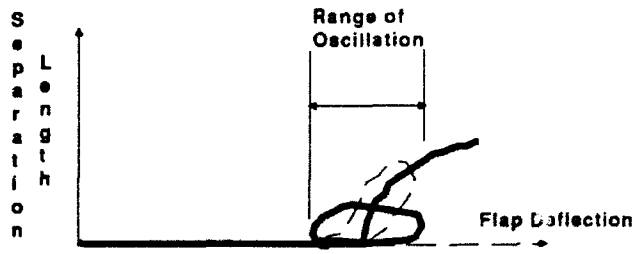
This expression could be non-dimensionalized to arrive at the equivalent to Eq. 5; however, for our purpose here this is not required. It will be sufficient to discuss how these terms impact the results. The inclusion of the $P(t)$ term allows all of the parameters, such as ρ, μ, and c in Eq. (5), to be treated as spatial variables, and hence the instantaneous local values can be used. The local wind term, $U_{aw}(t)$, could be added directly to the velocity of the vehicle, $U(t)$, provided the wind components are put in the appropriate body coordinate representation and appropriately normalized. This is the traditional way of handling the atmospheric wind terms. Note, this can not be done for the time rate of change of these two terms, $U'(t)$ and $U'_{aw}(t)$. This is because the acceleration of a body through a uniform air mass does not produce the same effect as the steady motion of a vehicle through a variable air mass.

The two gradient terms in Eq. (9) introduce first order effects of the spatial variation. The gradient of the properties, $\text{Grad}P(t)$, is a very small quantity and its effect would be expected to be small. This may not be the case for the gradient of the winds. It is easier to think about the wind gradient term if it is considered as the sum of

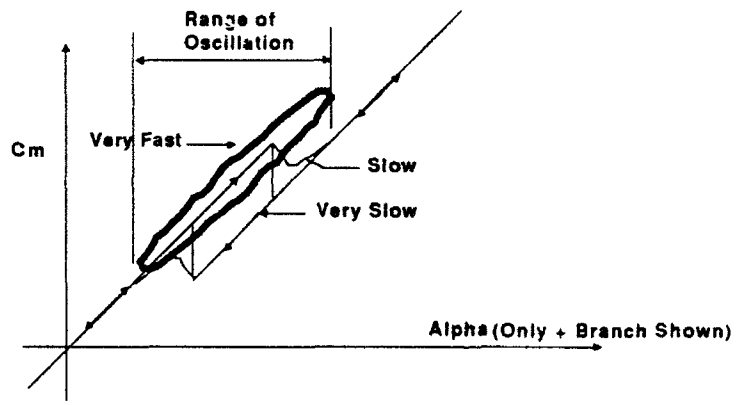
a symmetric and anti-symmetric matrix. These matrices represent the deformation and rotation of the local atmosphere. The interpretation of the deformation term is still difficult. However, the rotation term is similar to the $\omega(t)$ term. Here again care must be exercised. The effect of a body rotating in a uniform atmosphere is not the same as a rotating atmosphere on a body moving at constant speed and orientation, even if the rotation rates are the same. Another way of saying this is that a vehicle in a constant radius turn, at constant speed and orientation, does not experience the same aerodynamic forces and moments as a vehicle at constant speed flying in a curved airflow. Until more information can be obtained about the winds and variations in atmospheric properties during testing, it is not worth utilizing them in the aerodynamic model. Hence, for flight testing the final results may be that we always have unmodeled effects that can contribute to significant corruption of the aerodynamic data. Work is desperately needed on this problem.

Assumption (3) concerning the time invariance of the flow over vehicles at constant speed and orientation can be very limiting. For example consider the periodic shedding of vortices from a stationary circular cylinder in cross flow. The existence of the periodic or, in other cases, more complex unsteady flows under these stationary conditions results in a very slow loss of the effects of the initial conditions. Hence, the retention of only the first order terms in the expansions of the time history terms is not adequate. One way of viewing this is that the phase relationship between the fluid motion and the vehicle motion is rooted in the initial conditions of the problem. Hence, modeling that only considers instantaneous values of orientations and rates is not sufficient. It would appear that the entire past history may be required. There have been attempts to model this behavior with differential equations (Ref. 29). If modeling aerodynamics of this type is important, more research will be required.

The truncation of the time derivative terms to first order, assumption (4), and the expansion of the rate terms about zero to first order, approximation (5), are consistent approximations for nondimensionalized rate terms that are much less than one. Both of these assumptions assume that the expansion coefficients to all orders are of the same order of magnitude. This may not be true under some of the circumstances illustrated in Fig. (3). These circumstances occur when the dynamics of the vehicle result in motion that takes the fluid dynamic/aerodynamic system near or through one of three conditions. These are: (a) a supercritical bifurcation, (b) a sub critical bifurcation, or (c) a fold resulting from a sub critical bifurcation in an unexcited variable. All three of these can cause problems for linear rate dependence, even when the rates may be small relative to characteristic times of the body motion. An example of a supercritical bifurcation is the separation of the flow over a flap at supersonic Mach numbers (Ref. 30). Here, if we make a plot of separation length as a function of flap deflection angle, as measured under conditions where the flap is moved very slowly, we get a curve very similar to the supercritical bifurcation diagram (Fig. 5a). Under more rapid dynamic conditions where the flap angle is increased across the point of onset of separation and back out, one might get a curve like the dashed curve in Fig. 5a. Because the time scales of the separated flow are much slower than those of the attached flow, which are similar to those for the overall vehicle, one might expect the flow lag to be so great as to cause the flow to remain attached throughout the motion under some conditions, as indicated by the dotted curve. This behavior is obviously not linear in rates. The more extreme cases of this behavior result . sub critical bifurcations and folds that lead to static hysteresis. For ex-



a) Dynamics of flow separation length--a supercritical bifurcation



b) Dynamics over a hysteresis in pitching moment resulting from a fold or a subcritical bifurcation

Fig. 5 Impact of dynamics on supercritical bifurcations and folds

ample, consider the pitching moment for a blunt-nosed cylinder flare at transonic speeds (Ref. 25). Here the static hysteresis in pitching moment is due to a static hysteresis of flow separation on the flare. A sketch of a hysteresis behavior of this type is indicated in Fig. 5b by the solid curve. Consider three motions with the same amplitude but increasing rates. A very slow motion would move along the solid curve, as the arrows indicate, with a very small linear contribution from the rate term. The static hysteresis loop can be treated with the modeling described to this point. A somewhat faster motion might produce the dashed curve, which still can be handled with linear rate terms. However, at a very high rate compared to attached flow time scales, the dotted curve may result. In this case separation never develops. This case can not be handled with a simple linear treatment of the rate terms. In some cases a swept delta has two stable trim points with roll angle. Jinkens and Myatt (Ref. 31) have found that classical modeling will not handle this under dynamic conditions.

There may also be some cases where the body rates are not small and extensions to the above need to be considered. The most common occurrence of this might be for spinning projectiles. Here the dependence on the spin rate, p, may be nonlinear. Other special cases may occur; they are normally associated with the body axis about which the moment of inertia is small.

The problem with the last assumption of small vertical and lateral acceleration, assumption (6), is that the pairs of terms that are brought together, for example $\{C_{m_q}(\alpha,\beta,\delta) + C_{m_{\dot{\alpha}}}(\alpha,\beta,\delta)\}$, come from completely different types of motion (e. g., Ref. 31). It just so happens they can not be distinguished when the vertical and lateral motions of the center of gravity of the vehicle are small. This is not necessarily true for certain maneuvers of VSTOL aircraft, high performance aircraft with lateral motion control independent of rotation, i. e., those with canard and or thrust vectoring, or possibly some high angle-of-attack maneuvers. Hence, the practice of replacing the pairs of terms with just the first term runs the risk of violating the assumption. It would appear that the differences are sufficiently large as to be detectable in the high angle-of-attack maneuver data used in Ref. 10 and illustrated in Fig. 1. A close examination of the α and q curves in Fig. 1, shows that $\dot{\alpha}$ and q can not be equated. The zeros in the q curve do not match up with the maximum in the α curves and vice-versa; similarly for the β and r curves. In these cases it appears that assumption (6) is violated, at least to a degree, in the high angle-of-attack maneuvers. The assumption should be used when appropriate but checked in each case.

Powered, Elastic Body with Controls

Most vehicles undergo some amount of elastic deformation during flight, particularly during high g maneuvers. An example of a body deforming in flight is shown in Fig. 6. Seen here are shadowgraph pictures of a very slender body in flight in the ballistic range at Wright Laboratory (WL) Armament Directorate at Eglin AFB in Florida. For the first 40 meters of flight, up to station 13, the model appears to be straight, and only a small amount of bending appears by station 31, a distance of nearly 100 meters down range, but from that point on the flexing increases rapidly. In this case damage to the fins, evident in the shadowgraphs, causes the body to spin up and go into resonance with the natural bending frequency of the body. The projectile actually bent beyond the elastic limit. Cases have been observed where the projectile breaks in half.

In considering the powered flight of an elastic body with controls the following assumptions will be made:

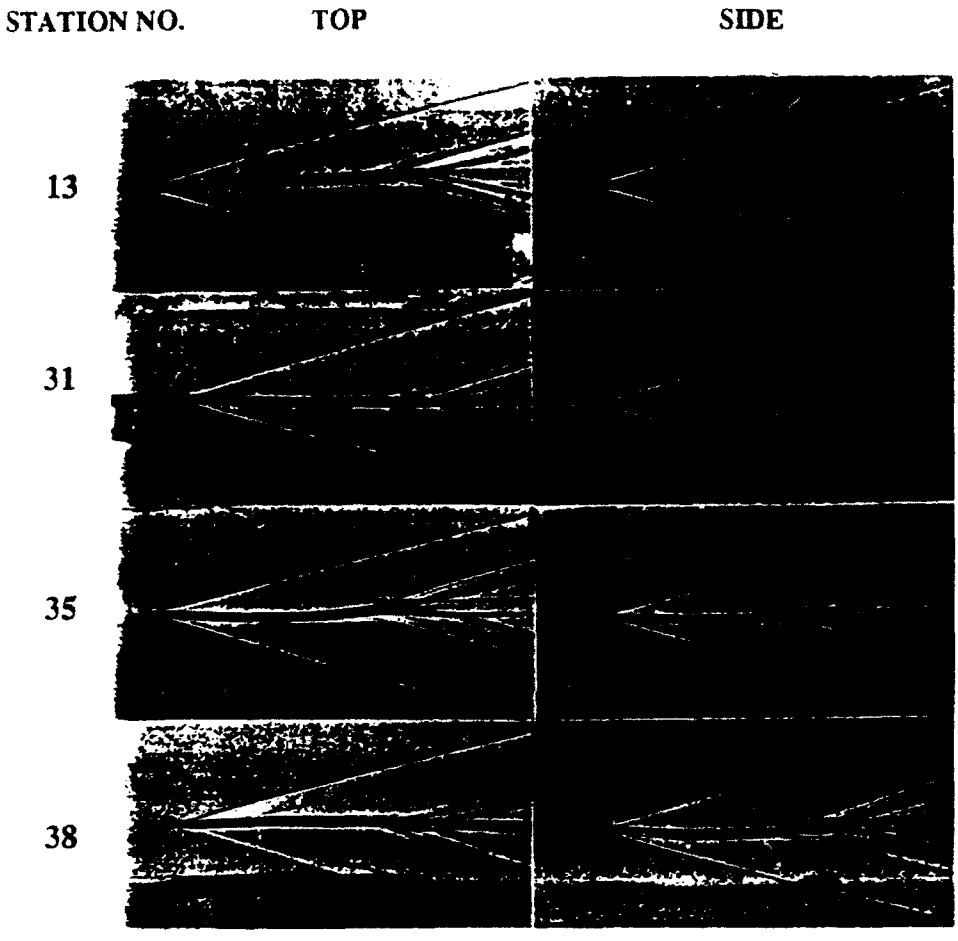


Fig. 6 Elastic body in flight through a ballistic range.

First, the atmosphere is uniform and there are no winds. Second, the control and propulsion systems do not depend significantly on the elasticity effects. Even under these conditions a comprehensive treatment of aeroelasticity is beyond this paper, s n a treatment can be found in Bisplinghoff and Ashley (Ref. 32). The equations for the aeroelastic motion for a very slender elastic body can be found in Ref. 33. To illustrate how elasticity enters into the problem, a modal representation approach will be used. In this approach the deformation of a body is represented as a sum over the eigen functions of the natural modes, called mode shapes, times an amplitude function for each mode. The deformation amplitudes are assumed small compared to the translational movement of the center of mass and the rotation about the center of mass.

The equations for translation of the center of mass is the same as Eq. (1a). In a general form it is

$$d/dt\iiint(d\mathbf{R}/dt)\,dm = m(\mathbf{U}' + \omega \times \mathbf{U}) = \mathbf{C}_1 + \mathbf{F}_a(t) + \mathbf{F}_{th}(t) + \mathbf{F}_g(\mathbf{R}_0) \qquad (10a)$$

Where $\mathbf{R}$ is the position of an elemental mass within the deformed body, $\mathbf{R} = \mathbf{R}_0 + s + \varepsilon$, $\mathbf{R}_0$ is the vector location of the center mass, $\mathbf{s}$ is the vector from the center of mass to the elemental mass in the undeformed vehicle, ε is the vector deformation, dm is an elemental mass at position $\mathbf{R}$ and d/dt is the total time derivative. This form is directly applicable to ballistic ranges where the earth base measurement system allows for measuring the location of the center of mass of the deformed vehicle. However, for flight testing where instruments are relative to the center of mass of the undeformed vehicle it might be more convent to write the equations relative to the location of the undeformed center of mass. If this is done Eq. (10a) becomes much more complex. An observer at the center of mass of the undeformed vehicle experiences accelerations due to the vibratory motion.

The equation for rotation about the center of mass of the deformed vehicle is given as

$$d/dt\iiint((s + \varepsilon) \times d(s + \varepsilon)/dt)\,dm = \mathbf{C}_2 + \mathbf{M}_a(t) + \mathbf{M}_{th}(t) + \mathbf{M}_G(t) \qquad (10b)$$

This expression becomes very messy when fully expanded in terms of the modal representation. One of the large effects is due to the changes in the moments of inertia as the vehicle deforms.

Using the modal representation the deformation ε is written as

$$\varepsilon(s,t) = \sum \mathbf{A}_i(s)\vartheta_i(t) \qquad (11)$$

Where the $\mathbf{A}_i(s)$s are the mode shapes, assumed known from either a free vibration analysis or from a vibration experiment, and the $\vartheta_i(t)$s are the amplitude functions which must be solved for as part of the dynamics. For small distortions the differential equations for the modal amplitudes are given as

$$m_i\,\vartheta_i(t)'' + m_i\,\lambda_i^2\vartheta_i(t) = E_i(t) \qquad (12)$$

where $m_i = \int |\mathbf{A}_i(s)|^2 dm$ is the generalized mass for the ith mode, λ_i is the natural frequency of the ith mode and $E_i(t) = \iint \mathbf{A}_i(s)\vartheta_i(t) \cdot d\mathbf{F}$ is the generalized aerodynamic force driving the ith mode, where $d\mathbf{F}$ is an elemental aerodynamic force vector at a position s on the body surface. In this form the modal equations, although as simple as possible, they are still complex. However, they would not capture the behavior exhibited in Fig. 9. In using the small displacement approximation the coupling between the spin and bending mode has been lost. To recapture the spin-bending coupling, Eq. (12) must be generalized and hence becomes even more complex. The only forcing of the modes in Eq. (12) comes from the aerodynamics which needs modeling to be useful in parameter estimation.

These generalized aerodynamic forces, like the regular aerodynamic forces and moments, depend on the

motion history and flow field parameters as well as on the deformation. To illustrate this we will consider the aerodynamic moment $\mathbf{M}_a(t)$ and the ith generalized aerodynamic force. These can be written as

$$\mathbf{M}_a(t) = \mathbf{M}_a[U(\xi),\omega(\xi),\delta(\xi),T(\xi),\vartheta_1(\xi),\vartheta_2(\xi),\ldots\vartheta_n(\xi);\mathbf{P},L]$$

$$E_i(t) = E_i[U(\xi),\omega(\xi),\delta(\xi),T(\xi),\vartheta_1(\xi),\vartheta_2(\xi),\ldots\vartheta_n(\xi);\mathbf{P},L] \qquad (13)$$

These expressions can be handled in the same manner as Eq. (2) in the derivation of Eq. (7). The additional terms due to the mode functions are given in the following equations

$$C_m = \text{Eq. (7)} + \Sigma C_{m\vartheta_i}(\alpha,\beta,\delta)\vartheta_i + \Sigma C_{m\dot{\vartheta}_i}(\alpha,\beta,\delta)\dot{\vartheta}_i$$

$$C_{Ej} = E_j/QA|A_j(s)| = C_{Ej}(\alpha,\beta,\delta) + \{C_{Ejq}(\alpha,\beta,\delta) + C_{Ej\dot{\alpha}}(\alpha,\beta,\delta)\}q + \{C_{Ejr}(\alpha,\beta,\delta) + C_{Ej\dot{\beta}}(\alpha,\beta,\delta)\}r + C_{Ejp}(\alpha,\beta,\delta)p + C_{Ej\dot{\delta}}(\alpha,\beta,\delta)\dot{\delta} + C_{EjNM}(\alpha,\beta,\delta)(N_M - N_{M0}) + C_{Ej\zeta_T}(\alpha,\beta,\delta)(\zeta_T - \zeta_{T0}) + \Sigma C_{Ej\vartheta_i}(\alpha,\beta,\delta)\vartheta_i + \Sigma C_{Ej\dot{\vartheta}_i}(\alpha,\beta,\delta)\dot{\vartheta}_i$$

As can be seen, the aerodynamic force and moment model becomes very complex when written for the general case. In most cases subsets of these can be used, but one must keep in mind that all these terms can impact the final results. Hence, when there are questions concerning the overall results from a parameter estimation procedure, one needs to examine the modeling to insure the appropriate terms have been retained or that an assumption has not been violated.

CONCLUDING REMARKS

A comprehensive treatment of modeling for nonlinear aerodynamic parameter estimation has been presented. The modeling of the flight dynamic system was considered. Modeling of the flight dynamics included the effects of aeroelasticity, atmospheric winds, propulsion, controls and aerodynamics. The modeling of nonlinear aerodynamics is the largest single source of discrepancy. Non uniform atmospheric conditions, aeroelasticity and propulsion effects can also give rise to uncertainties in the estimated aerodynamics. A more complete review of nonlinear parameter estimation and model structure determination was conducted during the AFOSR Summer Research Program of 1992 including a discussion of measurement systems and algorithms (Ref. 34).

REFERENCES

1. Lanchester, F. W., "Aerodenetics," A Constable and Co. Ltd., London, 1908.
2. Bryan, G. H., "Stability in Aviation," MacMillan and Co., London, 1911.
3. Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," ASME Trans., J. Basic Engineering, Series D, 82, 1960.
4. Proceedings of the "Symposium on Nonlinear Estimation Theory and Its Application," Western Periodicals Co. North Hollywood, Ca, Sept. 1970.
5. Sieff, A. and Wilkins, M. E., "Experimental Investigation of a Hypersonic Glider Configuration at Mach Number of 6 and Full Scale Reynolds Number," NASA TN D-341, 1961.
6. Rasmussen, M. L. and Kirk, D. B., "On the Pitching and Yawing Motion of a Spinning Projectile Governed by an Arbitrary Nonlinear Restoring Moment," NASA TN D-2135, 1964.
7. Chapman, G. T. and Kirk, D. B., "A Method for Extraction of Aerodynamic Coefficients from Free-Flight Data," AIAA J., Vol. 8, No. 4, April 1970.
8. Sabot, S. M., Winchenbach, G. L. and Chapman, G. T., "Comparison of Various Drag Coefficient Expansions Using Polynomials and Splines," J. of Spacecraft and Rockets, Vol. 23, No. 3, May-June 1986.
9. Chapman, G. T., "Ballistic Range Aerodynamic Data Reduction Using Polynomials and Continuous and Discontinuous Splines," Proceedings of the 38th Aeroballistic Range Association Meeting," Tokyo, Japan, Oct. 5-8, 1987.

10. Stalford, H. L., "The EBM System Identification Technique and Its Application to High α/β Modeling of Aircraft," Proceedings of the AIAA Atmospheric Flight Mechanics Conference, Danvers, Mass., Aug. 11-13, 1980, pp. 619-625.
11. Klein, V. and Batterson, J. G., "Determination of Airplane Model Structure From Flight Data Using Splines and Stepwise Regression," NASA TP 2126, March 1983.
12. Mendel, J. M., "Discrete Techniques of Parameter Estimation, The Equation Error Formulation," Control Theory Series, Vol. 1, Marcel Dekker, Inc., New York, 1973.
13. Billings, S. A., "Identification of Nonlinear Systems: A Survey," IEEE Proc., Vol. 127, Pt D, No. 6, 1980.
14. Tobak, M., "On the Use of the Indicial Function Concept in the Analysis of Unsteady Motions of Wings and Wing-Tail Combinations," NACA Rept. 1188, 1954.
15. Tobak, M. and Pearson, W. E., "A Study of Nonlinear Longitudinal Dynamic Stability," NASA R-209, Sept. 1964.
16. Tobak, M. and Schiff, L. B., "Aerodynamic Mathematical Modeling-Basic Concepts," AGARD Lecture Series No. 114 on Dynamic Stability Parameters, Lecture No. 1, March 1981.
17. Tobak, M., Chapman, G. T. and Unal, A., " Modeling Aerodynamic Discontinuities and the Onset of Chaos in Flight Dynamical Systems," Des Annales Des Telecommunications, Vol. 42, No. 5-6, 1987, pp. 300-314.
18. Schlichting, H., "Boundary-Layer Theory," MaGraw-Hill Publishing Comp., Seventh Ed., 1979.
19. Chapman, G. T. and Tobak, M., "Nonlinear Problems in Flight Dynamics," Proceedings of theBerkeley-Ames Conference on Nonlinear Problems in Controls and Fluid Dynamics, Math Sci Press, 1985.
20. Reese, D. E. and Wehrend, W. R., "An Investigation of the Static and Dynamic Aerodynamic Characteristics of a Series of Blunt Nosed Cylinder-Flare Models at Mach Numbers from 0.65 to 2.20," NASA TM X-110, 1960.
21. Bearman, P. W., "Vortex Shedding from Oscillating Bluff Bodies," Annual Review of Fluid Mechanics, Vol. 16, 1984, pp. 300-314.
22. Chapman, G. T., Yates, L. A. and Szady, M., "Atmospheric Flight Mechanics and Chaos--Some Issues of Modeling and Dimensionally," Proceedings of the Applications of Chaos Conference, Ed., Kim, J., Wiley, 1992.
23. McCrosky, W. J., "Unsteady Airfoils," Annual Review of Fluid Mechanics Vol. 14, 1982, pp. 285-311.
24. McAlister, K. W. and Carr, L. W., "Water-Tunnel Experiments on Oscillating Airfoils at Re = 12,000," NASA TM-78446, March 1978.
25. McAlister, K. W., Carr, L. W. and McCroskey, W. J., "Dynamic Stall Experiments on the NACA 0012 Airfoil," NASA TP-1100, Jan. 1978.
26. Etkins, B., "Dynamics of Flight-Stability and Con-trol," Wiley, New York, 1958.
27. Truong, D. V. and Tobak, M., "Aerodynamic Mathematical Modeling Based on Indicial Response Approach: Derivation from Navier-Stokes Equations. Part I.- Time-Invariant Equilibrium State," NASA TM-102856, Oct. 1990.
28. Bridgeman, P. W., "Dimensional Analysis," Paper Bound ed., Yale University Press, New Haven and London, 1963.
29. Leishman, J. G. and Nguyen, K. Q., "State-Space Representation of Unsteady Airfoil Behavior," AIAA J., Vol. 28, No. 5, May 1980, pp. 836-844.
30. Coon, M and Chapman, G. T., "An Experimental Study of Separated Flow on an Aspect Ratio Three Flap at Mach Number 2.4" AIAA Paper No. 91-1620, AIAA Fluid Dynamic, Plasmadynamic and Lasers Conference, Honolulu, HI, July 1991.
31. Jenkins, J. E. and Myatt, J. H., "Modeling Nonlinear Aerodynamic Loads for Aircraft," AGARD Workshop on Stability in Aerospace Systems, Toulouse, France, June 1992.
32. Bisplinghoff, R. L. and Ashley, H., "Principles of Aeroelasticity," Dover Publications, Inc., New York, 1975.
33. Nissim, E. and Gilon, N., "The Aeroelastic Equations of a Slender Body," Israel J. of Technology, Vol. 10, No. 1-2, 1972.
34. Chapman, G. T. and Yates, L. A.., "Nonlinear Aerodynamic Parameter Estimation and Model Structure Identification," , AIAA Paper No. 92-4502, AIAA Atmospheric Flight Mechanics Conference, Hilton Head, SC, Aug. 10-12,1992

BUILT-IN SELF-TEST DESIGN OF PIXEL CHIP

Chien-In Henry Chen
Assistant Professor
Department of Electrical Engineering
Wright State University
Dayton, OH 45435

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

BUILT-IN SELF-TEST DESIGN OF PIXEL CHIP

Chien-In Henry Chen, Assistant Professor
Department of Electrical Engineering
Wright State University
Dayton, OH 45435

ABSTRACT

Presented in this paper is the Built-In Self-Test (BIST) design for Wright Laboratory, ELED's PIXEL graphics chip. Based on the inadequacy of the COMPASS design tools to provide an efficient BIST methodology, a custom BIST plan was designed. Numerous combinational and sequential fault simulations support the efficiency of this design. Sequential testing reduces hardware overhead and concurrent testing reduces test time. In addition, an in depth study of a derivative of Circular Built-In Self-Test was performed and evaluated on a number of sequential benchmark circuits. While the results were extremely favorable, Circular BIST required more hardware overhead than the proposed BIST design with similar fault coverage.

BUILT-IN SELF-TEST DESIGN OF PIXEL CHIP

Chien-In Henry Chen

1. INTRODUCTION

Presented within is a progress report summarizing work completed (and subsequent conclusions drawn) over the twelve weeks of our research period at Wright Laboratories, Electronics Directorate, Wright Patterson Air Force Base, Dayton, Ohio.

A brief preliminary evaluation of the COMPASS design tools is presented in section 2. Section 3 is devoted to an evaluation of the ALU and Linear Tree elements of the PIXEL graphics chip and their suitability to Built-In Self-Test techniques. Section 4 discusses an initial plan to implement BIST to the ALU and Linear Tree elements. Section 5 discusses the updated BIST plan. Section 6 offers a brief conclusion and presents an overview of the Circular BIST study.

2. COMPASS DESIGN TOOLS

COMPASS offers a wide variety of powerful design tools for both Standard Cell and Gate Array design methodologies. As our focus was intentionally limited to Standard Cell design, all discussion of design tools is made within the context of the Standard Cell design methodology.

The Logic Assistant is the main tool for Standard Cell design and layout, but the Terminal Shell is equally important in regards to simulation and timing. Within the Logic Assistant are a number of powerful tools including the Design Assistant, the Timing Verifer, the HDL Assistant, and the Test Assistant. Exploration into COMPASS' design tools was made primarily in the Test Assistant, but enroute many other aspects of the Logic Assistant were utilized.

COMPASS is complete with an extensive set of Standard Cell libraries, including both technology specific and portable technology libraries. In the Logic Assistant, through the use of the Portable and Synthetic libraries, we completed a datapath design, and with the use of the State Machine library we completed a state machine design from a state diagram representation. Both design examples were constructed within the graphical design environment, although the latter could have been completed in the Shell.

Also in the Logic Assistant, we built an 8-bit Adder/Subtracter. Shown in figure 1, this is a 93 gate combinational circuit that is used as a test bench for combinational fault simulation.

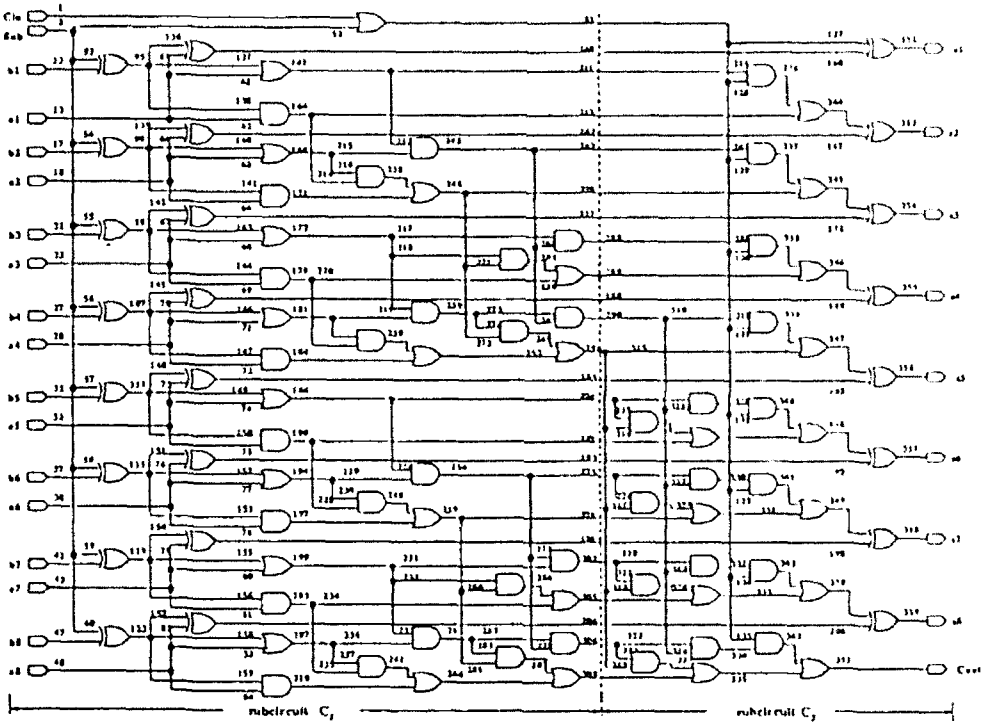


Figure 1. 8-bit Adder/Subtracter

In the Shell, using qsim, we attempted to run COMPASS' PROFAULT fault simulator. This was aborted as PROFAULT was not under a licensing agreement. Upon advice from the COMPASS Hotline, we attempted to run another fault simulator, MACH1000, from the Netlist Utilities in the Shell and met with a similar response.

COMPASS' Test Assistant offers four main test methodologies which may be used independently or in conjunction with one another. These Methodologies are: Boundary Scan Test (BST), MUX Isolation, Threshold Detection, and Full Scan Design. Within COMPASS, BST and Threshold Detection are mutually exclusive.

COMPASS Boundary Scan Test is compatible with IEEE Std. 1149.1. The BST13E000d library contains a large number of BST register cells and controllers. The controllers are dependent on which BST instructions are supported. Mandatory functions are EXTEST, SAMPLE/PRELOAD, and BYPASS. Optional instructions include INTEST, IDCODE, RUNBIST, and SCANM.

Boundary Scan Test was incorporated into the 8-bit adder/subtracter above. Upon examination of the schematic created from the Test Assistant, the functionality was determined to follow the IEEE Std. 1149.1. The only optional instruction supported was the INTEST. Examining the controller, a few observations were made. The Test Access Port (TAP) controller was generated using a standard 16-state state machine shown in figure 2.

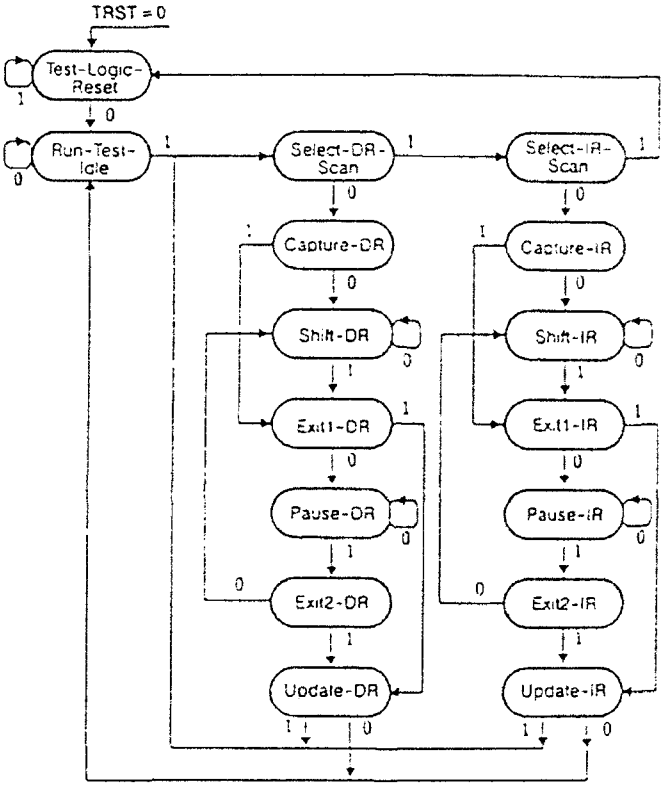


Figure 2. TAP controller state diagram

It is interesting to note that here the state machine was implemented in a conventional manner (combinational gates and Flip-Flops) rather than the PLA implementation used in the state machine compiler. COMPASS' use of the Single Transport Chain (STC) Concept to reduce silicon overhead is evidenced in the sharing of the instruction and bypass registers. The instruction decoder decoded a 3-bit instruction into five control signals that control the value of the BS register, the update stage of the BS register, the BS test mode, and the EXTEST/INTEST mode.

Test Assistant's MUX Isolation test methodology was also independently implemented into the 8-bit adder/subtracter. The MUX partitioning line is shown dotted in figure 1. Due to confusion in specifying an

instance list, the partition did not occur exactly on the line specified. However, the correct number of multiplexers were inserted and test control logic was generated. Upon preliminary examination of the schematic an interesting observation surfaced. It appears that only ONE MUX can be accessed at any given time. This would seem to prohibit the use of COMPASS' MUX Isolation for any partitioning scheme. Also, it is required that additional test pads be added to a combinational circuit for additional test control, including pads for a test clock, test reset, and test enable. Examination of Test Assistant's Threshold Detection circuit was interrupted due to a software error and Full Scan Design was disabled. Further evaluation is in progress.

The Test Assistant has the capability to add Built-In Self-Test (BIST) architecture to RAMs, ROMs, and multipliers via its BIST compilers. Furthermore, Linear Feedback Shift Registers (LFSR) may be generated automatically from Test Assistant's LFSR Compiler. However, Test Assistant can NOT automatically add ANY custom BIST circuitry to any component. All BIST MUST be done manually. BIST can be incorporated into Boundary Scan Test by supporting the RUNBIST instruction.

3. SUITABILITY OF PIXEL ELEMENTS FOR BUILT-IN SELF-TEST

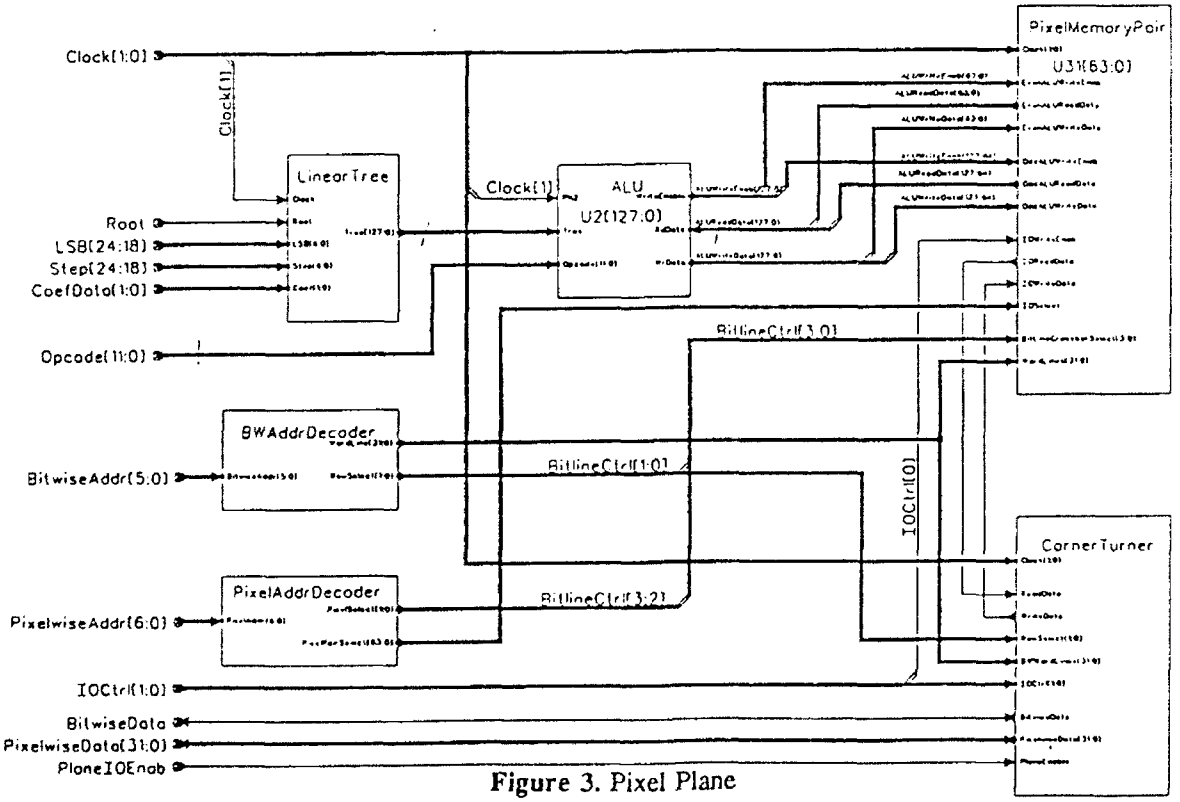


Figure 3. Pixel Plane

Figure 3 shows the PIXEL Plane. As can be seen from figure 3, the PIXEL chip is composed of a number of large blocks each of which can be evaluated to determine its suitability to Built-In Self-Test techniques. Our initial focus was on the 128-bit ALU block and the Linear Tree structure.

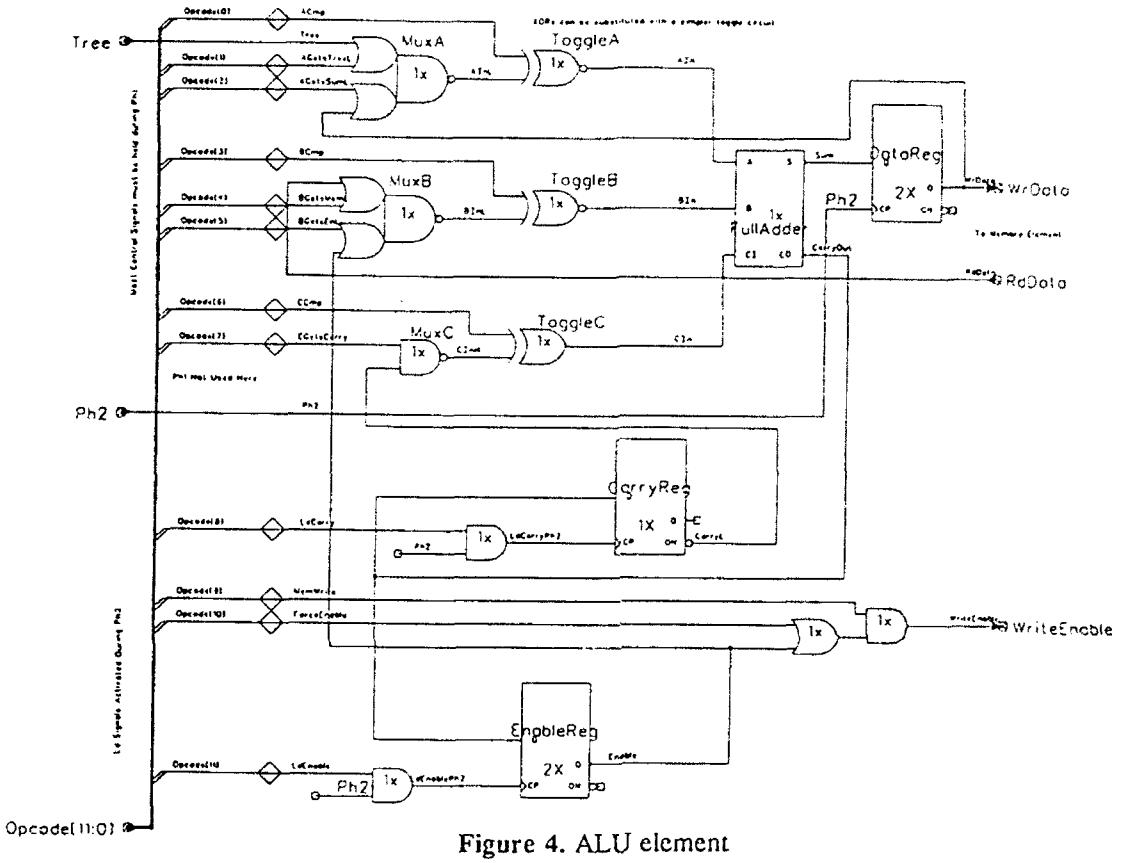


Figure 4. ALU element

A bit-slice of the 128-bit ALU block is shown in figure 4. The ALU element is a sequential circuit employing 14 simple gates, a full adder, and three D-type flipflops. The Linear Tree block is divided as shown in figure 5, with the Tree3Stage shown in figure 6. The basic block of the Linear Tree, shown in figure 7, is also a sequential circuit containing three simple gates, a full adder, and three D-Type flipflops. For analytical purposes, the full adder of figure 8 is assumed for both the full adder of the ALU and for the full adder of the single tree stage.

The first step in our analysis of the elements of the PIXEL chip was to determine if the elements could achieve high fault coverage in a simple combinational model. In order to construct a combinational model of a sequential circuit, the feedback paths must be broken and replaced with extra inputs. A combinational model of the ALU element is shown in figure 9. Note that the feedback paths have been cut and replaced by extra

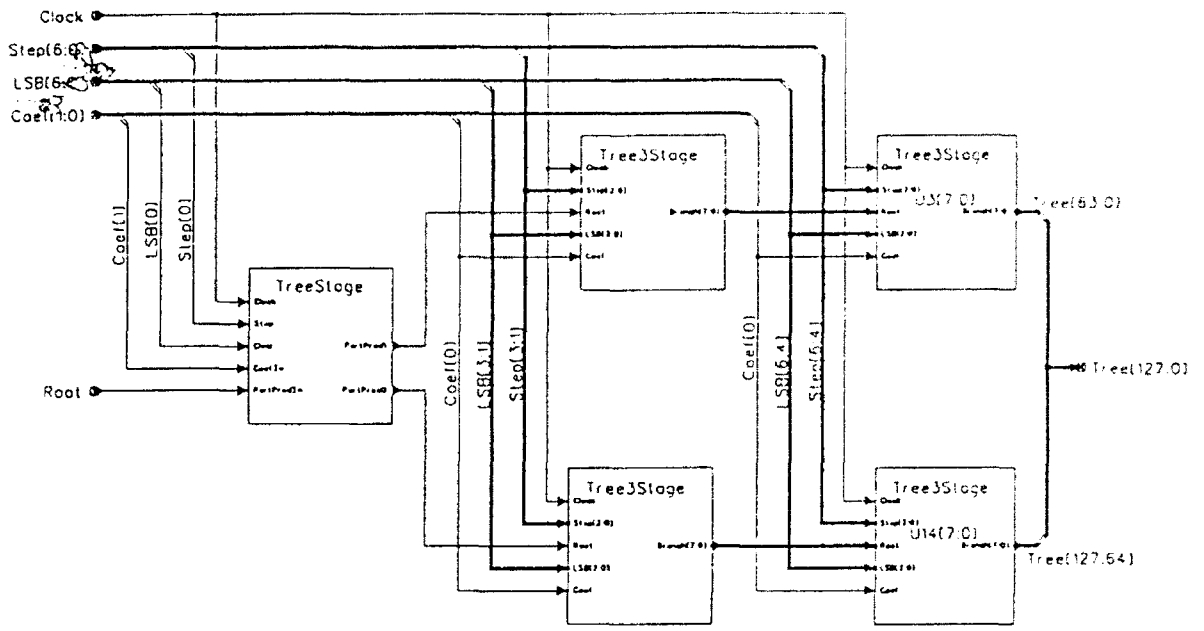


Figure 5. Linear Tree

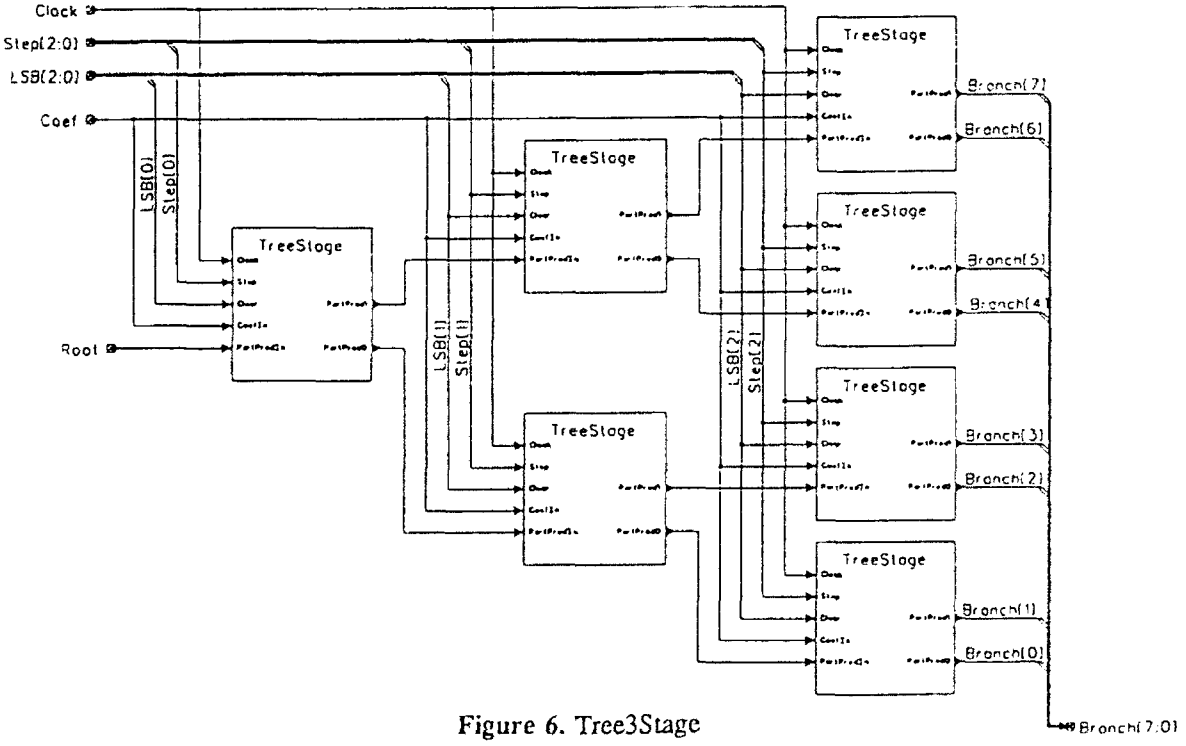


Figure 6. Tree3Stage

inputs (shown in Fig. 9) and that extra outputs now exist. The combinational model of the Tree stage can be constructed similarly. From these models, gate level and transistor level fault simulations were performed. The results are tabulated in table 1. Examination of table 1 reveals that the sub-circuits are 100% testable for

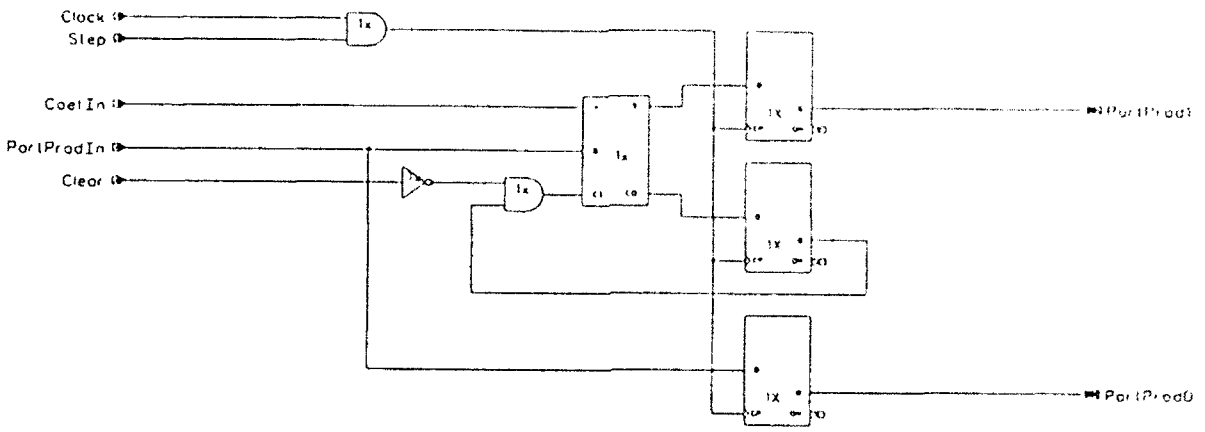


Figure 7. TreeStage

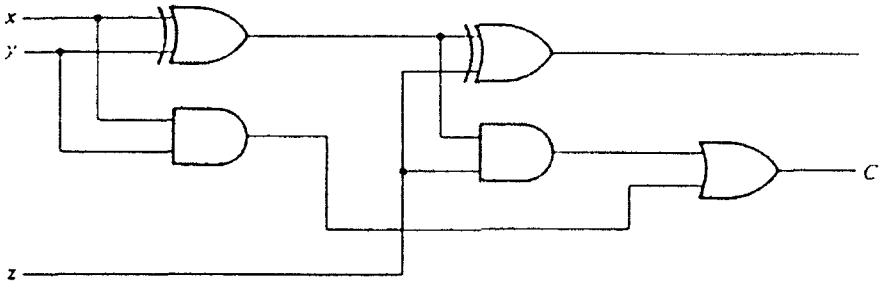


Figure 8. Full Adder

stuck-at-1, stuck-at-0, and stuck-open faults in the combinational model. This does not guarantee that the sequential circuit is 100% testable.

Table 1. Combinational fault simulation of ALU and Linear Tree

Circuit	PI	PO	GATE LEVEL				TRANSISTOR LEVEL			
			Vectors	Faults	Fault Coverage	Test Time	Vectors	Faults	Fault Coverage	Test Time
ALU Slice	18	5	9	70	100.00%	0.133 sec	20	62	100.00%	0.167 sec
Linear Tree	6	3	6	32	100.00%	0.067 sec	12	25	100.00%	0.100 sec

Sequential models of the ALU element and the single tree stage were also constructed, and sequential fault simulation was performed. Note that in sequential testing, gated signals connecting to clock inputs are omitted from testing. Results from the sequential fault simulation are shown in table 2. Examination of table 2 indicates that the sub-circuits are 100% sequentially testable provided the test patterns are applied in the

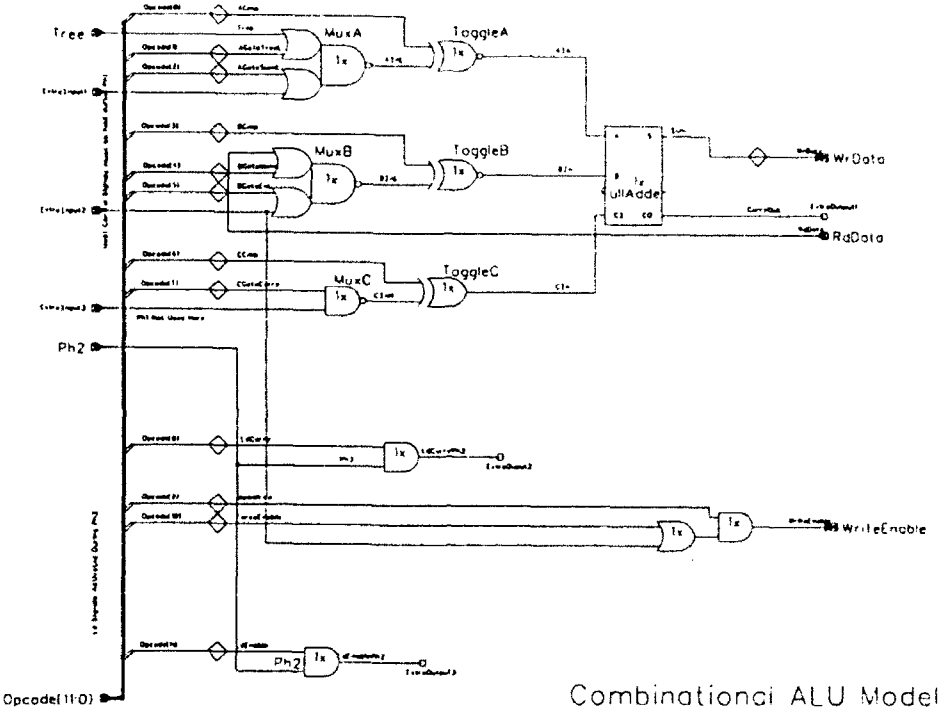


Figure 9. Combinational model of ALU element

proper order. It should be noted that both the combinational and sequential models were tested with random patterns.

Table 2. Random sequential fault simulation

Circuit	PI	PO	Vectors	Faults	Fault Coverage	Test Time
Alu Slice	12	2	17	68	100.00%	0.388 sec
Linear Tree	3	2	15	34	100.00%	0.317 sec

The most efficient method to implement Built-In Self-Test is to use Linear Feedback Shift Registers (LFSR) as test pattern generators. Linear feedback shift registers produce non-zero, non-repetitive patterns which are dependent on the previous state of the register. Given an initial seed, the LFSR the cycles through a set of test patterns automatically. However, since the patterns generated from the LFSR are no longer arbitrarily chosen and placed, high fault coverage can not be guaranteed for a given test set.

The patterns generated from LFSR were then used in sequential fault simulation of the ALU slice, the single tree stage, and the DualTree stage of figure 10. The results are tabulated in table 3.

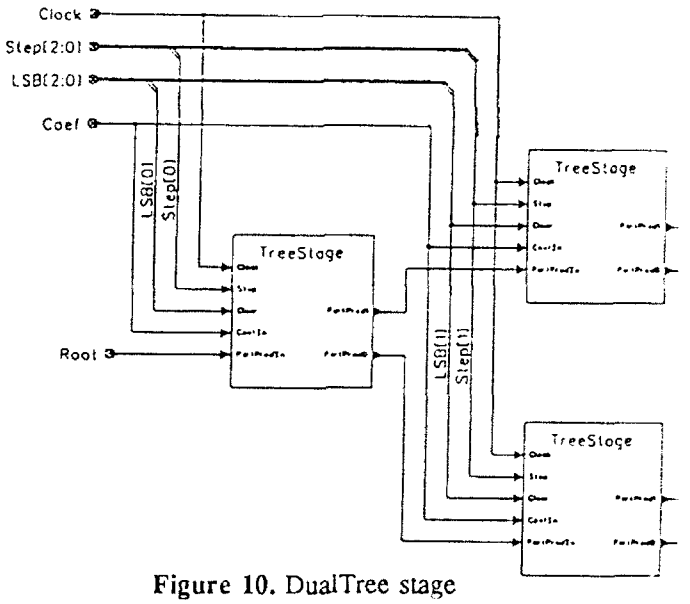


Figure 10. DualTree stage

Table 3. LFSR sequential fault simulation

Circuit	PI	PO	Vectors	Faults	Faults Detected	Faults Undetected	Fault Coverage	Test Time
ALU Slice	12	2	31	68	68	0	100.00%	0.450 sec
TreeStage	3	2	7	34	27	7	79.412%	0.283 sec
DualTree	3	4	7	96	68	28	70.833%	0.433 sec

From table 3 we can see that, even with exhaustive testing (applying all of the non-zero patterns 2**n - 1), the Tree and DualTree stages have poor fault coverage, 79.412% and 70.833% respectively. It can easily be concluded that the fault coverage will decrease more as the nested-tree structure gets deeper. On the other hand, table 3 shows that by applying only 31 tests, all 128 bit slices of the ALU can be concurrently tested with 100% fault coverage.

As a possible solution to sequentially testing the Linear Tree block, a reciprocal pattern was generated using the reciprocal primitive polynomial of the LFSR. The test set produced from such a LFSR steps through the first 2**n - 1 patterns and then steps down towards the initial seed. For example, a 3-bit LFSR with an initial seed of 100 would yield a test set of: 100,110,111,...,010,001,010,...,110,100. Using a reciprocal LFSR,

a test set was generated and sequential fault simulations were again performed on the single Tree and Dual-Tree stages. The results are given in table 4.

Table 4. LFSR sequential fault simulation with Reciprocal sequence

Circuit	PI	PO	Vectors	Faults	Faults Detected	Faults Undetected	Fault Coverage	Test Time
TreeStage	3	2	25	34	32	2	94.118%	0.367 sec
DualTree	3	4	25	96	89	7	92.708%	0.467 sec

Although the fault coverage is more encouraging, it is still not 100% and will tend to decay faster as the tree branches out. Sequential testing did not appear a viable option for the Linear Tree structure.

Since sequential testing appeared to be grossly inefficient for the Tree stage, we had to revert back to the previous combinational models. The feedback paths were again broken and replaced by extra inputs. Unlike the previous combinational fault simulations, which used randomly generated patterns to form a test set, further fault simulations were performed, which used test sets generated by LFSRs, in order to determine the Tree stages' suitability to BIST. Neglecting the gated clock signals, the fault simulation was also performed on the DualTree stage. The results are tabulated in table 5. Examining table 5, we note that for a negligible increase in test length, we can completely generate the entire test pattern, given an initial seed, from BIST hardware. While sequential testing did not provide an efficient means to test the Linear Tree structure, combinational testing provides 100% fault coverage at the expense of test length and hardware overhead.

Table 5. LFSR Combinational fault simulation

Circuit	PI	PO	Vectors	Faults	Faults Detected	Faults Undetected	Fault Coverage	Test Time
TreeStage	3	2	10	32	32	0	100.00%	0.283 sec
DualTree	3	4	11	80	80	0	100.00%	0.317 sec

4. INITIAL PLAN FOR BIST TO ALU AND LINEAR TREE ELEMENTS OF PIXEL CHIP

Having determined the suitable BIST methods for the ALU and Linear Tree elements of the PIXEL chip, we proposed the following plan to implement the BIST hardware into the existing design.

As was previously determined, the Linear Tree was to be tested from a combinational model. There are 127 single tree stages in the Linear Tree, each containing a feedback path to be broken. This requires the addition of an extra input and an extra output. Hardware requirements include 127 multiplexers to be inserted to break the feedback paths. The test mode inputs to the MUXs can all be tied to a common signal (the extra input) thus reducing the number of test signals from a possible 130 to four.

When test mode (T1, T2) = 00 (phase 1), the nested Linear Tree is tested using the test patterns generated from Test Pattern Generator TPG 1 (see figure 11).

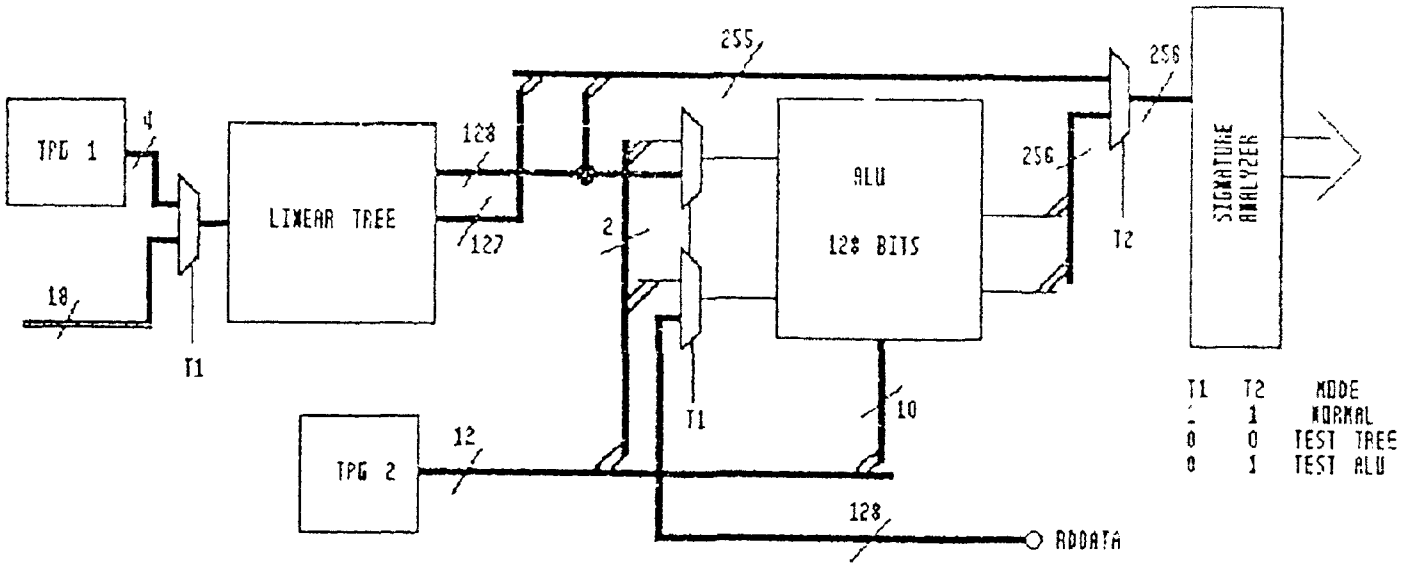


Figure 11. Initial BIST plan for ALU and Linear Tree

Since the Linear Tree is not tested sequentially, the feedback paths are broken and a common signal is simultaneously sent to each of the 127 single tree stages. Furthermore, each of the tree stages produces an extra testable output which must be routed to the signature analyzer, thus resulting in a 255-bit output from the Linear Tree.

As fault simulation has only been performed on a DualTree stage, the test length may increase rapidly as the tree branches out. If the test length exceeds the normal limit 2**25, partitioning of the Linear Tree structure into smaller subcircuits may have to be considered. However, we are fairly optimistic that this will not be necessary since the Linear Tree has only seven levels.

Again referring to previous discussion, the 128 bits of the ALU may be tested concurrently using a 12-bit test pattern generated from TPG 2 (see figure 11). Since it was determined that the ALU can be efficiently tested sequentially, no feedback paths need to be broken. However, since the ALU element contains two gated clock signals, MUXs must be inserted to break the clock path, thus enabling a single test clock (normally the system clock) to simultaneously shift data in all of the registers on the same clock edge. An example of this is shown in figure 12 for the LdEnablePh2 gated clock signal. The insertion of the muluplexers results in a hardware increase of 256 MUXs.

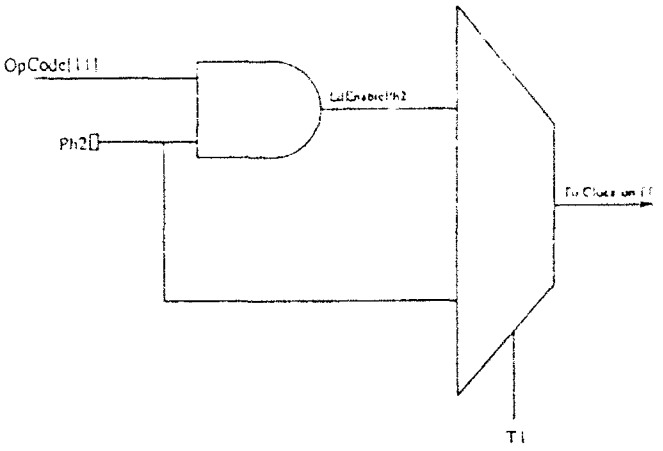


Figure 12. MUX insertion to gated clock

When test mode = 01 (phase 2), TPG 2 provides a 12-bit test pattern: 10 bits to OpCode[0-7,9-10] and one each to Tree and RdData. At the same time, OpCode[8] and OpCode[11], which control the gated clock signals LdCarryPh2 and LdEnablePh2 respectively, are disabled by the internal MUXs and the system clock is wired directly to the Enable and Carry registers. When test mode = 11 (normal operation), TPG 2, which is a Built-In Logic Block Observer (BILBO), is in the parallel register load mode, and provides the 12-bit OpCode, including the gated clock signals from OpCode[8] and OpCode[11].

In either test mode, the output signals are sent through the selection MUX to the Signature Analyzer (BILBO/MISR), modeled in figure 13, where they may be scanned out. The test pattern generators are simple LFSRs generated from the standard polynomial. No reciprocal patterns are necessary.

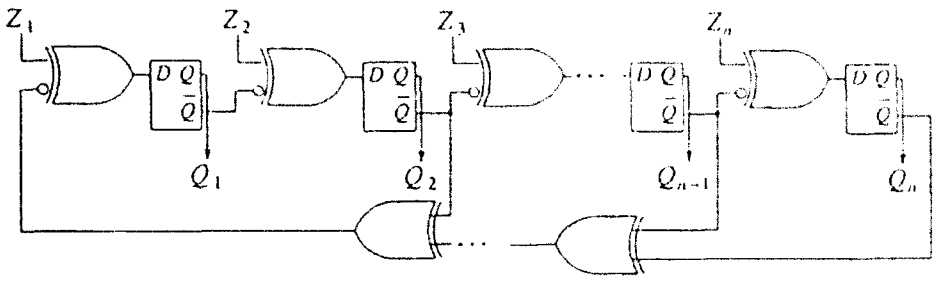


Figure 13. BILBO as a MISR

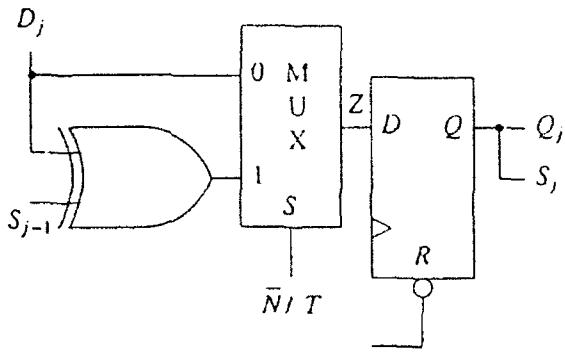
5. REEVALUATION AND UPDATED BIST PLAN

While the BIST plan for the Linear Tree structure proposed in section 4 had a low hardware overhead and a guaranteed 100% fault coverage, we did not consider routing difficulties. Based on space limitations, it was not possible to route the extra 127 outputs from the broken feedback paths of the Linear Tree to the 256 bit MISR. We concluded that a sequential model must be used to implement BIST in the Linear Tree and thus eliminate the extra output lines.

In Section 3, we showed that the Linear Tree can be tested sequentially while achieving 100% fault coverage (see table 2). However, the patterns generated to perform this simulation were not generated using a LFSR as a test pattern generator. Referring to table 3, we showed that the test patterns generated from the LFSR test pattern generator did not provide a high fault coverage. The lower fault coverage is a result of the LFSR removing the randomness in the test pattern generation. While the LFSR test pattern generator will generate non-repetitive, non-zero test patterns, the same sequence of patterns always develops after every n-1 clock cycles. In order to improve the fault coverage while using LFSR generated test patterns, we must scramble the patterns, thereby restoring some randomness to the test set.

A derivative of Circular Built-In Self-Test was implemented on the single tree and DualTree stages. The Circular BIST idea is to chain all of the input, output, and selected internal registers into a Circular Self-Test Path where the next state of the registers depends on the current state of the register as well as the combinational logic. To implement Circular BIST, all of the registers in the circular path must be modified as shown in figure 14. While this method yielded from 98% to 100% fault coverage, the increase in hardware overhead

(albeit a small increase) and the inability to set or reset the registers to an initial state forced us to abandon the Circular BIST approach.



$\bar{N}/T$	Z	Mode
0	D_j	System
1	$D_j \oplus S_{j-1}$	Test

Figure 14. Circular BIST Register

By scrambling the connections between the LFSR test pattern generator and the inputs of the Linear Tree, as shown in figure 15, we were able to increase the fault coverage to 94.118%. Although an improvement, the fault coverage is still too low.

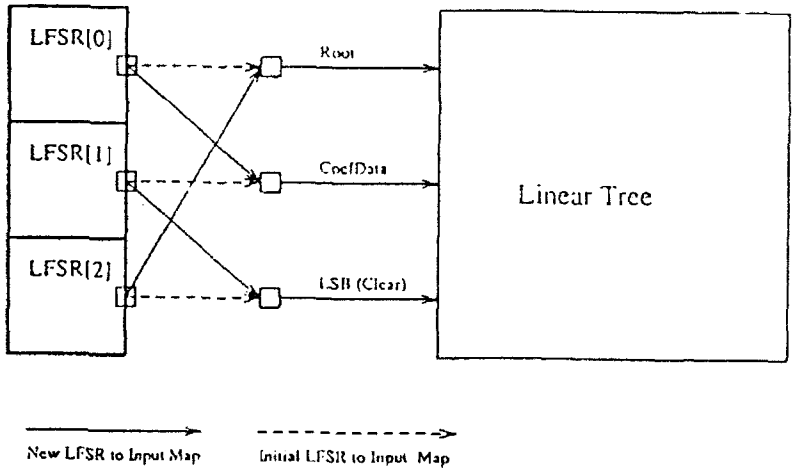


Figure 15. Scrambled test signals

Another method to increase the randomness in the test pattern generation, is to increase the length of the LFSR. Since only the first three bits will be used, the test pattern is no longer non-repetitive and non-zero. While this method provides an advantageous increase in the randomness of the test patterns, the increase in

hardware overhead is its obvious disadvantage. Disturbingly, this method was employed and yielded the same results as table 2.

Combining the previous two methods, we increased the length of the LFSR and we also scrambled the input connections. In addition, an easily testable Full Adder was used (shown in figure 16).

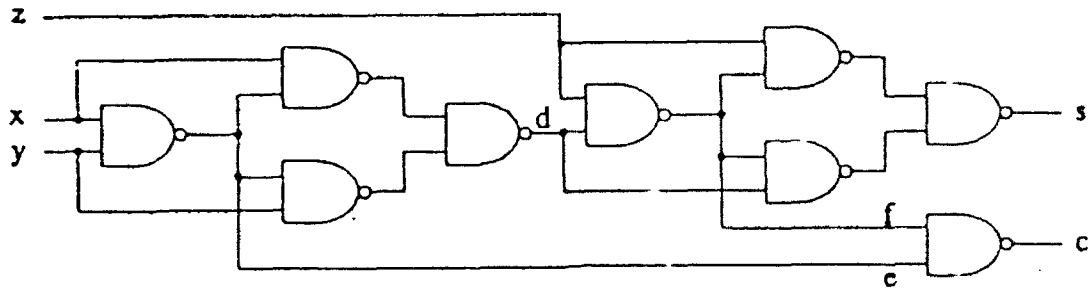


Figure 16. Testable full adder

The sequential fault simulation results for the single tree stage, the DualTree stage, and the entire Linear Tree block using various LFSRs, input sequences, and Full Adder models are tabulated in table 6.

Table 6 shows that by increasing the randomness of the test pattern generation using larger LFSRs and scrambling the input sequence, the Linear Tree can be sequentially tested with 100% fault coverage.

By sequentially testing the Linear Tree, the extra 127 output lines are eliminated, and only the 128 outputs of the Linear Tree must be sent to the MISR. In order to reduce the size of the MISR, we attempted to limit observation of the ALU slice to the WriteEn output only. Since the WriteData output was connected to a feedback path, it was our hope that the ALU could still be sequentially tested and achieve 100% fault coverage by propagating all of the faults through the WriteEn output. Sequential fault simulation results showed that by increasing the test pattern length from 31 vectors to 107 vectors, the ALU could still be tested with 100% fault coverage. Therefore, the size of the MISP can be reduced from 256 bits to 128 bits as both the Linear Tree and the ALU need only 128 outputs to be observed. An updated BIST plan for the ALU and Linear Tree is shown in figures 17 and 18.

Note also that the test pattern generators are no longer local to the Pixel Plane. Since the gated clocks in the ALU were globalized in the Padring, the 12-bit LFSR test pattern generator used to test the ALU has

been moved to the Padring. Also, since the Linear Tree is to be sequentially tested, the 8-bit LFSR TPG has likewise been moved to the EMCTrunk inside the Padring. Using multiplexers, the test signals can be sent along the appropriate buses to the subsystems under test.

Table 6. LFSR sequential fault simulation

Circuit	LFSR	FA	Vectors	Faults	Faults Detected	Faults Undetected	Fault Coverage	Test Time	Input Map
TreeStage	4bit	xor	255	34	27	7	79.412%	0.42 sec	0:0,1:1,2:2
TreeStage	4bit	xor	255	34	32	2	94.118%	0.39 sec	0:2,1:1,2:0
TreeStage	4bit	nand	255	42	33	9	78.571%	0.47 sec	0:0,1:1,2:2
TreeStage	4bit	nand	255	42	41	1	97.619%	0.40 sec	0:2,1:1,2:0
TreeStage	4bit	nand	255	42	41	1	97.619%	0.42 sec	0:2,1:0,2:1
TreeStage	4bit	xor	255	34	32	2	94.118%	0.37 sec	0:2,1:0,2:1
TreeStage	5bit	xor	255	34	33	1	97.059%	0.38 sec	0:2,1:0,2:1
TreeStage	5bit	nand	14	42	42	0	100.000%	0.27 sec	0:2,1:0,2:1
TreeStage	5bit	nand	14	42	42	0	100.000%	0.30 sec	0:2,1:1,2:0
TreeStage	6bit	xor	58	34	34	0	100.000%	0.32 sec	0:2,1:0,2:1
DualTree	4bit	xor	255	96	91	5	94.792%	0.70 sec	0:2,1:0,2:1
DualTree	4bit	xor	255	96	85	11	88.542%	0.62 sec	0:2,1:1,2:0
DualTree	5bit	xor	255	96	88	8	91.667%	0.60 sec	0:2,1:1,2:0
DualTree	5bit	xor	255	96	93	3	96.875%	0.63 sec	0:2,1:0,2:1
DualTree	6bit	xor	255	96	90	6	93.750%	0.63 sec	0:2,1:1,2:0
DualTree	6bit	xor	59	96	96	0	100.000%	0.45 sec	0:2,1:0,2:1
DualTree	5bit	nand	255	120	109	11	90.833%	0.72 sec	0:2,1:1,2:0
DualTree	5bit	nand	255	120	111	9	92.500%	0.72 sec	0:1,1:2,2:0
DualTree	5bit	nand	255	120	119	1	99.167%	0.65 sec	0:2,1:0,2:1
DualTree	6bit	nand	39	120	120	0	100.000%	0.47 sec	0:2,1:0,2:1
Linear Tree	10bit	xor	255	3816	3557	259	93.213%	27.4 sec	0:0,1:1,2:2
Linear Tree	8bit	xor	255	3816	3803	13	99.659%	32.3 sec	0:2,1:1,2:0
Linear Tree	7bit	xor	255	3816	3808	8	99.790%	52.9 sec	0:2,1:0,2:1
Linear Tree	8bit	xor	227	3816	3816	0	100.000%	28.5 sec	0:2,1:0,2:1

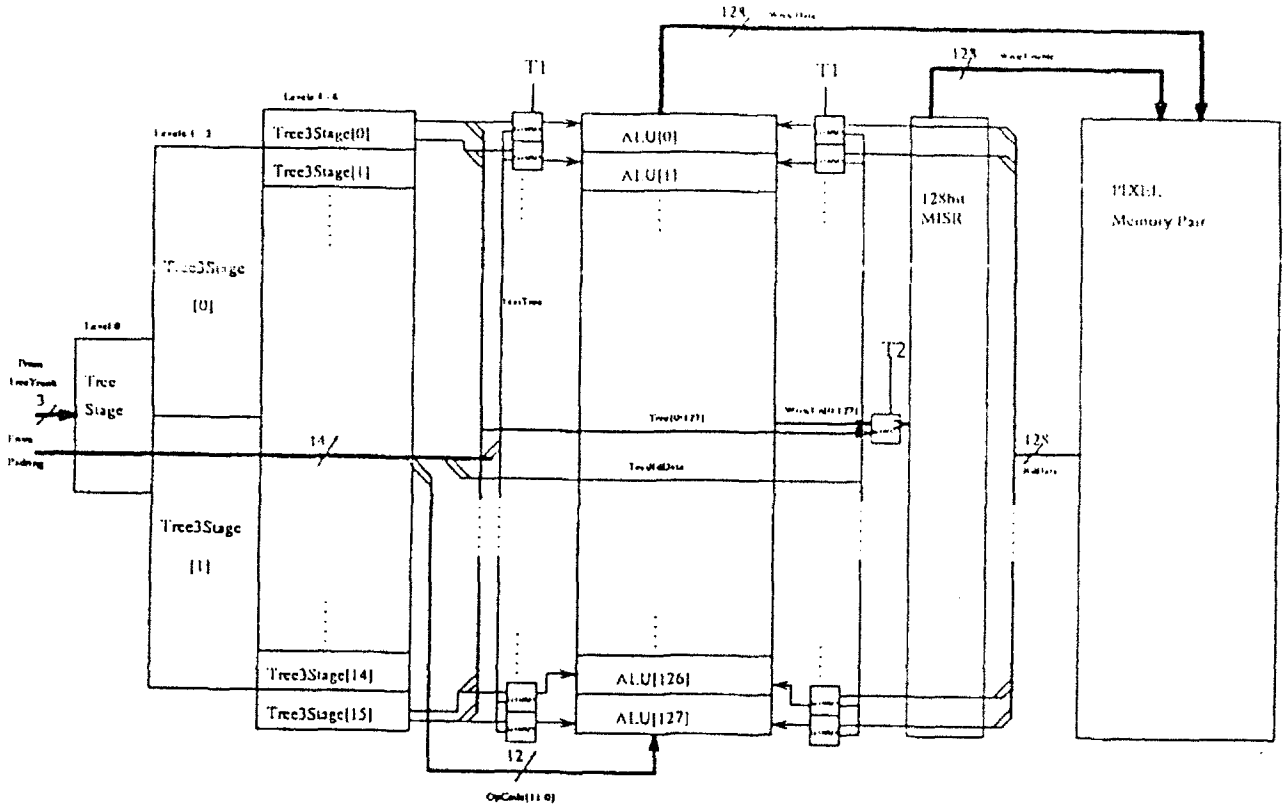


Figure 17. Final BIST plan for the ALU and Linear Tree

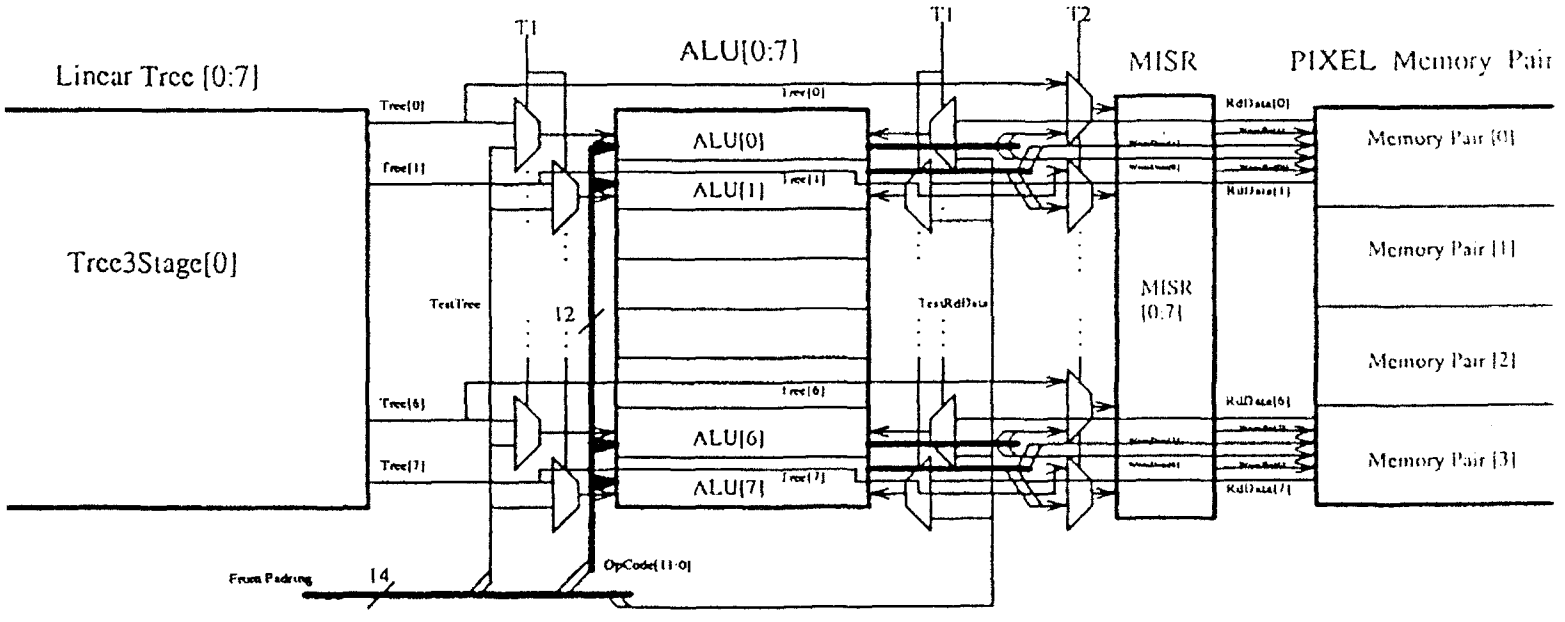


Figure 18. Final BIST plan slice

6. CONCLUSION AND TOPICS FOR FUTURE RESEARCH

As can be seen from the above results, sequential testing offers a much more efficient test methodology than combinational testing, evident most notably in a much lower hardware overhead and ease of signal routing. Although straightforward sequential testing does not always provide a higher fault coverage than its combinational counterpart, new methods of sequential testing have been developed that drastically improve fault coverage.

A new technique for designing self-testing sequential circuits, referred to as Circular BIST, was analyzed in this research. To implement Circular BIST, all of the registers in the sequential circuit must be modified as shown in figure 14. The test patterns are generated from LFSRs and fed into the primary inputs of the sequential circuit. The output response from the primary outputs is fed into a MISR for the signature analysis. The stuck-at fault simulation results from benchmark circuits shown in table 7 indicate that the circular BIST on sequential circuit achieves a higher fault coverage in comparison to several previous work on sequential circuit testing. Employing the partial scan into the Circular BIST design to achieve even higher fault coverage is currently under investigation and also a topic for future research.

Table 7. Circular BIST fault simulation

Circuit	PI	PO	FF	Gates	Faults	Vectors	Faults Detected	Faults Undetected	Fault Coverage	Run Time	Hardware Overhd
s298	3	6	14	150	366	395	366	0	100.000%	2.30 sec	17.07%
s344	9	11	15	199	418	256	417	1	99.761%	1.82 sec	14.02%
s349	9	11	15	200	426	398	423	3	99.296%	2.45 sec	13.95%
s382	3	6	21	203	495	766	495	0	100.000%	5.15 sec	15.75%
s400	3	6	21	207	520	1358	514	6	98.846%	7.95 sec	18.42%
s420	19	2	16	247	494	20000	472	22	95.547%	83.40 sec	12.17%
s444	3	6	21	226	570	478	556	14	97.544%	4.75 sec	17.00%
s510	19	7	6	242	588	20000	562	26	95.578%	109.82 sec	4.84%
s526	3	6	21	238	641	6538	635	6	99.064%	38.73 sec	16.22%
s641	35	24	19	452	581	20000	574	7	98.795%	144.30 sec	8.07%
s713	35	23	19	466	695	20000	650	45	93.525%	167.40 sec	7.84%
s820	18	19	5	317	870	20000	847	23	97.356%	213.30 sec	3.11%
s832	18	19	5	315	890	20000	852	38	95.730%	223.20 sec	3.13%
s838	35	2	32	489	985	20000	885	100	89.848%	199.60 sec	12.25%
s953	16	23	29	469	1241	20000	1211	30	97.583%	155.68 sec	11.65%
s1196	14	14	18	579	1338	20000	1333	5	99.626%	203.20 sec	6.03%
s1238	14	14	18	558	1449	20000	1375	74	94.893%	267.50 sec	6.25%
s1423	17	5	74	822	1835	20000	1819	16	99.128%	331.00 sec	16.52%
s1488	8	19	6	673	1510	512	1338	172	88.609%	14.20 sec	1.77%
s1494	8	19	6	667	1530	2048	1390	140	90.850%	36.50 sec	1.76%
s5378	35	49	179	3172	5397	20000	4742	655	87.864%	1873.01 sec	10.68%
s9234	19	22	228	6072	8001	20000	5514	2487	68.916%	6109.73 sec	7.24%
s13207	31	121	669	9320	13167	20000	8930	4237	67.821%	9727.57 sec	13.39%
s35932	35	320	1728	19556	46004	20000	42020	3984	91.340%	17394.3 sec	16.24%
Average	17	31.4	133	1910	3750.4	13031.2	3246.67	503.79	93.647%	1554.82 sec	10.77%

CHARACTERIZATION OF PART SHRINKAGE FOR LARGE, THICK INJECTION MOLDED ARTICLES

Joe G. Chow
Associate Professor
Industrial and Systems Engineering Department

Florida International University
Miami, FL 33199

Final Report for:
Summer Research Program
Wright Patterson Laboratory

Sponsored by:
Air Force of Scientific Research
Bolling Air Force Base, Washington, D. C.

September 1992

CHARACTERIZATION OF PART SHRINKAGE FOR LARGE, THICK INJECTION MOLDED ARTICLES

Joe G. Chow
Associate Professor
Industrial and Systems Engineering Department
Florida International University

Abstract

The shrinkage of large, thick injection molded parts was investigated to support the directly formed aircraft transparency program at WL/FIVR. To determine part thickness and shrinkage, a previously developed measuring procedure using a coordinate measuring machine (CMM) was employed. The part and mold thickness data were computed and plotted in the form of color contours for ease of interpretation. The causes of variable shrinkage were identified and the correlations between shrinkage and process parameters established. Based on this study and previous researches, recommendations were made for the shrinkage characterization and the mold cavity design for the molding of full scale frameless transparencies.

CHARACTERIZATION OF PART SHRINKAGE FOR LARGE, THICK INJECTION MOLDED ARTICLES

Joe G. Chow

Introduction

The increasing demand for more efficient transparency design, low cost manufacturing, and longer service life has lead to the concept of the directly formed frameless transparency program. The goal of the frameless transparency program is to eliminate the metal frame and all components necessary for attachment of the transparency panel to the frame. This can be achieved by forming the aircraft transparency directly from the molten plastic resin using an injection molding process. The hardware facilitating a direct interface between the transparency and aircraft can also be incorporated into the mold during molding.

The directly formed fabrication method offers many potential advantages for aircraft transparencies. The process permits tailoring of transparency thickness in the optical area to achieve optical quality and structurally adequate cross-sections to accommodate transparency to aircraft interfacing. Compared to current forming techniques, injection molding is primarily a one process technique which can be closely controlled and repeated. Preliminary experimental results have shown that the directly formed parts can be produced with less labor and in less time, with less variation in finished material properties.

The thickness distribution in the optical area is critical to optical quality of a transparency. Excessive thickness variation would result in distortion of scenes viewed through the transparency. Dimension control objectives for the aircraft transparency thickness shall include achieving design thickness in the

optical area within plus or minus .008 inches and thickness gradients of less than .004 inches in 4.0 inches in any direction from any point in the optical area. This stringent requirement for dimensional accuracy has made the injection molding an ideal process for fabricating transparencies.

To produce closely toleranced transparencies using an injection molding process, it is essential to be able to characterize part shrinkage after cooling. Shrinkage, an inherent characteristic of all plastics, is usually reported along with other physical properties on most manufacturer's plastic data sheets. The current industrial practice is to calculate shrinkage by multiplying the shrinkage rate of material by a linear dimension, which, in theory, indicates how much that dimension will contract after the part cools down. The calculated shrinkage again dictates how much oversize the mold should be machined. This approach works well for parts of simple shapes. For some complex shapes and odd-sized parts, however, some critical key dimensions of the part will not be as predictable from the shrinkage factor, particularly if the item is long and thick.

Objective

The main objective of this research was to characterize shrinkage which occurred in thermoplastic injection molded articles in support of the directly formed frameless aircraft transparency program. In particular, this work aimed to achieve the following goals:

(1) Measure part shrinkage values efficiently and accurately.

(2) Correlate part shrinkage to molding conditions used.

(3) Make recommendations for the molding of frameless transparencies.

Review of Previous Work

The increased use of the injection molding process in the plastics industry has spawned a great deal of researches on part shrinkage during the past two decades. A computer literature search at Wright Laboratory Technical Library has resulted in over 20 research articles on this topic. This section summarizes the findings from previous research efforts.

Work done by other researchers (1970-1992)

Many studies have been carried out to characterize part shrinkage for different plastic resins, part geometries, and molding conditions. However, most of previous work have been limited to molding relatively small plastic parts. Shot size of only a few ounces in weight are common. Molding of these small parts may be different from that of large, thick-wall aircraft transparencies, since the latter requires low injection pressure and long cycle time. In spite of these marked differences, previous researches on part shrinkage would have some bearing on this study and provide some insights into shrinkage behavior. What follows summarizes the results pertinent to this work.

1. It is generally recognized that the shrinkage is a local phenomenon, instead of a global one. Therefore, shrinkage values in a part would vary from location to location. For simple parts, the variation is minimal. But for large articles, the variation can be quite significant. Variable shrinkage in large parts often leads to many modifications of the cavity and core and occasionally to a complete retooling.

2. Part shrinkage is a function of molding process conditions and geometric factors. It depends on parameters such as hold (pack) pressure, mold temperature, injection pressure, melt temperature, cooling time, flow length, and part thickness. To completely characterize part shrinkage, the dominant process variables must be identified and the correlations between those process conditions and the shrinkage established.

3. Prediction of part shrinkage has also been attempted by two different approaches: mathematical and statistical. The mathematical prediction is still very inaccurate and limited in its capability due to complicated nature of the process. Thus the statistic approach has been favored in the past and found some success in predicting part shrinkage based on the process conditions and geometric factors used.

Work done by WL/FIVRB (1985-1992)

The injection molding experiments for the frameless transparency program were initiated by WL/FIVRB in '85. Since then, a number of trials have been performed using process evaluation test panels. The test panels fabricated consisted of two shapes. One shape was relatively flat, and the other shape was a half cone. Molded panels conforming to the two geometries have been referred to as "flat panels" and "conical panels", respectively throughout this document. The history of the process development and shrinkage study are described below.

1. The first investigation of direct forming processes for aircraft transparencies was initiated by Loral Defense Systems-Arizona in October '85. After exploring five different direct forming processes, they concluded that injection molding was the most viable candidate. To evaluate the feasibility of using injection molding for large, thick parts, they molded a number of flat and conical panels at Eimco Molded Products of Salt Lake City, Utah. The subsequent mechanical and optical testing of these panels had led to a conclusion that the injection molding process was capable of producing large closely toleranced parts at a much lower cost. In their study, the overall dimensions and the thickness of the flat and conical panels were also measured and the results presented in the Report [1].

2. Additional trials of flat panel molding took place at Eimco in January '90. The objective of this experiment was to evaluate the moldability and the mechanical properties of different resin materials. Test results showed that the

Dow-Calibre Polycarbonate materials were easiest to mold and capable of producing parts with best qualities.

3. In October '90, flat panels were molded using equipment at Hettinga of Des Moines, Iowa. Mechanical inserts for latching a transparency directly to the aircraft were also molded with the panels. This test was not completely successful mainly because the nozzle used during experiment was too small.

4. After some modification to the nozzle size, flat panels were molded at Hettinga again earlier in '91. The newer machine design and control system at Hettinga eliminated surface defects such as brown streak, and orange peel commonly found in articles molded at Eimco. Thus, the parts molded during this trial usually possessed better quality than the ones molded at Eimco. Differential heating of the mold (higher temperature in the middle and colder one at both sides) was also tested in the experiment. The short shots indicated that the flow on the thickened edge had slowed down as a result of colder temperatures on the sides. The tensile testing of the molded inserts indicate that they would be sufficient to hold the transparency to the aircraft frame during flight.

5. To characterize part shrinkage, a dimensional mapping study of molds and molded parts was undertaken at Sheffield Measurements Inc. of Dayton, OH during February to May '91. WL/FIVRB and Sheffield mapped a number of previously mol'ed flat and conical panels, and their respective molds using a coordinate measuring machine (CMM) at Sheffield. The raw CMM data were then processed at WL/FIVRB on a Silicon Graphics Iris Workstation by PATRAN/PCL routines. The thickness and shrinkage of each panel were computed and tabulated. The shrinkage data for the conical panels was later used to influence the cavity thickness for the CFT (confirmation frameless transparency) mold.

6. In April '92, WL/FIVRB engineers molded both flat and conical panels at Envirotech (formerly Eimco), Salt Lake City, Utah using Dow-Calibre 300 resins with different MFI (melt flow indices): 4, 5.5, 7, and 15. A higher melt flow

index implies a low resistance to flow and thus would enable the resin to fill the mold more quickly. Flat panels with larger thickness, 0.75", in the optical section were also molded. After molding, some of the flat panels were cut up for thickness measurement using a micrometer, while others were tested for impact resistance.

A more detailed description of these experiments and their results has been documented by WL/FIVRB [2]. Some additional observations and comments on the test data are given in the following paragraphs.

1. The CMM data shows that the mold cavity thickness in the optical area for both flat and cone molds is not a constant. Compared to the nominal size (0.500"), the thickness variation may be as high as 0.006" in the flat mold and 0.010" for the cone mold.

2. Along the flow direction, the panels molded at Eimco are generally thicker near the gate and become progressively thinner toward the aft or arch edge. The maximum thickness variation between the locations near and far away the gate may be as much as 0.024". This taper effect is mainly due to the fact that the parts are very long and long flow tends to increase the shrinkage at the far end of the part because packing pressure progressively decreases with distance from the gate.

3. For the flat panels molded at Hettinga, the variable shrinkage along the flow direction also causes a taper in the part. However, the thickness for the locations near the gate generally do not exceed 0.006" more than that near the arch end. A careful examination of these data shows that in the vicinity of the gate the panels molded at Hettinga shrank more than those molded at Eimco, while the shrinkages away from the gate were approximately the same in both cases. The excessive shrinkage near the gate for the Hettinga panels had reduced the thickness variation in the optical area.

4. One of the key differences between the moldings carried out at Hettinga and Eimco was that a smaller nozzle was used at Hettinga. The main effect of a smaller nozzle is that it tends to freeze prematurely during packing and terminates packing too soon. When this happens, the material near the gate would be less densely packed and result in a smaller thickness variation in the optical section, as discussed in the previous paragraph. The evidence of this premature freezing can be seen by the presence of a void near the gate area in the Hettinga panels.

5. The flat panels are approximately twice as long as the conical panels (24" vs 13"). Therefore, there should be more pressure drop in the mold during packing and larger part shrinkage along the flow direction for the flat panels. However, the experimental results show that the tapers in the flat panels are smaller than the ones in the conical parts, probably due to different geometry and pressure distribution during packing.

6. The effect of mold temperature on part shrinkage was not clear. For experiments carried out at Eimco, a constant mold temperature (approximately 220 $^{\circ}$F) was used. For Hettinga's trials, differential temperatures were used to heat the mold; a higher temperate in the optical section (approximately 200 $^{\circ}$F) and a colder one for the thickened sills (150 $^{\circ}$F). Since the mold temperature in the optical section was very much the same for both Hettinga and Eimco experiments, the effect of mold temperature on shrinkage cannot be inferenced from the data currently available.

7. Part dimensions from panel to panel show quite a bit variation. For both flat and conical panels, the maximum thickness variation in the optical area can be as high as 0.011" for the same location. These excessive variations may be due to different resins and process variables used during molding.

8. The measurements made by WL/FIVRB engineers indicate that the overall dimensions of the panels did not shrink much (less than 0.01 %). Thus the

thickness in the optical section would be the main focus of the dimensional control for frameless transparencies.

9. The thickness data generated by CMM and PATRAN/PCL routines were not validated. The coordinate measuring machine is very accurate, efficient, and flexible. However, the way thickness data was obtained for both the molds and parts was indirect. A more direct method to measure the part and mold cavity thickness should be used to confirm the thickness measurement obtained by a CMM.

10. The optical qualities and surface finish of the molded parts were not satisfactory. A great deal of these defects were caused by the poor mold quality and molding environment.

11. For some of the flat panels, the material on the thickened edges actually expanded, making the sill edges thicker than the mold cavity. This happened possibly because of elastic recovery after the pack pressure was released. This should not be a concern as the thickness of the thickened edge is not critical for the transparency's performance.

Present Study

One of the principal activities during this study was to dimensionally map the flat and conical panels molded at Envirotech, Utah. during April '92. The panels mapped and some of their process variables are shown in Table 1. Some of the process conditions are also included at the bottom of the table. The parts were mapped at Giddings and Lewis (formerly, Sheffield Measurements) at Dayton, OH. One of the goals in this mapping study was to further improve previous dimensional mapping procedure so that it would be more efficient and adaptable to the future mapping of the CFT. A brief description of the mapping procedure used is included here.

TABLE 1. APRIL '92 MOLDED PARTS MAPPED

	Proc #	Panel ID	Material	Part Description	
1)	MX2049	920413-08	DOW 300-15	3/4" FLAT	
2)	MX2049	920413-09	DOW 300-15	3/4" FLAT	
3)	MX2049	920413-03	DOW 300-15	3/4" FLAT	
4)	MX2047	920416-04	DOW 300-04	3/4" FLAT	
5)	MX2047	920416-05	DOW 300-04	3/4" FLAT	
6)	MX2047	920416-08	DOW 300-04	3/4" FLAT	
7)	MX2048	920414-03	DOW 300-06	3/4" FLAT	
8)	MX2048	920414-05	DOW 300-06	3/4" FLAT	
9)	MX2048	920414-07	DOW 300-06	3/4" FLAT	
10)	MX2050	920415-04	DOW XU7-5.5	3/4" FLAT	
11)	MX2050	920415-05	DOW XU7-5.5	3/4" FLAT	
12)	MX2050	920415-06	DOW XU7-5.5	3/4" FLAT	
13)	MX2052	920414-04	DOW 300-06	1/2" FLAT	
14)	MX2052	920414-05	DOW 300-06	1/2" FLAT	
15)	MX2052	920414-07	DOW 300-06	1/2" FLAT	
16)	MX2051	920415-04	DOW 300-04	1/2" FLAT	
17)	MX2051	920415-05	DOW 300-04	1/2" FLAT	
18)	MX2051	920415-07	DOW 300-04	1/2" FLAT	
19)	MX2054	920415-04	DOW XU7-5.5	1/2" FLAT	
20)	MX2054	920415-05	DOW XU7-5.5	1/2" FLAT	
21)	MX2054	920415-06	DOW XU7-5.5	1/2" FLAT	
22)	MX2053	920414-04	DOW 300-15	1/2" FLAT	
23)	MX2053	920414-06	DOW 300-15	1/2" FLAT	
24)	MX2053	920414-08	DOW 300-15	1/2" FLAT	
25)	MX2057	920421-07	DOW 300-15	1/2" CONE	
26)	MX2057	920421-05	DOW 300-15	1/2" CONE	
27)	MX2057	920421-06	DOW 300-15	1/2" CONE	
28)	MX2058	920421-01	DOW XU7-5.5	1/2" CONE	UNPACKED
29)	MX2058	920421-03	DOW XU7-5.5	1/2" CONE	
30)	MX2058	920421-04	DOW XU7-5.5	1/2" CONE	
31)	MX2058	920421-05	DOW XU7-5.5	1/2" CONE	
32)	MX2055	920422-01	DOW 300-04	1/2" CONE	UNPACKED
33)	MX2055	920422-02	DOW 300-04	1/2" CONE	
34)	MX2055	920422-04	DOW 300-04	1/2" CONE	
35)	MX2055	920422-05	DOW 300-04	1/2" CONE	
36)	MX2056	920421-01	DOW 300-06	1/2" CONE	UNPACKED
37)	MX2056	920421-03	DOW 300-06	1/2" CONE	
38)	MX2056	920421-04	DOW 300-06	1/2" CONE	
39)	MX2056	920421-05	DOW 300-06	1/2" CONE	

Process conditions used:

1. Pack time : 30 minutes.
2. Pack pressure : 750 psi.
3. Injection speed : 50 rpm
4. Melt temperature : 540°F - 570°F
5. Mold temperature : approximately 200 F
6. Fill time : 7-16 seconds for 0.50" flat panels
   12-18 seconds for 0.75" flat panels

a. Mapping Procedure:

1. Calibrate all the probes that will be used during mapping against a 1" diameter tooling ball. The calibration result will be used by the machine controller to determine the accuracy of various probe tip orientations.

2. Set up the part on the machine table.

3. Determine the reference planes in the part so that the machine coordinate system conforms to the part coordinate system.

4. Record X, Y, and Z coordinates for each point in a approximately 1" X 1" grid pattern. Store the coordinates for all the points in a 3.5" floppy disk.

5. Process the CMM data using the PATRAN/PCL routines at WL/FIVRB.

6. Generate thickness data and contours for both the molds and parts.

7. Obtain shrinkage values for the part.

The inspection program for the coordinate measuring machine was written in a programming language developed by Giddings and Lewis. This program consisted of four different modules; flat mold, conical mold, flat panels, and conical panels respectively. For the panels, the mapping started from the outside surface; in the order of port stations, optical area, and starboard stations. Then the corresponding points on the inside surface were mapped. For the molds, the core was mapped first, followed by the cavity. In addition, extra points on the top surface of both halves were taken to facilitate the matching of the two halves during computer processing.

After dimensionally mapping the first conical panel, the accuracy of the thickness data obtained by a coordinate measuring machine (CMM) and PATRAN/PCL routines was confirmed. Circles representing points recorded by the CMM were first marked on the inner surface of a conical part. The panel was then sawed

into several pieces so that the thickness along the centerline and the edges of the optical section could be checked. Approximately forty points were measured using a micrometer. The thickness at the points measured compared very well with the CMM data. The largest discrepancy between these two different measuring methods was less than 0.001". The close agreement demonstrates that the coordinate measuring machine, the measuring procedure, and the PATRAN/PCL routines used were very accurate and reliable.

b. Results and Discussion

The mapping results are divided into two parts according to the geometry: conical and flat. Conical parts and its mold will be discussed first.

Conical Parts and Mold

1. The mold was repolished in earlier '92 to eliminate the scratches on the surface, As a result, the polished mold should produce parts with larger thickness, compared to the parts molded previously. Indeed, the thickness data for the parts molded during April '92 indicate that they are generally 0.001" - 0.007" thicker than the parts molded before.

2. Most of the thickness contours for the conical parts display a very regular pattern, with contours symmetric to the centerline. Figure 1 shows the contour plot for the part MX205707_C15.

3. The part dimensions are quite repeatable from panel to panel, generally not more than 0.003" variation.

4. While the thickness contours in most of the parts are symmetric to the centerline, the contours for the repolished mold are not. Also the thickness for the polished mold cavity are generally larger than the mold cavity thickness before polishing. Due to these two problems, it was decided that the cone mold

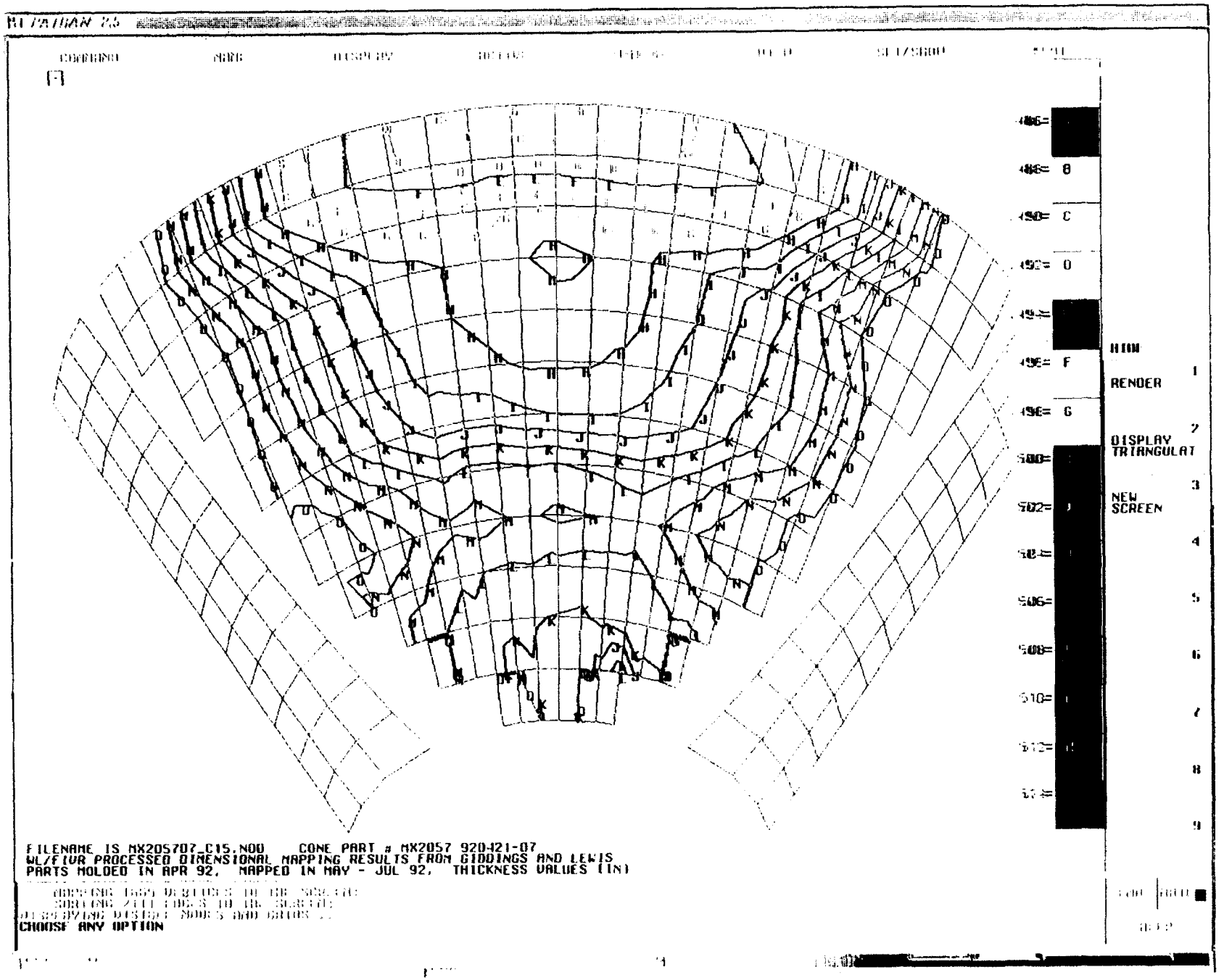


Figure 1 Thickness contours for Conical
Part MX205707_C15

should be remapped at Giddings and Lewis.

5. The part thickness for the unpacked panels are significantly smaller than the ones with full packing (by 0.015"). This is expected since packing would force additional resin to flow into the mold, making the part much more dense after cooling.

6. Melt flow index might have an influence over the filling time, but did not seem to affect part shrinkage. Therefore, this confirms that packing phase is more critical for the amount of part shrinkage after cooling than filling phase.

[1]Flat Panels and Molds

1. There is no significant difference in part shrinkage for parts with different thicknesses.

2. Part repeatability from panel to panel is excellent. In the optical section, the dimension variation are generally within 0.002". This dimensional consistency is a significant improvement over previous results.

3. Just like conical panels, melt flow index did not make any difference to the shrinkage values in the flat panels.

Conclusions and Recommendations

a. Conclusions for molded flat and conical panels

1. The results from molding process evaluation test panels, both flat and conical, indicated that the injection process was a viable technique for directly

[1] As the mapping results for flat panels are not available at the time of this writing, the micrometer results have been used to draw the conclusions for this subsection.

forming the frameless transparencies. It exhibited many advantages that were clearly superior to the current manufacturing method.

2. Based on the limited experiments carried out by WL/FIVRB and other researchers, it can be said that the most dominant parameter in determining the part shrinkage is hold pressure. The progressive pressure drop along the flow length during packing and other dominant parameters determines the amount of shrinkage after cooling.

b. Recommendations for the Molding of Confirmation and Full-scale Frameless Transparencies (CFT and FST)

Currently, the cavity sizing for the CFT mold was solely based on the shrinkage data obtained from the conical panels. Those shrinkage results were obtained using one single geometry and a very narrow range of process conditions. Furthermore, extrapolation was required in order to apply the shrinkage values for the conical panels to the CFT mold design as the CFT is significantly longer than the conical panels, This approach is a significant improvement over the current industrial practice, which simply uses a single number to uniformly scale up the mold. However, this method may suffer a severe drawback in that it neglects the effect of the process conditions and geometries on the shrinkage. Once process conditions and geometries change (such as much longer flow length in a FST), a different shrinkage pattern may emerge, rendering this approach ineffective.

A more desirable approach would be to include the process variables and geometrical factors in characterizing the part shrinkage. The dominant parameters should first be identified and the correlations between those parameters and shrinkage established. A predictive model based on statistic regression analysis techniques should be developed that would predict shrinkage for a wider range of process conditions and geometries. In addition to shrinkage, other aspects of part quality such as impact resistance, surface finish, etc.

should also be considered in selecting the molding conditions. The following recommendations are aimed to obtain a frameless transparency with good part quality and acceptable thickness variation in the optical section.

1. For the molding of CFTs, a wider range of process conditions such as injection pressure, mold temperature, pack pressure, pack time, etc. should be attempted. The goal of testing different combinations of process conditions is to determine the optimum conditions that would produce best part qualities, except the dimensional accuracy in the optical area.

2. As different combinations of process variables are used, the temperature and pressure distributions in the mold will be recorded. In addition, the shrinkage of each transparency will be determined using a coordinate measuring machine. A quantitative relationship between part shrinkage and dominant process conditions should be established.

3. Develop a statistical model which would allow shrinkage to be predicted based on the process conditions and geometry used.

4. Verify the accuracy of this model. In particular, it should be tested against the molding conditions and geometries other than the ones used to develop the model.

5. Use the predictive model to influence the mold cavities design for full scale frameless transparencies.

References:

1. "Development of Forming Process for Frameless Aircraft Transparencies", D. B. Fuller, B. Penunuri, and T. Taylor, Loral Defense Systems-Arizona, Litchifield Park, AZ, October 1988.

2. "WL/FIVR Dimensional Mapping Study of Molded Parts and Molds", K. Roach, Internal Report, Wright-Laboratory WPAFB, OH, 1991.

DETERMINATION OF MULTIPLE-SOURCE SCHLIEREN SYSTEM CAPABILITIES

Steven H. Collicott
Assistant Professor
School of Aeronautics & Astronautics
Purdue University
1282 Grissom Hall
West Lafayette, IN 47907-1282

Final Report for:
Summer Research Program
Wright Laboratories

Sposored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

August 20, 1992

DETERMINATION OF MULTIPLE-SOURCE SCHLIEREN SYSTEM CAPABILITIES.

Steven H. Collicott

Assistant Professor

School of Aeronautics and Astronautics

Purdue University

1 Abstract

Original analysis is performed to determine the noise-reduction potential of multiple-source (sharp-focusing) schlieren systems. System geometry is constrained to that arising from modification of existing traditional single-source schlieren systems. Applications in the Tri-sonic Gasdynamics Facility (TGF), the Mach 6 tunnel, and the 20-inch hypersonic tunnel are considered. Analysis based on propagation of the mutual intensity function is shown to describe the small perturbation case. Results of the analysis indicate that the smallest length scale disturbances are simple to remove, but that the largest length scales possible (on the order of the field of view) are not likely to be removed.

DETERMINATION OF MULTIPLE-SOURCE SCHLIEREN SYSTEM CAPABILITIES

Steven H. Collicott

2 Introduction

Schlieren systems have long been used to visualize density fields in wind tunnel research. A schlieren system is simply an imaging system with specialized illumination and aperture. The illumination is a collimated beam of light, and the unique feature of the aperture is a straight opaque cut-off, typically blocking half of the light. A typical wind tunnel schlieren system is sketched in Figure 1. It is assumed that the reader is familiar with the elementary operating principles of the traditional schlieren system. That is, how a density field bends light rays onto or off of the knife edge, thereby forming light and dark regions in the image. A useful explanation of this process may be found in Chapter 6 of "Elements of Gasdynamics" by Liepmann and Roshko[1].

Multiple-source, or sharp-focusing, schlieren systems were developed to eliminate one major limitation of the traditional schlieren systems. This limitation is the fact that the schlieren system will detect density gradients not only in the wind tunnel, but also in the entire region between the source slit and the knife-edge. Thus any density gradients which exist outside of the test facility will also contribute light and dark regions to the image. Multiple-source schlieren attempts to eliminate this image noise by probing the flow with a number of collimated beams in different directions. Operation of a multiple source schlieren system with incoherent illumination can be viewed as the sum of a number of individual schlieren systems. The sharp-focusing effect will be described in detail below.

"Sharp-focusing" is an unfortunate choice of modifiers to describe the system. Since a traditional schlieren system focuses the object space (wind tunnel model) into the image space (photographic film), the term "sharp-focusing" does not convey any differences to the unenlightened. To compound the problem, "focusing" is a term used to describe an image in a transverse plane, and "sharp" implies a well-focused image. However, primary difference between the performance of the two schlieren systems is the longitudinal, or axial, response. Thus "multiple source schlieren" is a more descriptive name.

2.1 Motivation

Noise in schlieren images can arise from density gradients anywhere between the source slit and the knife-edge. Examples of these density gradients are thermal currents in the room and turbulent boundary layers on the windows of the wind tunnel test section. The severity of the noise depends not only on the magnitude

of these superfluous density gradients but also on the magnitude of the density gradient that is to be visualized. As the Mach number of a blow-down facility increases, the density decreases, and thus density gradients decrease. Hence extraneous thermal currents which cause negligible noise in the image of a Mach 2 flow may obscure the entire signal in a higher Mach number flow

In order to reduce the image noise, a multiple-source schlieren system superimposes the schlieren images from a series of traditional schlieren systems. In place of a single source slit, a series of slits is used. These are called the source slits or the source grid. Then in place of the traditional knife-edge, a second series of slits, the cut-off slits or cut-off grid, is used. For each source slit, there exists a corresponding cut-off slit. As implied by the name, the cut-off slits function analogously to the knife-edge in a traditional schlieren. As sketched in Figure 2, the centerlines of all of the beams formed from the source slits are coincident at both the ideal object plane and the ideal image plane. Thus the images of the ideal object plane created by each individual schlieren system coincide at the ideal image plane. The key to the operation of the multiple-source schlieren is simply that the centerlines of the beams are not coincident in the remainder of the region between the source and cut-off slits. This causes the individual images of the regions outside the ideal object plane to not be coincident in the ideal image plane. The shift between the various images results in reduced visibility of that plane in the image plane. The further a disturbance is from the ideal object plane, the more it is spread out in the image. While a given disturbance now affects a larger region of the image plane, the magnitude of the effect is diminished. Thus the multiple-source schlieren possesses a certain degree of axial selectivity

Additional interest in multiple-source schlieren has been generated by the hope that axial selectivity could be optimized to be highly selective. Then the selected region could be scanned through a three-dimensional flow, mapping out, for example, the three-dimensional structure of a detached bow shock on an arbitrary body. This has not been accomplished, to the author's knowledge, and the present analysis does not address this task. Difficulties exist in designing such a system because published analyses of sharp focusing schlieren systems are sketchy, and sometimes in error.

2.2 Shortcomings of previous work

Previous work in multiple-source schlieren systems began about 1950[2, 3, 4]. This initial period of interest is terminated by what appears to be the most comprehensive analysis to date, a master's thesis by L. R. Boedeker in 1959 at MIT[5]. The analysis is a mixture of the geometric analyses of traditional schlieren operation and depth-of-focus in an imaging system. Boedeker's approach is mostly sound, and is used to compare several different arrangements of multiple-source schlieren systems. The interpretation of the results

of his analysis is sufficient for his immediate needs, but is lacking in rigor and generality. Boedeker's primary goal was to maximize the field of view of the system, so the analysis has a definite slant towards that point of view.

Other papers of the era of initial interest are brief in comparison with Boedeker's thesis. These papers describe the basic operation in a qualitative manner. The descriptions are useful for those wishing to acquire an insight into the device and its operation, but are insufficient for design purposes. That is, these papers are an insufficient description of multiple-source schlieren if your goal is to reduce the effect of thermal currents in the room by 90%. Because this initial era preceeds the invention of the laser, possible use of a highly coherent light source was obviously not considered in these efforts.

The present era of interest in focusing schlieren has grown from work performed by Weinstein at NASA Langley Research Center in the late 1980s. An AIAA conference paper[6] has resulted from his work. It describes a number of multiple-source schlieren systems which he constructed and demonstrated.

Papers from both eras share some common shortcomings. In general issues, the primarily flaw is in the interpretation and discussion of the analyses of system performance. With regards to a specific item, the lack of a quantitative definition of the axial selectivity has hindered comparison of different systems and prevented the logical design of all systems. Note that with no quantitative measure of axial selectivity, no analysis to determine how selectivity depends on various system parameters (number of slits, slit spacing, inter-beam angles, magnification, etc.), is possible. Boedeker addresses this in regards to the depth-of-focus of an imaging system with a point object, but the results do not appear to apply directly to sharp focusing schlieren systems for viewing finite-sized objects.

3 Goal of this work

The goal of this work is to assess the feasibility and benefits of implementing a multiple-source schlieren system in three wind tunnels of the Experimental Engineering Branch of the Aeromechanics Branch of Wright Laboratories (WL/FIMN). These tunnels are the Trisonic Gasdynamic Facility (TGF), the Mach 6 tunnel, and the 20-inch hypersonic tunnel operating at nominal Mach numbers of 12 or 14. The motivation for the study is to determine whether multiple-source schlieren systems could reduce noise in the images. The noise is presumed to be generated by thermal currents in the rooms which enclose the test sections. In the case of the TGF, this includes a control room with numerous cooling fans ejecting streams of turbulent warm air into the room. Because of the state of knowledge in the design of multiple-source schlieren systems existing at the beginning of this work (see above), the majority of the work performed to meet this goal was

original analysis. Application of the previous analytical efforts did not provide a sound method to design a multiple-source schlieren system or to predict its performance.

To meet the goal, analysis of the axial selectivity of multiple-source schlieren systems was performed. This began with a review of the previous efforts. This review, and analogous elementary analyses to verify or modify the reports, provided a base from which to proceed. The present work defines a practical quantitative measure of the axial selectivity and then determines how this is governed by system parameters. This permits intelligent design of focusing schlieren systems for noise reduction in the TGF, Mach 6, and 20-inch wind tunnels. One benefit of such an approach is that the knowledge gained in the analysis can be presented in a logical manner. A journal paper will be written based on the analysis summarized in this report and measurements to be performed by the author and his graduate students at Purdue University.

4 Analysis

Geometric analysis describes the basic operation of the multiple source schlieren system. Boedeker's geometrical optics work is reviewed and modified. Fourier optics analysis can be applied in some places, but generally fails when applied to beams of incoherent light trans-illuminating an object at different angles. The effects of pure phase objects in an incoherent system, such as a shock wave, also fail to be described well by Fourier optics. Since both of these situations exist in a multiple source schlieren, simple Fourier optics methods are not applicable. Complicating analysis further is the presence of the knife edge or cut-off slits. These cut-off devices permit visualization of the density gradient field, but by doing so, they destroy the linearity of an incoherent system. That is, the image plane intensity is no longer determined by a superposition integral of the geometric image intensity and an intensity point spread function. Thus a more comprehensive analysis method is required. Description of the light by use of the mutual intensity function is sufficient to describe incoherent illumination at arbitrary angles, the effect of pure phase objects, and the propagation through the optical system. Additionally, varying degrees of coherence can be modeled. Unfortunately, analysis becomes difficult when the knife edge or cut-off grid is included with a phase object. Nevertheless, a small density gradient model is examined which provides insight into the operation of the multiple slit schlieren system.

4.1 Geometric analysis

The multiple-source schlieren system can be regarded as a combination of a series of individual schlieren systems, or it can be considered as an imaging system with an unique extended source and a corresponding

aperture. For an incoherent source, this combination occurs by a summing of the individual image intensities. For partially or fully coherent light sources, the combination is more complex, and is not addressed here. In this work, and in Boedeker's, it is assumed that a typical schlieren system has a source slit small enough such that the depth-of-focus of the system is approximately infinite. This is a safe assumption because the sensitivity of a schlieren system requires a small source slit (see Chapter 6, Liepmann and Roshko). Images of objects outside of the ideal object plane are displaced from each other in the ideal image plane, as shown in Figure 3. The axial selectivity is best described by considering the distance from the ideal object plane at which a disturbance in the density field has individual images completely displaced from each other. That is, the value of $\tilde{z}_o - z_o$ to produce $h = ML_f$, where $M = z_i/z_o$ is the imaging magnification, and L_f is the transverse length scale of the disturbance in the object space. Call twice this value (for both sides of the object plane) the integration length, L_i. This definition differs from Boedeker's since it depends on a length scale of the disturbance. This may appear to be a complication, but it is necessary to arrive at a useful description of the axial selectivity. From geometry, it can be seen that

$$L_i = 2L_f/\theta \ . \tag{1}$$

Note that the axial selectivity depends only on the inter-beam angle and the length scale of the disturbance. Thus different length scales in the fluid are selected differently. This is consistent with how the depth-of-focus in an imaging system depends on the transverse length scale chosen.

Boedeker performs similar analysis, based on an assumption that an incoherent extended source will emit rays at all possible angles in the direction of the imaging system. This is generally a safe assumption for incoherent sources, but it is improperly applied in this case. Light from the source passes through the source slits and then through a lens to form a converging series of ray bundles. Thus rays are incident at a finite number of angles.

4.2 Mutual intensity Analysis

Fourier optics describes the amplitude of the electric field and the intensity in a series of transverse planes. For example an object plane and an image plane. Similarly, a property termed the "mutual intensity" may be defined in these transverse planes. The mutual intensity function is defined to be the time-averaged product of the electric field amplitude at point P_1 and the conjugate of the field at point P_2:

$$J(P_1, P_2) = \langle U(P_1)U^*(P_2) \rangle \ . \tag{2}$$

Note that for $P_1 = P_2$, the mutual intensity function describes the intensity at the point P_1. Since P_1 and P_2 are general, the entire intensity field is described. One reason why the mutual intensity function is required,

rather than simply the intensity, in the analysis of multiple-source schlieren systems can be seen from a simple example.

Consider a collimated beam of incoherent light traveling parallel to the optical axis as shown in Figure 4. The intensity of such a beam would be given by $I(x,y) = I_o f(x,y)$, where I_o is a constant and $f(x,y)$ is a function bounded by zero and unity. Now consider a similar beam travelling at an angle to the axis. One could create a beam such that the intensity of this beam could also be described by $I_o f(x,y)$. Thus the intensities of the two beams are the same, and yet they will propagate through the schlieren system differently; the beam at an angle could be blocked entirely by the knife-edge. The mutual intensity function is capable of describing these two beams and the differences between them. To demonstrate this, denote the (complex phasor) amplitude of the electric field of the on-axis beam by $U(x,y)$. We know that the intensity is given by $I(x,y) = U(x,y)U^*(x,y)$. Now turn this beam off-axis, but still in the y-z plane. The field amplitude will be given by $U(x,y)\exp[-j2\pi \sin\alpha x/\lambda]$. Thus the intensity remains $I(x,y) = U(x,y)U^*(x,y)$ and fails to differentiate between the on-axis and off-axis beams. The mutual intensity of the on axis beam is given by $J(x_1,y_2;x_2,y_2) = \langle U(x_1,y_1)U^*(x_2,y_2)\rangle$. It has been shown[7] that frequently a good description of an incoherent source in a imaging system can be arrived at by assuming that the electric fields at any neighboring points are perfectly incoherent:

$$J(x_1,y_1;x_2,y_2) = I(x_1,y_1)\delta(x_1-x_2,y_1-y_2)\,. \tag{3}$$

The mutual intensity of the tilted beam is found similarly:

$$J(x_1,y_1;x_2,y_2) = \langle U(x_1,y_1)\exp[-j2\pi\sin\alpha x_1/\lambda]U^*(x_2,y_2)\exp[+j2\pi\sin\alpha x_2/\lambda]\rangle \tag{4}$$

$$J(x_1,y_1;x_2,y_2) = \langle U(x_1,y_1)U^*(x_2,y_2)\rangle\exp[-j2\pi\sin\alpha(x_1-x_2)/\lambda] \tag{5}$$

$$J(x_1,y_1;x_2,y_2) = I(x_1,y_1)\delta(x_1-x_2,y_1-y_2)\exp[-j2\pi\sin\alpha(x_1-x_2)/\lambda] \tag{6}$$

Note that both beams have the same intensity, and that the mutual intensity function may be complex valued. Of course, the mutual intensity function must be positive and real valued for $(x_1,y_1) = (x_2,y_2)$.

Propagation of the mutual intensity through thin phase objects, lenses, and free space is described by methods analogous to Fourier optics, but in four dimensions rather than two[7]. To apply these tools to the multiple-source schlieren system, a valid description of the mutual intensity incident on the wind tunnel test section is required. The multiple-source schlieren system has source slits, beam angles and cut-off slits which are a function of only one transverse direction, say x. In the $y-z$ plane, the system acts like a simple incoherent imaging system. Thus the y-dependence of the mutual intensity will be ignored. This is entirely consistent with the x-y separability of the functions which describe the source slits, lenses and

mirrors, cut-off slits, and the separability of the free-space propagation integrals. Thus it is assumed that $J(x_1, y_1; x_2, y_2) = J_x(x_1; x_2)J_y(y_1; y_2)$, and that the present task is to solve for the propagation of the J_x factor.

The mutual intensity for incoherent illumination of the source slits is described by the following mutual intensity function:

$$J_x(x_1; x_2) = \sum_{i=-N}^{N} \text{rect}(x_1 - ia_2/a_0)\delta(x_1 - x_2) \ , \tag{7}$$

where a_0 is the slit width, and a_2 is the distance between centerlines of adjacent slits. Beginning from the source slits, propagating to the first mirror, then to the test section, produces the mutual intensity in the object plane. The mutual intensity is, ignoring a leading constant, a two-dimensional Fourier transform of the source plane mutual intensity:

$$J_{obj}(u_1; u_2) = \text{sinc}\left(\frac{a_0}{\lambda f}(u_1 - u_2)\right) \sum_{i=-N}^{N} \exp\left(j2\pi \frac{i \sin\alpha}{\lambda}(u_1 - u_2)\right) \ . \tag{8}$$

The reader may verify by setting $u_1 = u_2$ that this is a uniform intensity field.

The object which this mutual intensity passes through in aerodynamics research is a compressible flow. The density gradients in this field bend the light rays, or equivalently, deform the wavefronts of the light. The goal of this analysis is to determine how the multiple source schlieren system reacts to density gradients located anywhere between the mirrors. To produce an analytical model which is soluble and remains relevant, the concept of an "effective transparency" approximation is introduced into the multiple-source schlieren model. This approximation is used to model density changes outside the ideal object plane, thereby avoiding the complications of computing the mutual intensity (or even the intensity) at planes which are not conjugate (object and image) planes or Fourier transform pairs. The effective transparency is assumed to be the object plane transparency which has the original disturbance replicated and shifted as each beam would carry it to the object plane. This is sketched in Figure 5 for a series of five beams. The shift is proportional to the beam angle and the distance from the object plane. Unfortunately, no diffractive effects of propagation are included. This might not be a problem because of the very large depth of focus of each individual schlieren system.

A localized phase disturbance at $(x, z) = (x_D, -L_D)$ in an otherwise uniform density field can be written as an electric field transmittance for the ith beam: $t_i(x) = \exp[j2\pi b\delta(x - \tilde{x})]$. The shift, $\tilde{x}$, is determined from geometry (Figure 5), $\tilde{x} = x_D - ia_2L_D/f$. For $b \ll 1$, this is approximated by the first two terms of the Taylor series, $t_i(x) \approx 1 + j2\pi b\delta(x - \tilde{x})$. The mutual intensity is modified by a thin transparency[7] according

to: $J_{out}(x_1, y_1; x_2, y_2) = J_{in}(x_1, y_1; x_2 y_2) t(x_1, y_1) t^*(x_2, y_2)$. The product of $t(u_1)$ and $t^*(u_2)$ is

$$t_i(u_1) t_i^*(u_2) \approx 1 + j2\pi b\left[\delta(u_1 - \tilde{x}) - \delta(u_2 - \tilde{x})\right] + (2\pi b)^2 \delta(u_1 - \tilde{x})\delta(u_2 - \tilde{x}) \ . \tag{9}$$

It can be shown that when the output plane intensity, and not the mutual intensity, is of interest, the contributions of the second term cancel out. This, and the linearity (in intensity) of the propagation integrals permits us to set

$$t_i(u_1) t_i^*(u_2) \approx 1 + (2\pi b)^2 \delta(u_1 - \tilde{x})\delta(u_2 - \tilde{x}) \ . \tag{10}$$

Multiplying the mutual intensity incident on the object plane by the above transmittance term results in the mutual intensity after the object plane, J'_{obj},

$$J'_{obj}(u_1; u_2) = \mathrm{sinc}\left(\frac{a_0}{\lambda f}(u_1 - u_2)\right)\left(1 + (2\pi b)^2 \sum_{i=-N}^{N} \delta(u_1 - \tilde{x})\delta(u_2 - \tilde{x})\right) \sum_{i=1}^{N} \exp\left[j2\pi \frac{i \sin\alpha}{\lambda}(u_1 - u_2)\right] \ . \tag{11}$$

The portion of interest is the part involving the product $\delta(u_1 - \tilde{x})\delta(u_2 - \tilde{x})$. The other part represents a uniform transparency which will produce a uniform intensity in the output plane, so it is of little interest here. Thus $J'_{obj} = J_{unif} + J_p$, where J_p in the perturbation mutual intensity. The perturbation mutual intensity is a summation of the individual beam perturbations, or

$$J_p(u_1; u_2) = \sum_{i=-N}^{N} J_{pi}(u_1; u_2) \ , \tag{12}$$

where the individual beam perturbation mutual intensities are, from Equation 11,

$$J_{pi}(u_1; u_2) = (2\pi b)^2 \mathrm{sinc}\left(\frac{a_0}{\lambda f}(u_1 - u_2)\right) \delta(u_1 - \tilde{x})\delta(u_2 - \tilde{x}) \exp\left[-j2\pi \frac{i \sin\alpha}{\lambda}(u_1 - u_2)\right] \ . \tag{13}$$

The imaging system processes this input perturbation mutual intensity in the x-z plane according to a two dimensional superposition integral, resulting in the output plane perturbation intensity (one could also compute output plane mutual intensity, but it is not of interest here)

$$I_{pi}(\tilde{x}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J_{pi}(u_1; u_2) K(\tilde{x}, u_1) K^*(\tilde{x}, u_2) du_1 du_2 \ . \tag{14}$$

The scaled and reversed output plane dimension is $\tilde{x} = -Mx$. The amplitude spread function, $K(\tilde{x}, u)$, is given by a Fourier transform of the pupil function[7]. For the imaging system in the schlieren, and ignoring a leading constant, this relation is

$$K(\tilde{x}, u) = \exp\left[j\frac{\pi d}{\lambda f^2}\tilde{x}^2\right] \int_{-\infty}^{\infty} P(\xi) \exp\left[-j\frac{2\pi}{\lambda f}(u - \tilde{x})\xi\right] d\xi \ . \tag{15}$$

For each source and cut-off slit pair in either a traditional or a multiple source schlieren, the pupil function $P(\xi)$ is determined by both the bandwidth of the illumination and the cut-off slit. This arrangement is sketched in Figure 6. Thus the pupil function is a rect-function with a width that depends, in general, on the source slit width, the portion of the image of the source slit cut by the cut-off grid, and to some extent, the deflection of rays by density gradients. In this work it has already been assumed that $b \ll 1$, so the deflection due to density gradients is neglected. This in turn linearizes the problem, permitting simple solution of the final intensity. These assumptions produce a pupil function for the ith beam which is, with reference to Figure 6,

$$P_i(\xi) = \text{rect}\left(\frac{\xi - ia_2 - a_3/2}{a_0 - a_3}\right) . \tag{16}$$

Substitution of this into Equations 14 and 15 results in the ouput plane perturbation intensity, again ignoring leading constants,

$$I_p(x) = \sum_{i=-N}^{N} \text{sinc}^2\left(\frac{a_0 - a_3}{\lambda z_i}(x - Mx_d - iML_d\alpha a_2/f)\right) . \tag{17}$$

The perturbation intensity is approximately the distribution of intensity in the image plane caused by a small disturbance at a general location in the object space.

Note that when this expression is applied to the traditional schlieren ($N = 0$), reduces to

$$I_p|_{N=0} = \text{sinc}^2\left(\frac{a_0 - a_3}{\lambda z_i}(x - Mx_d)\right) . \tag{18}$$

The dependence on L_d has disappeared, demonstrating the fact that a disturbance anywhere between the source slit and knife edge will be detected. The x-direction functional form is simply the incoherent point spread function of an imaging system with an aperture of width $a_0 - a_3$.

To simplify interpretation of the perturbation intensity in the multiple source schlieren, set $x_d = 0$ and let L_d remain general. This constrains the model to considering density disturbances on the optical axis. A non-zero x_d simply shifts the response up or down. Let us further assume that $a_3 = a_0/2$, which means that each slit in the cut-off grid cuts half of the image of the corresponding source slit. This is not unusual for traditional schlierens, and appears reasonable for the multiple source schlieren. The value of $(a_0 - a_3)$ and the integration length are inversely proportional; so it is simple to reproduce the following work for other values of a_3. With these assumptions, the perturbation intensity is reduced to

$$I_p(x) = \sum_{i=-N}^{N} \text{sinc}^2\left(\frac{a_0}{2\lambda z_i}(x - iML_d a_2/f)\right) . \tag{19}$$

This expression is plotted versus x and L_d for $N = 1$ and $Ma_2/f = 0.01$ in Figure 7. In this plot, L_d run left and right, and x back and front (since the sinc-function is even, positive and negative arguments

are equivalent). In the center of the plot, the peak represents the strongest response. That is, the system responds best to a density disturbance in the ideal object plane. This is good news; some axial selectivity has been created with only three slits.

Several features of this three-slit response besides the central peak should be noted because the same features exist in the perturbation intensities for all numbers of slits. The walk-off of the individual squared sinc-functions is analogous to the walk-off of individual images of a fluid disturbance discussed in the geometrical section. This tells us that the shortest possible integration length is governed not by the fluid length scales, but by the diffractive response of optical system itself. Thus the minimum integration length may be considered to be $\min(L_i) = \lambda z_i \theta/(a_0 - a_3)$.

The second item to note is that the response of the multiple source schlieren system never goes to zero for any value of L_d. This is seen in Figure 7, where the central left to right ridge is approximately 1/3 the strength of the peak response. This ridge represents the response of the central pair of source and cut-off slits; a plain old schlieren system with no axial selectivity. Note too the similar ridges for each pair of source and cut-off slits. This is another manifestation of the incoherent multiple source schlieren system acting as a sum of a number of individual traditional schlieren systems. Thus while the integration length may be reduced, the response of the schlieren system to density disturbances throughout the system is not eliminated, only reduced. Actually, it appears that the response is not so much reduced outside the integration length, as it is enhanced within the region.

The third item to note is the strength of the peak relative to the strength of a traditional schlieren response. Whether the axial selectivity is due to an enhancement of one portion or attenuation of another makes no great difference in the small perturbation case. The response to large density gradients, such as shock waves, may shift light from the images of the source slits entirely onto or off of the cut-off slit, thereby saturating the response. Even though the analytical model does not represent large density gradients, one expects that the walk-off phenomena and the central peak shall still exist. Consider a shock wave sufficiently strong to cause a saturated response in a traditional schlieren system with, for example, a source slit width of 1 mm. Then the central ridge in the response of a multiple-source schlieren system with 1 mm slit width will also be (Figure 7) saturated. Clearly then all other ridges will also be at saturation, and the peak could be the sum of $(N + 1)$ bright saturated responses. If the individual systems are saturated at black, then the central peak would simply be black. To adjust for this situation, one could enlarge the slit width to reduce the sensitivity[1], thereby reducing the individual responses below saturation but maintaining the central peak at or near saturation. If this is indeed the case it would be very fortuitous; for sufficiently wide source slits each individual schlieren system within the multiple source system could begin to have an integration

length of less than the mirror separation distance. This would manifest itself in the response plotted in Figure 7 as a decay of the ridges in the regions away from the central peak.

Finally, the issue of whether the inter-beam angle or the maximum beam angle should be used in computing the integration length is addressed. The simple answer is that the inter-beam angle is the most relevant, and that the implications of this merit examination. Figure 8 shows one quarter of the response for a system of 11 pairs of source and cut-off slits. The symmetry of the response in Figure 7 permits examination of just one quarter. Note that if the peak response is normalized to unity, then the individual ridges are at a value of $1/(2N+1)$. Thus to reduce the response by a factor of A outside the integration length, one requires $N \geq (1/A - 1)/2$ slits. For example, a single order of magnitude reduction requires only 5 slits on either side of the center slit, for a total of 11 slits. Then the desired integration length and the relevant length scale, fluid or optical, for the purpose at hand is used to specify the inter-beam angle. The inter-beam angle multiplied by $2N$ gives the maximum beam angle.

4.3 Computer modeling potential

In conjunction with attempts to extend the analytical model, the feasibility of a numerical model is examined. The purpose of the modelling is to consider arbitrary objects, not just small disturbances. The modeling is found to be possible, but was not practical within the summer and on the available computers.

In this model, the mutual intensity is computed step by step through the optical system. This computation starts from the mutual intensity at the source slits (Equation 7). The optical system is described by two basic processes: propagation and thin transparencies affecting the phase and amplitude of the harmonic electric field. For example, mirrors and lenses are represented by thin quadratic phase screens, apertures and lens rims are real-valued transparencies of transmittance zero (opaque) or unity (clear). By properly combining these basic elements, a multiple source schlieren system can be modeled.

At every step through the optical system, the array of values representing the mutual intensity function, $J(x_1; x_2)$, must resolve the entire spatial frequency bandwidth. For the source and cut-off grids, this is not a problem. For the lenses or mirrors, and for the quadratic phase factors in the propagation integrals, identification of the necessary resolution is simple. A lens of focal length f has an amplitude transmittance

$$t(x, y) = \exp\left[-j\frac{\pi}{\lambda f}(x^2 + y^2)\right] . \tag{20}$$

A local frequency approximation shows that the local x-direction spatial frequency, μ, of this phase factor is

$$\mu \approx x/(\lambda f) . \tag{21}$$

Similarly, for the paraxial approximation to the x-z plane propagation integral[7].

$$J(x_1;x_2) = \frac{1}{(\lambda z)^2}\exp\left[-j\frac{\pi}{\lambda z}(x_2^2-x_1^2)\right] \tag{22}$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J(\xi_1;\xi_2)\exp\left[-j\frac{\pi}{\lambda z}(\xi_2^2-\xi_1^2)\right]\exp\left[j\frac{2\pi}{\lambda z}(x_2\xi_2-x_1\xi_1)\right]d\xi_1 d\xi_2 . \tag{23}$$

the local spatial frequency approximation is applied to the two quadratic phase factors. This results in two frequencies, $\mu_1 \approx \xi/(\lambda z)$ and, $\mu_2 \approx x/(\lambda z)$.

The elements of the mutual intensity array are separated by a distance s_1, or in other terms, $\xi = s_1 i_1$, where i_1 is the row or column index. Because both ξ_1 and ξ_2 are in the same direction and in the same plane, the scaling factor is identical, and the label ξ will designate either ξ_1 or ξ_2, and analogously for x. In the x-plane, we can write $x = s_2 i_2$. However, s_1 and s_2 are related by the propagation integral to be $s_2 = \lambda z/(s_1 N)$, where N is the dimension on one side of the array. Thus to resolve twice the highest spatial frequency in the quadratic phase factors (doubling the frequency allows room for freqency addition in the process of multiplying the mutual intensity by the phase factor), the following conditions must be satisfied for m_1 and m_2, defined to be the maximum absolute values of i_1 and i_2

$$\frac{4m_2\lambda z}{N^2} \le s_1^2 \le \frac{\lambda z}{2m_1} . \tag{24}$$

Comparison of the left and right sides of this relation gives a restriction on the maximum indices, and setting $m_1 = m_2$ produces

$$m_1 \le N/\sqrt{8} . \tag{25}$$

This indicates that only 1/8 of the points in the array representing the mutual intensity are actually used to represent the system. The remaining points pad the signal in order to permit better resolution in the Fourier transform. With reference to Equation 24, it is seen that setting $m_1 = m_2 = N/\sqrt{8}$ results in a maximum scaling factor of $s_1 = \sqrt{\sqrt{2}\lambda z/N}$. For visible wavelengths, $z = 10^3$ mm, and $N = 1024$, this gives $s_1 = 0.026$ mm. Thus the maximum extent in the input plane is 9.5 mm. For this situation, the scaling factor for the output plane is $s_2 = \lambda z/(s_1 N) = 0.019$ mm. This limits the extent of the output plane to 6.8 mm. Thus very large numbers of elements are required to produce a computer model which will simulate a typical schlieren system, let alone a focusing schlieren.

Continued work in this area could produce a functional model, but it is likely that computation times would be substantial. Multiplication by the quadratic phase factors and transparencies is an order N^2 process, and the two-dimensional fast Fourier transforms involved are order $N^2\log_2(N)$ process.

5 Conclusions from Analysis

Reduction of the integration length is described quantitatively for small disturbances. It is shown to depend on the length scale of the disturbance. The fundamental behavior of the optical system is shown, for the incoherent small disturbance case, to be a scalar summation of the intensities produced by individual schlieren systems. It is interesting to note that this result is similar to the result one could obtain from geometric optics analysis. However, the result obtained with the mutual intensity analysis provides the additional information about the minimum possible integration length as determined by diffractive effects.

Extension of the small disturbance theory to larger disturbances, such as shock waves, slip lines, compressible boundary layers, compressible wakes and mixing layers, and so forth is difficult. Some general observations may be made. For example, since the integration lengths of the individual schlieren systems persist when the integration length of the system is reduced, a single shock wave could lead to "ringing" effects. That is, a shock wave outside the system integration length could show up in the image as a series of weaker shocks. This type of ringing has been observed by a group at USC. This does not bode well for the hope of probing three-dimensional shock structure, unless the question of enhanced sensitivity is resolved with a favorable result.

A computer model of the mutual intensity analysis for the system can be constructed. Currently it appears that application to even the simplest multiple source schlieren system would be prohibitively time-consuming. Professor Collicott is currently exploring the possibilities of reducing the computer memory and time required by constraining the analysis to shift-invariant imaging systems. This has been shown to be a good approximation to many imaging systems[8, 9]. Additionally, Professor Collicott is considering the possibility of applying commercially available ray-tracing codes for the analysis of the large disturbance multiple source schlieren systems.

6 Recommendations

The simplest change which can be made to reduce noise in the TGF schlieren system is to move the mirrors closer together. As close as 360 inches (twice the focal length) would be practical. In traditional schlieren operation, this will reduce the integration length by decreasing the total optical path length. If the primary noise source is turbulent boundary layers on the windows, and not currents in the room, then this change will not have a noticeable effect on a traditional schlieren. In a multiple beam schlieren, this reduction of optical path length will still exist, and a second benefit will exist as well. With the mirror separation reduced from 558 inches (3.1 focal lengths) to 360 inches (2.0 focal lengths), the maximum possible beam

angle will increase by 50%. This will assist in reducing the integration length, as is seen in the geometric analysis section above. Of course, other considerations such as tunnel access and use of the room may affect the mirror position. Presumably the current choice of mirror separation is derived from a desire to have the mirrors placed safely against the walls of the room.

6.1 Multiple source modification

Design of the multiple source system based on the existing schlieren systems is simple. There is no doubt that the system will reduce noise; this effect has been shown several times[2, 3, 4, 6]. Just how to design a certain level of integration length reduction has not been shown before; it is the goal of the above analysis. It is useful to note that the final imaging lens used after the cut-off grid has no effect on the ability of the schlieren to detect density gradients. For this reason, the simplest arrangement is chosen for illustration purposes. As long as the F-number of the final imaging lens is slightly less than the F-number of the mirrors, any imaging lens and magnification should have identical behavior. Because the above analysis is original, it is at this point unconfirmed by measurement. Such measurements are being undertaken by Professor Collicott and his graduate student, Mr. Terry Salyer, at Purdue University this fall. It is recommended that this work at Purdue be supported by a *AFOSR Research Initiation Program* award. This award will free the graduate student of grading and teaching assistant duties, provide for presentation and publication of results, and other support.

The purpose of initial modification of the schlieren system would be to examine the noise reduction by a small number of source slits. This approach has several advantages over a large number of source slit system. For example, the interbeam angle is larger, vignetting would be minor, and the grid is smaller in extent than the mirrors. This last item will assure that the image of the source slits will likely be of sufficient quality to use a pre-printed cut-off mask. This will maintain simplicity and enhance the robustness of the system. On the TGF schlieren, a grid of 11 horizontal source slits each 1 mm high separated by 15 mm, and about 25 mm wide would be a good place to start. Such a system would show reduced tolerance to noise in the room. As described above in the analysis, the reduction in integration length will vary with the length scale of the density gradients, but all length scales would be reduced; there does not appear to be any danger of enhancing noise. Reduction of noise from room currents may be observed by comparing photographs from the existing and the modified schlieren when the tunnel is not running. It is strongly recommended that the symmetry of the existing schlieren system be maintained; there should be unity magnification in the imaging of the source slits onto the cut-off slits. Of course, this modification of the TGF schlieren to confirm the analysis may tie up the facility for a while. It is recommedend that this initial step be performed at Purdue

University on a scaled-down version of the schlieren system on the Purdue Mach 2 tunnel. This would likely be cost the Air Force less than tying up one of the Wright-Patterson facilities for use in instrumentation development.

One potentially important benefit of the multiple source schlieren arrangement is the potential increase in sensitivity discussed above. This will be investigated further at Purdue during this winter and spring. If indeed a larger source width is possible, then the integration length will be further reduced, or fewer slits will be needed.

Based on the results of the examination of the 11-slit schlieren, it should either be used, or modified to spread out the slits, or to incorporate a larger number of slits. This may be necessary if the primary noise source proves to be turbulent compressible boundary layers on the windows, rather than currents in the room.

References

[1] H. W. Liepmann and A. Roshko. *Elements of Gas Dynamics*. Wiley & Sons, New York, 1957.

[2] R. A. Burton. A modified schlieren apparatus for large areas of field. *Journal of the Optical Society of America*, 1949:907–908, Nov. 1949.

[3] R. W. Fish and K. Parnham. Focusing schlieren systems. Technical Report IAP 999, ARC, Nov. 1950.

[4] A. Kantrowitz and R. L. Trimpi. A sharp-focusing schlieren system. *Journal of Aeronautical Science*, 17(5):311–314, May 1950.

[5] Laurence R. Boedeker. Analysis and construction of a sharp focusing schlieren system. Master's thesis, MIT, 1959.

[6] Leonard M. Weinstein. An improved large-field focusing schlieren system. In *AIAA 29th Aerospace Sciences Meeting*, 1991. AIAA Paper number AIAA 91-0567.

[7] J. W. Goodman. *Statistical Optics*. John Wiley and Sons, New York, 1985.

[8] J. W. Goodman. *Fourier Optics*. McGraw-Hill, New York, 1968.

[9] Jack D. Gaskill. *Linear Systems, Fourier Transforms, and Optics*. John Wiley, New York, 1978.

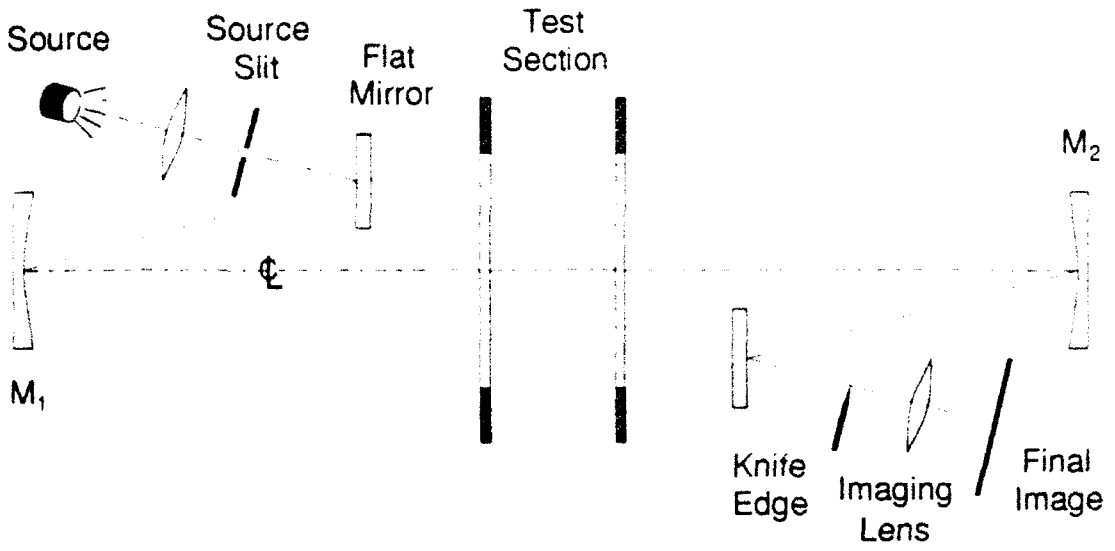


Figure 1. Traditional single-source schlieren system.

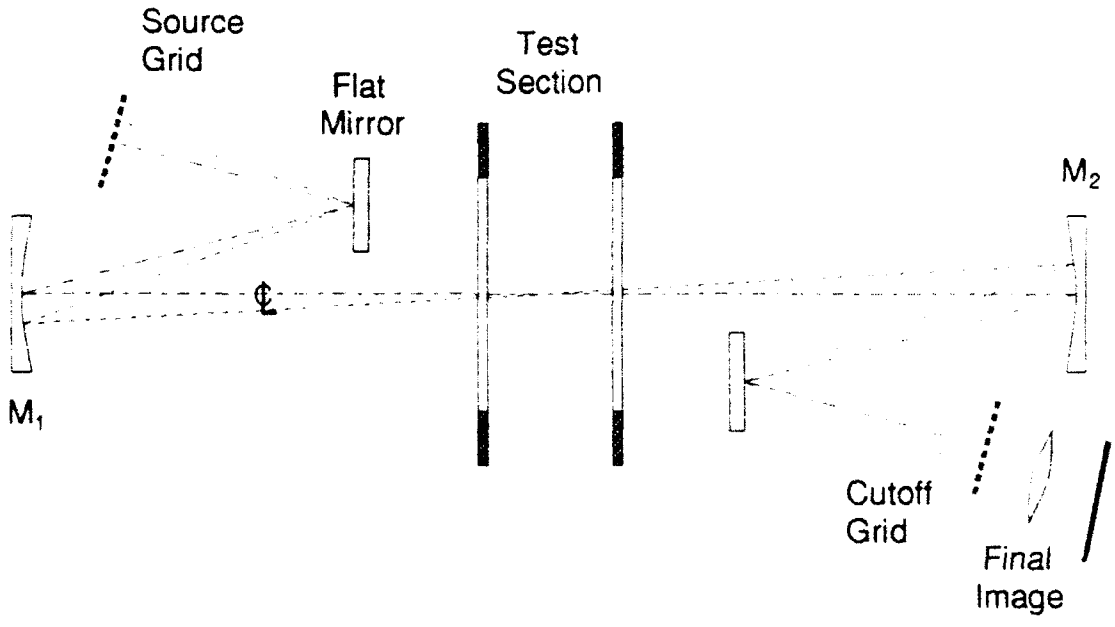


Figure 2. Multiple-source (sharp-focusing) schlieren system.

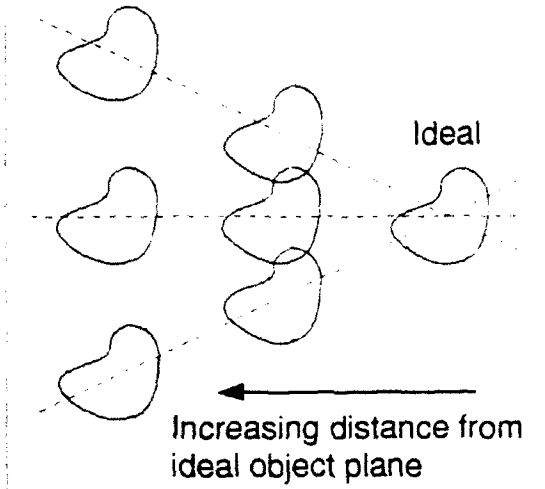


Figure 3. Shift of images from regions outside of the ideal object plane.

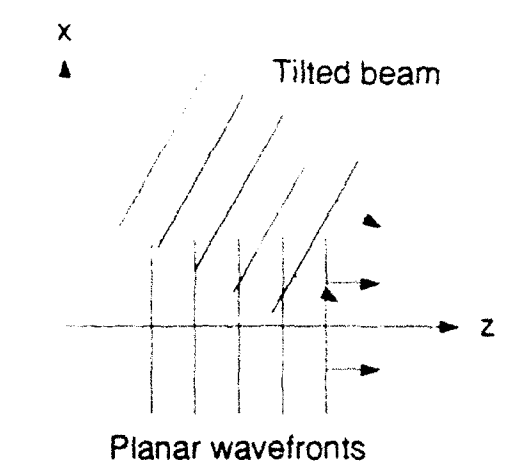


Figure 4. Geometry used to describe incoherent beams.

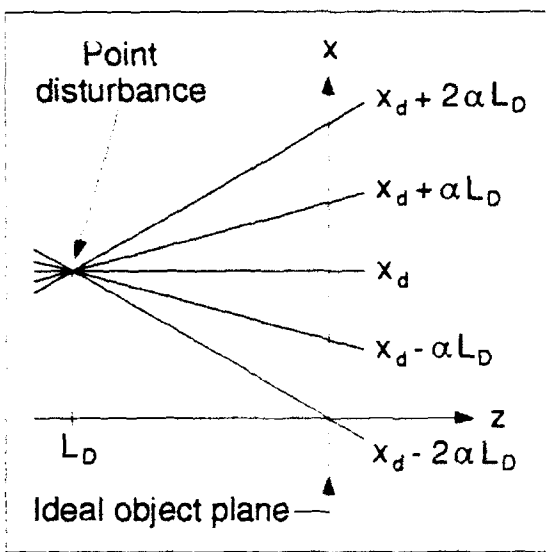


Figure 5. Geometry used to formulate the effective transparency of a small magnitude point disturbance.

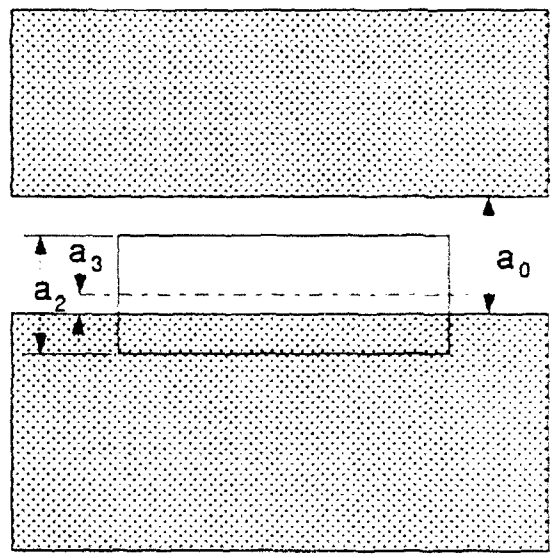


Figure 6. Cut-off grid and image of source grid: limits of integration.

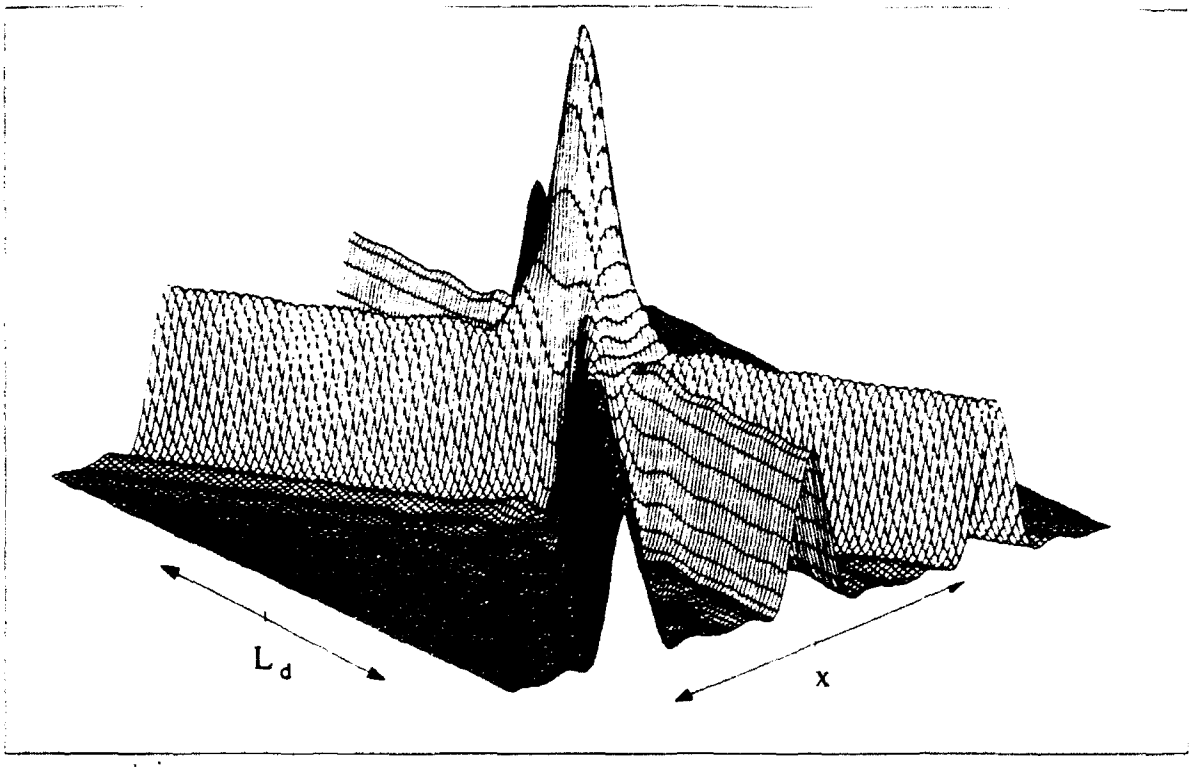


Figure 7. Triple-slit small-disturbance response in TGF system.

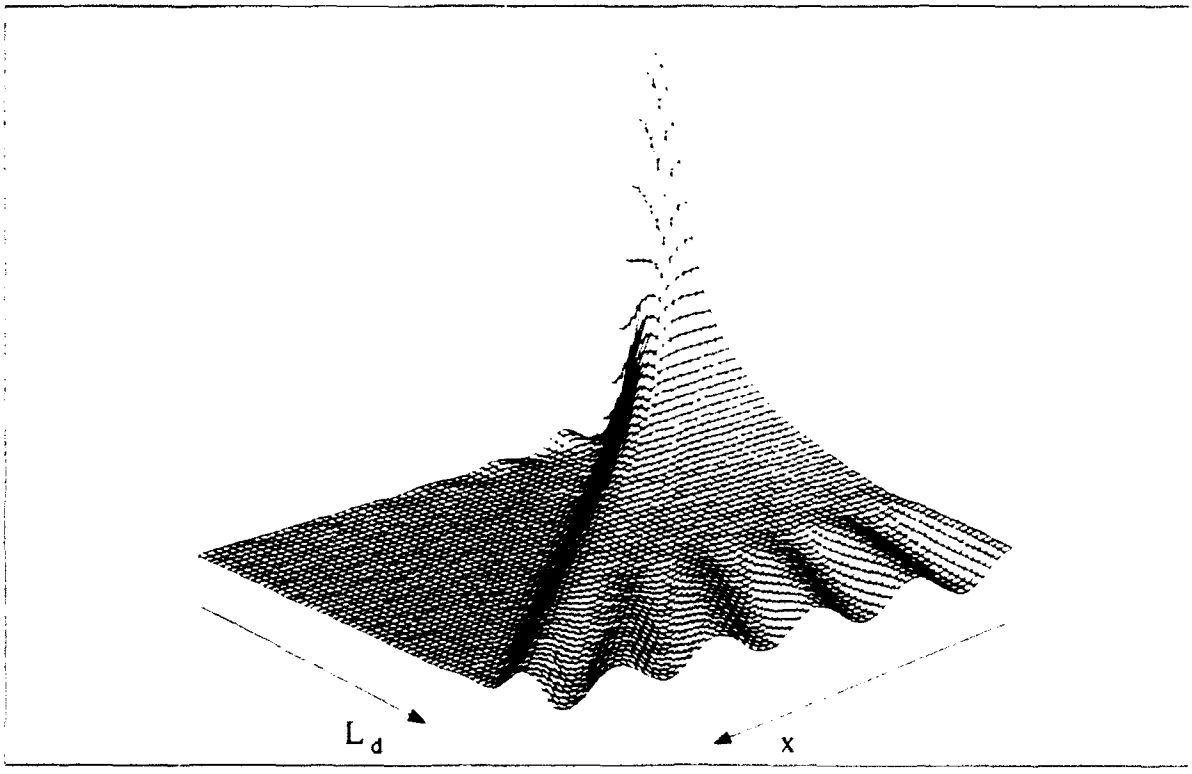


Figure 8. Proposed TGF noise-reduction system: response to small-disturbances.

Computational Studies on Rigid Rod Model Polymer and NLO Model Substances

John W. Connolly
Professor
Dept. of Chemistry

University of Missouri-Kansas City
Kansas City, MO, 64110-2499

Final Report for:
Summer Research Program
WL/MLBP

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base,.Washington, D. C.

September, 1992

Computational Studies on Rigid Rod Model Polymer and NLO Model Substances

John W. Connolly
Professor
Dept. of Chemistry

Abstract

Section 1: Using a heterocycle-phenylene-heterocycle type of model structure, semiempirical AM1 calculations indicate that either removing and electron from or adding an electron to the closed-shell model PBO and PBZT structures brings about a substantial increase in torsional barrier.

Section 2: Semiempirical AM1 calculations of conformational energies using large model structures, i. e. three polymer repeat units for PBO and 2.5 polymer repeat units for PBZT, result in calculated torsional barriers of zero for PBZT and 1 Kcal/mol for PBO for a torsional angle of 90° down the entire model structure. These barriers are much lower than those calculated using small model structures as consequence of cooperativity between the rotating structural units.

Section 3: *Ab initio* calculations of ionization energy of some candidate NLO materials have been begun. The technique involves calculating the difference between the total energy of neutral molecules and the corresponding cation radicals. Structures are optimized at the *ab initio* or semi empirical level depending on size. Benchmark calculations on malononitrile and malonic acid successfully reproduce experimental values.

Computational Studies on Rigid Rod Model Polymer and NLO Model Substances

John W. Connolly

This report is in three sections, each of which describes an individual topic which was worked on during the 1992 Summer Research Fellowship. Section 1 is presented in the form of an abstract of a paper submitted for presentation at the Fall 1992 Materials Research Society Meeting in Boston, the other two sections are more detailed.

Section 1:

AM1 SEMIEMPIRICAL MO CALCULATIONS OF TORSIONAL BARRIERS AND STRUCTURAL REORGANIZATION IN RADICAL IONS OF PBO AND PBZT MODEL STRUCTURES, John W. Connolly, Dept. of Chemistry, UMKC, K. C., MO 64110-2499 and Douglas S. Dudis, WL/MLBP, Wright-Patterson AFB, OH 45433-6533.

Using a heterocycle-phenylene-heterocycle type of model structure, our calculations indicate that either removing and electron from or adding an electron to the closed-shell model PBO and PBZT structures brings about a substantial increase in torsional barrier. Examination of the calculated energy-optimized structures indicates that the closed-shell species undergo changes toward quinoidal phenylene configurations in all cases. These results suggest an explanation of the measured increase in conductivity of PBZT reported earlier[1].

Comparison of results obtained using different model structures will be discussed.

1 DePra, P. A., Gaudiello, J. G., and Marks, T. J., *Macromolecules*, 1988 , 21, 2297Part I: Torsional Cooperativity as Demonstrated by Semi-empirical Calculations on Large Rigid Rod Model Structures.

Section 2: Torsional Cooperativity, an Important New Concept in Polymer Chemistry.

An important aspect of the conformational properties of the rigid rod polymers PBO and PBZT is the barrier to rotation about the carbon-carbon sigma bonds in the polymer backbone. The computational efforts at estimating the barrier to this rotation have focused on rotation about one bond or two equivalent bonds[1]. Figure 1 illustrates the approach we have taken which has led us to an important insight.

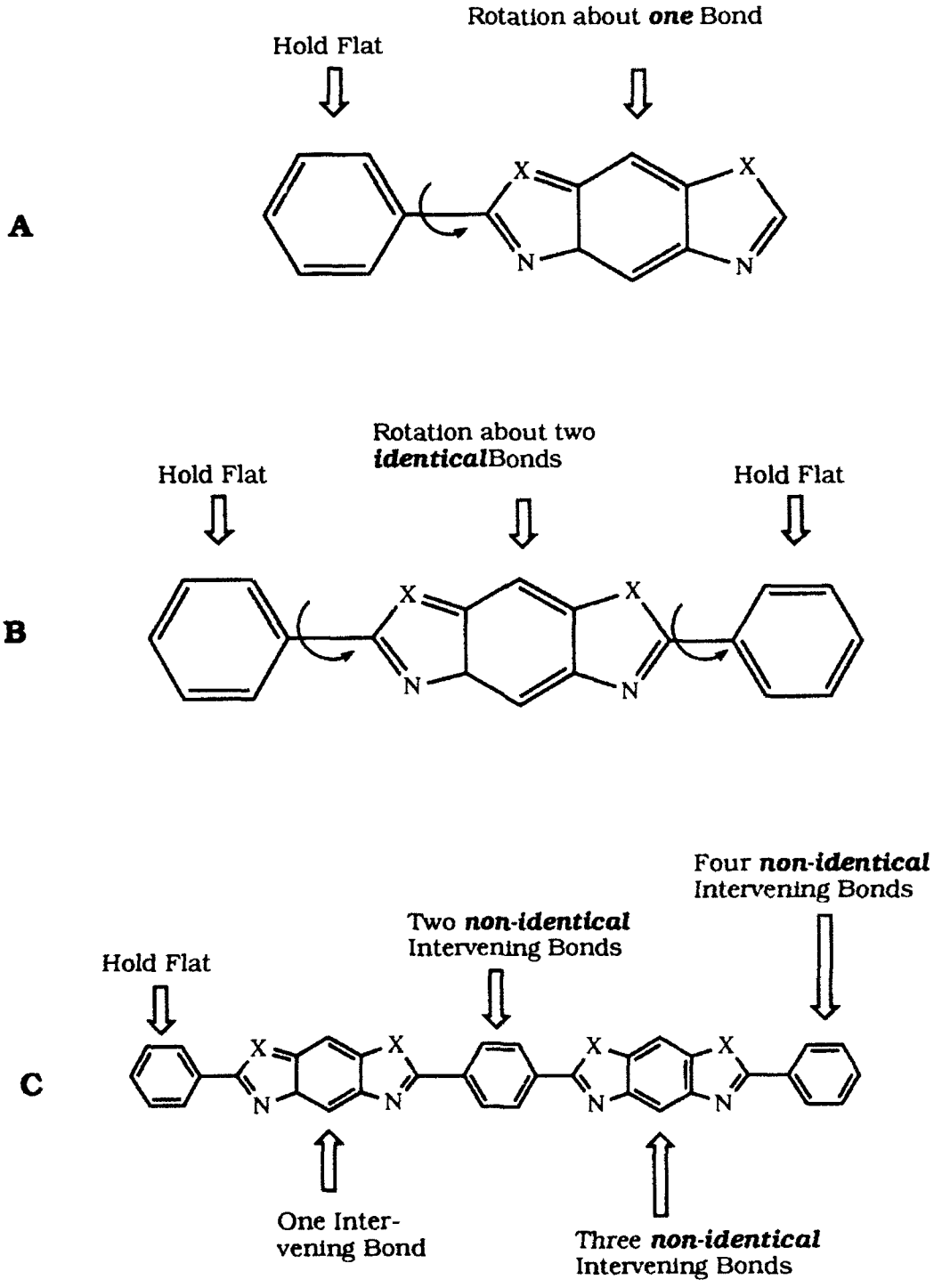
The structures A and B in Figure 1 are those used to calculate torsional barriers in PBO and PBZT. Implicit in the use of these structures is the idea that each bond in the polymer backbone rotates independently of all others. Of course that is not very realistic in conjugated polymers such as these. Structure C in Figure 1 is large enough that it is possible to consider the effect of a given bond rotation on the rotation of adjacent bonds. Interaction between adjacent groups as they rotate is what we refer to here as *torsional cooperativity.*

Using structure C as a model for PBZT we have done a series of torsional barrier height calculations. In each series of calculations the end phenyl group is held flat while some other part of the structure is rotated. (Torsional barriers are calculated as the difference between the minimum energy structure and the structure with the highest energy. In all cases the highest energy structure was the structure in which the torsional angle was 90°.) In the first calculation using structure C we rotated the heterocyclic group adjacent to the end phenyl group; the resulting torsional barrier was identical to that for structure A. Next we calculated the barrier height for rotating the phenylene group is structure C. The structure is now undergoing torsion about two bonds simultaneously. The heterocyclic group between the end phenyl and the rotating phenylene does not track with the phenylene group but takes on an intermediate torsion angle. This minimum energy configuration is one in which the loss in conjugation over the three units, phenylene-heterocycle-phenylene, is minimized. The important consequence of this is that the calculated torsional barrier is *reduced substantially* from that for adjacently rotating groups. We then considered the rotation of the

fourth unit in the model, which requires torsion over three bonds and finally we considered the rotation of one end of the model relative to the other, which involves torsion over four bonds.

Figure 1

Comparison of Small and Large Model Structures



The energetics of each of these torsional processes is shown, for the cis and the trans structure, in Figure 2. As can be seen, for rotation over four bonds in this PBZT model structure, there is zero calculated torsional barrier. That means that any minimum energy conformation beginning with the totally flat conformation and ending with the conformation in which the two ends are perpendicular are energetically equivalent.

The results of our calculations on PBO model structures are also shown in Figure 2. All previous calculations on PBO model compounds have shown it to have a higher barrier to rotation than PBZT. Here we find that torsion even through five bonds does not produce a zero torsional barrier, but is appears that if the model were made sufficiently large a zero torsional barrier would be found for PBO also.

In these calculations we have determined the increase in energy (torsional barrier) which results from twisting one end of a polymer segment until it is approximately 90^{o} relative to another segment. Even though that is not the highest energy conformation possible in these model structures it is the limit of our computational model. For example, if the torsion angle is increased to 100^{o} the energy optimization process simply produces the mirror image of the 80^{o} torsion angle conformation, so the energies of the conformations beyond 90^{o} torsion angle are the same as the energies of the conformations approaching 90^{o} torsion angle. This does not limit the validity of these results since we have, in effect, calculated the increase in energy created by having segments of the polymer chain at right angles to each other.

Figure 2

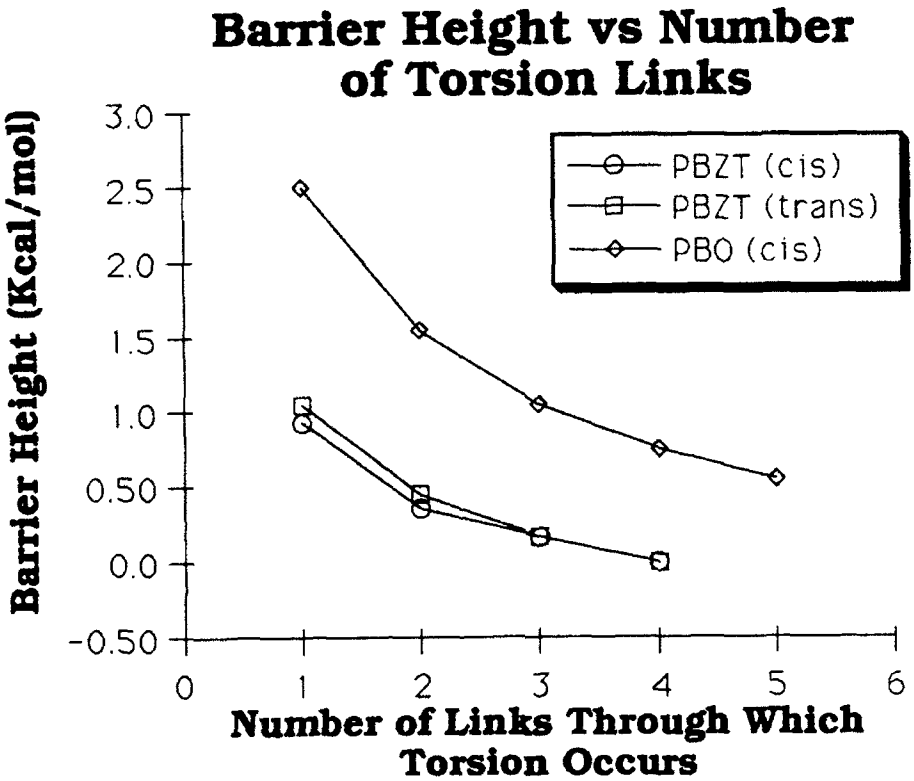
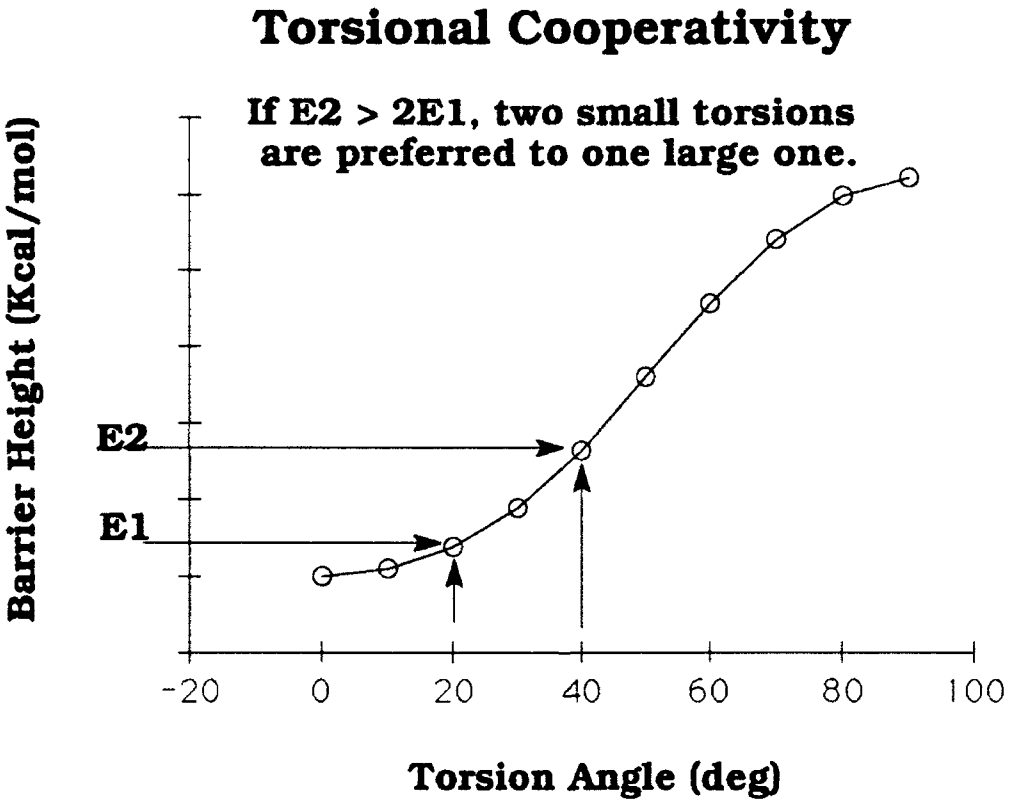

Barrier Height vs Number of Torsion Links
Barrier Height (Kcal/mol)
3.0
2.5
2.0
1.5
1.0
0.50
0.0
-0.50
0
1
2
3
4
5
6
PBZT (cis)
PBZT (trans)
PBO (cis)
Number of Links Through Which Torsion Occurs

Figure 3 shows a general requirement for torsional cooperativity. If torsion through any given angle requires more than half the energy for torsion through half that angle then two half-angle torsions will be energetically preferred to one full angle torsion. Our calculations are limited to conjugated polymers but torsional barriers tend to be sinusoidal in general so we would expect to see this effect in other types of polymers as well.

Figure 3



1a) Welsh, W. J. and Yang, Y., *Comp. Poly. Sci.*, 1991 , 1, 139

1b) Farmer, B. L., Wiershke, S. G., and Adams, W. W. *Polymer*, 1990 , 31, 1637

1c) Connolly, J. W. and Dudis, D. S., accepted for publication in POLYMER

Section 3: *Ab initio* Ionization Energy Calculations of Some NLO Candidate Materials.

In the presence of the intense electromagnetic radiation produced by a laser it is known that some materials exhibit, in addition to the first order properties of absorption and reflection, higher order properties such as frequency doubling, frequency tripling, intensity dependent refractive index, etc[1]. Such nonlinear optical (NLO) materials are the subject of a world-wide research effort involving many disciplines. The computational research proposed here,which has already been begun, is in support of a synthetic chemistry effort at WL/MLBP to make high temperature stable materials which have large χ^2 (the frequency doubling parameter) values. Large χ^2 candidate materials should have a large dipole moment, a large transition moment for the appropriate electronic absorption and be noncentrosymmetric. High χ^2 values have been found in compounds containing the benzthiazole group; that, and its high temperature stability make it a good candidate for inclusion into compounds to be screened for high χ^2 values. Thus compounds containing several benzthiazole groups attached to a saturated carbon atom have been synthesized in the project.

The benzthiazole group plays a significant role in a separate synthetic effort at WL/MLBP, one aimed at producing 'self doping' polymers which will have high intrinsic conductivity. The synthetic effort is aimed at producing PBZT type polymers containing 'spacers' consisting of a saturated carbon atom attached to one other electron withdrawing group. Thus the remaining hydrogen atom on the saturated carbon may become acidic enough to transfer (as a proton) to an adjacent basic site in the polymer, which may produce intrinsic conductivity in the polymer.

The present proposed research is computational and meant to support both the above projects. In both cases the electron withdrawing power of the benzthiazole group is of paramount importance. At present, candidate compounds for both projects are being screened by measuring the chemical shift of methylene protons located between benzthiazole groups and other electron withdrawing groups in the candidate compounds. We have begun to calculate the ionization energy of the

same compounds, in an effort to see if the ionization energy correlates with the methylene chemical shift data. The reason to consider this possibility is that both the chemical shift value and the ionization energy sample the tightness with which electrons are held in a molecule. However this sampling is done in a very different way. The chemical shift of the methylene protons is affected by the polarity of the C-H bond in addition to the ring current of the nearby heterocyclic rings. The ionization energy is a measure of how tightly the electrons are held in the highest occupied molecular orbital (HOMO) of the molecule. Since the HOMO may or may not involve the methylene protons the chemical shift approach and the ionization energy approach may look at different parts of the molecule.

For the compounds already synthesized if a correlation is found between measured chemical shift values and calculated ionization energies we will be in a position to suggest potential candidate materials by calculating ionization energies of compounds which do not yet exist. If a correlation is not found between the ionization energies of the compounds already synthesized and their chemical shifts it may be possible to find a correlation between one of the orbital energies in the neutral molecule, some of which are primarily C-H sigma orbitals, and the chemical shift.

The present status of this work is summarized in Table 1. The first two entries, malononitrile and malonic acid are primarily benchmark compounds from a computational standpoint. Their chemicals shifts are presumably a measure of the acidity of the methylene protons.

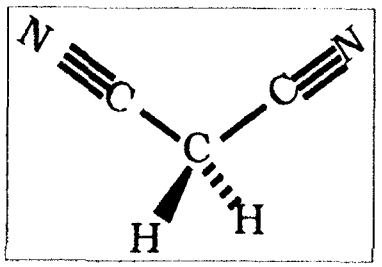
Using a rather straight-forward but high level calculational approach we have been able to successfully calculate the ionization energy of these two materials. Our approach is simply to take the difference between the calculated *vide infra* total energy of the neutral species and the cation radical. Since these differences are approximately 0.1% of the total energies of these compounds this attests to the accuracy of these calculations. We also see a correlation between the ionization energies of these compounds and the chemical shifts of their methylene protons. For the third entry, benzothiazolacetonitrile, the

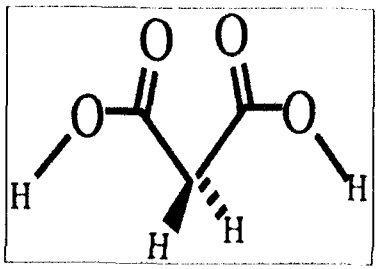
ionization energy has been calculated (by us) but not measured. As can be seen the correlation with the chemical shift does not hold here. Of course we would need to finish this work to see if this apparent lack of correlation is 'scatter' or real. In any case a detailed description of the electronic structure of the benzthiazole-containing compounds in Table 1, which would be a byproduct of this work, would be valuable in assessing their potential as NLO materials.

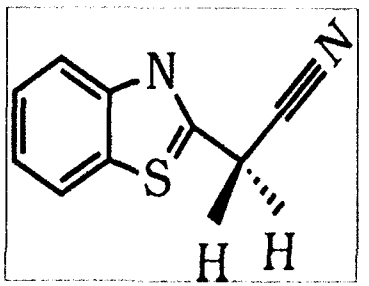
Calculational details: The *ab initio program* package GAMESS[2] has been used in this study. Ionization energies were obtained by taking the difference between the total energy of the neutral molecule and the total energy of the corresponding cation radical. Basis sets at the 6-31g* level have been used for all calculations. The smaller neutral structures were optimized at the *ab initio* level, the larger ones were optimized using the AM1 semi empirical hamiltonian in the MOPAC[3] program package. Since the ionizations being calculated are vertical (no structural change) the same structure is used the neutral molecule and the cation radical.

The first two calculations in Table 1 serve to benchmark our calculational procedures. The energy differences which we are calculating are less that 0.1% of the total energy and therefore each calculation must be very accurate for these results to be of any value. As can be seen from the table, the experimental ionization energies of malonotrile and of amlonic acid are reproduced to within 0.7 eV by these calculations, which is within one order of magnitude of the precision of the measurement. On the basis of these results we can have confidence e in the calculational results were obtained which are not supported by measurement.

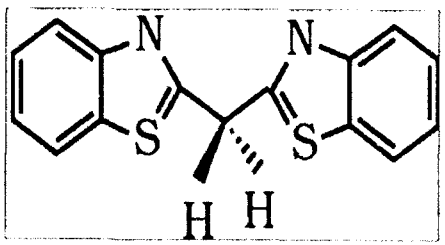
Table 1

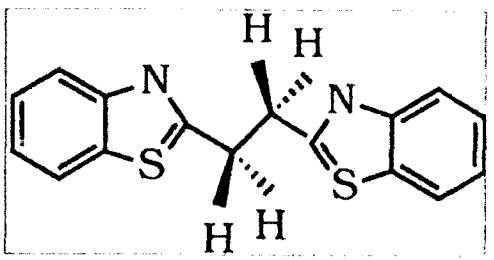






IE (ev, exp)	12.68*	11.05**	
IE (ev, calc)	11.92	11.39	7.88
Chem Shift	4.39***	3.24***	4.77***





IE (ev, exp)		
IE (ev, calc)		
Chem Shift	4.96***	3.74***

* ref 4
** ref 5
*** determined at MLBP

References:

1) Prasad, P. N. and Williams, D. J. *Introduction to Nonlinear Optical Effects in Molecules and Polymers*, Wiley, New York, 1991

2) Schmidt, M. W., Balkrdige, K. K., Boatz, J. A., Jensen, J. H., Koseki, S., Gordon, M. S., Nguyen, K. A., Windus, T. L., and Elbert, J.: **QCPE Bulletin**, 10, 52, (1990)

3) Quanutm Chemistry Program Exchange, Indiana University, Bloomington, IN 47405, #506

4) Bock, J and Statfast, H., **Z. Naturforsch.**, 28b, 746 (1973)

5) Ajo, D., Ciliberto, E., Fragala, I., and Granozzi, G., **J. Molecular Structure**, 62, 189 (1980)

Performing Target Classification Using a Fuzzy Morphology Neural Network

Jennifer L. Davidson
Assistant Professor
Department of Electrical Engineering

Iowa State University
319 Durham Center
Ames, Iowa 50011

Final Report for:
Faculty Summer Research Program

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

August 1992

1. Performing Target Classification Using a Fuzzy Morphology Neural Network

Jennifer L. Davidson
Assistant Professor
Department of Electrical Engineering
Iowa State University

Abstract

This report details a neural network application to target classification using a different type of neural network called a fuzzy morphological neural network. Neural networks are used mostly as pattern recognizers, and the main goal of this research was to give proof of concept that a simple form of this new neural network could perform object classification. The type of data used was the range data from ladar data on tanks. A two-tiered approach was used on the data. First, a linear convolution with a fixed template was performed. Then this data was input to a neural network for classification. The image was divided up into zones, with the expectation that a partially occluded object could still carry forward enough information from fewer zones to the neural network to identify the object. The simplified neural network was essentially a multilayer perceptron with backpropagation, so this network was used to train the data. The results showed that with proper training, the network can perform very good classification. The numerical results are given in this report, and substantiate our initial ideas. With the proof of concept in hand, the next step is to develop training algorithms and network architectures which take advantage of the complex range of parameters available in the network. This is the main topic of research described in a follow-up minigrant proposal.

Performing Target Classification Using a Fuzzy Morphology Neural Network

Jennifer L. Davidson

1. Introduction

The Armament Lab at Eglin Air Force Base, Florida, has been dedicated to performing seeker technology for a number of years. Crucial to this research development are solutions to the autonomous target classification problem. Initial development of algorithms is supported under basic or 6–1 research and often supplemented by visiting faculty, such as those on the Summer Research Program. Some types of data that the military are interested in are Forward Looking Infrared (FLIR), ladar, and millimeter wave (MMW). Each type of data has its advantages and disadvantages. For example, FLIR imagery is not too expensive, but cannot be used as all weather data since it cannot see through foggy or rainy weather conditions. The same holds for ladar data. Adverse weather conditions can dramatically reduce the effectiveness of these sensors to collect meaningful data. However, both FLIR and ladar sensors can obtain fairly high resolution data, an advantage over MMW data. The latter currently does not have as high lateral (vertical and horizontal) resolution as the first two types. One main advantage of MMW over FLIR and ladar sensors is that MMW can penetrate adverse weather conditions to produce an image of the scenery that contains a reasonable amount of information. However, current MMW sensors are expensive and, together with its low resolution, makes it not the best all round choice of sensor. One way to combine the advantages of both types of data is to group both sensors on one unit and use the multi-sensed data gathered from both types of sensors to classify targets. Since data fusion techniques developed for one type of data often extend to multi-data processing techniques, it is useful to have algorithms which can classify targets for one type of data. To this end, this report presents research on a target classification algorithm for ladar data using a neural network approach. The type of targets we are interested in are tanks. The ladar data was chosen due to its low expense and high resolution.

There are many approaches to target classification. These include segmentation, syntactic, and various pattern recognition techniques, including artificial neural networks. Neural networks have been applied successfully to different types of problems, producing acceptable results in many cases. While rigorous mathematical foundations have not been laid for much of neural networks, nonetheless they have seen success in many applications areas; see [8] for a list of applications. The type of neural network used in this report for target classification of tanks is a different type of network that, in theory, can model a wide range of mappings from linear to nonlinear, depending on the parameters of the network. The initial goal of this research was to show proof of concept that this particular neural network could solve a target classification problem for the military, which we believe has been achieved. Future research is presented in the minigrant proposal, and will allow a much more detailed investigation to be carried out.

In order to save time encoding the algorithm, two software packages were used. One, the IAF FORTRAN Preprocessor [3], was used for most of the image processing except for the neural net part, where a software package call NeuralWare was used [5].

The rest of the report is organized as follows: the next section discusses the data available for our purposes and the type of data that was selected. Section 3 presents a discussion of neural networks and their role in target classification as well as a detailed discussion of the particular network used. Section 4 discusses the preprocessing of the data and gives an outline of the algorithm used. Section 5 presents a discussion of the results, and the last section presents conclusions and directions for future research. Acknowledgments of the help of base and other personnel is noted, followed by the references.

2. The Data

The type of data available for this summer research included FLIR, ladar and MMW data. Ladar data consists of two pieces of information at every pixel location. One is the value, in some kind of units, of the distance from the sensor to the sensed object in the real world, and the other value is the intensity of the returned signal. Often it is easier to see more detail in the intensity image, as it exhibits greater visual

contrast. Nonetheless, it is the range data that contains much more information. From the range data an elevation or height image can be constructed using straightforward geometric and trigonometric relations. Using a ground level for a reference, actual shapes of the objects in the image (disregarding shadowing) can be displayed as a 2D surface in 3 dimensions. In addition, the range data has excellent horizontal and vertical resolution, typically on the order of 6" or a foot; the range accuracy (along the line of vision) is just as good or better. However, for our purposes the data is more useful in elevation form, so the first step is to create an elevation image. Data gathered by Schwartz was available in height form, so we used this data for our research. In future research as presented in the minigrant proposal, we plan to use data gathered by LTV, and develop our own algorithm to produce the elevation data. To our knowledge the LTV data has not been utilized much and since it is visually very good, we propose to focus on this data during research performed on the minigrant. The Schwartz data did not have as high resolution as the LTV data, however, so we expect better results using the LTV data in future research efforts.

The Schwartz images are each 24 x 60. A subset of the image of size 19 x 13 is extracted, and it is this data that is used for classification. The procedure, once it is automated, is to extract a possible target, scale to a pre-determined size (such as a 19 x 13 array), and perform the classification on the scaled data. This avoids the problem of varying perspectives and ranges from the same object resulting in different sized images of the same object.

Preprocessing was performed on each tank image. First, a histogram equalization of the image was calculated, followed by the extraction of a subimage of size 19 x 13, followed by a linear convolution on the 19 x 13 subimage. The linear convolution reduced each 19 x 13 array to a 5 x 4 array. Two sets of subimages from the original image were extracted. One was one image of the tank to be identified, and the other set was 20 subimages of nontank noise extracted from the original image. To the one image of the tank was added random uniform noise to create 20 additional noisy tank images. Thus, a total of 41 images were available to be used in the classification scheme, of which 40 were used. The original tank image was not used. Each image was reduced to a 5 x 4 array through the linear convolution. The

5 x 4 arrays were used as inputs to the neural network.

3. Neural Networks and the Model used for this Research

Artificial neural networks are most often used as pattern classifiers. Essentially a neural network performs (after training) a mapping from the input space of objects to the output space of classification. The way that these mappings are determined are typically through some optimization process whereby the error between the input and the known output is minimized. This phase, called the training stage, must generally be conducted using data which, in order to get good results, is statistically representative of the data to be classified. In addition, neural nets are often fairly good at determining complex relationships between input data and the classification representations. This was the main reason we chose to solve our classification problem using a neural network classification. While neural networks do not have a very rigorous or broad mathematical foundation, they have been applied to solve many practical problems and have produced very acceptable results in many cases. The system used to classify our data consisted of a preprocessing part where a histogram equalization was applied, followed by a regular linear convolution; Then the neural network was used to classify. Thus, a convolution was performed on the data, producing output, which was then input to a neural network for classification. The convolution template values were chosen by hand; in future research, the convolution values will be "learned" by the network as well.

As mentioned above, the image data is convolved with a template and it is this data that is used as input into the neural network. Only a subset of all points in the image are transformed by the convolution. This is done to save computation time. For the subimage of size 19 x 13, 20 points were chosen to be convolved on. It is these 20 points, post-convolution, that are input to the neural network. The 19 x 13 array was divided into 4 overlapping subarrays, or zones, of size 10 x 7. Each zone represents a part of the tank object to be classified. From each zone there are four distinct convolution points, and there are 4 more points which lie in the overlapping part of 2 zones. See Fig 1 for a picture of the zones and the 20 convolution points. The reason for dividing the image into zones is to help identify partially

occluded objects. The idea is that if enough information is contained in the points not occluded, then the object can be identified. This is especially important in range data due to the fact that "shadowing" can occur behind the object; See Fig. 2 for a picture.

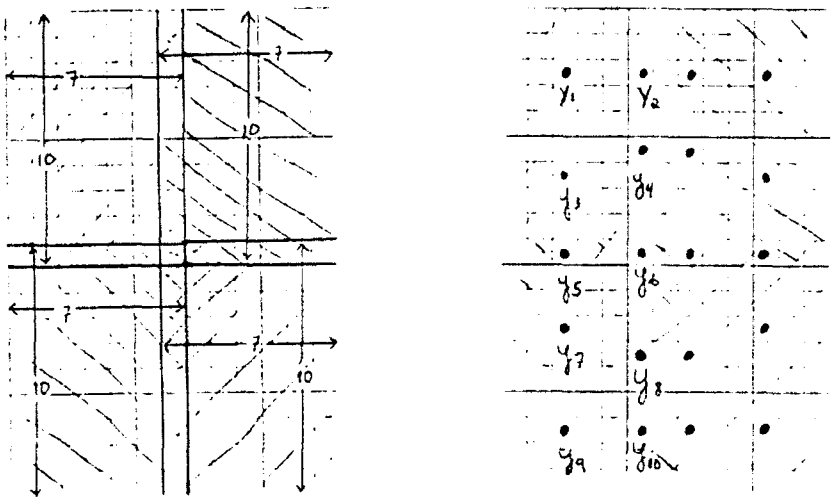


Fig. 1. Four 10 x 7 Zones and 20 Convolution Points in the 19 x 13 subarray.

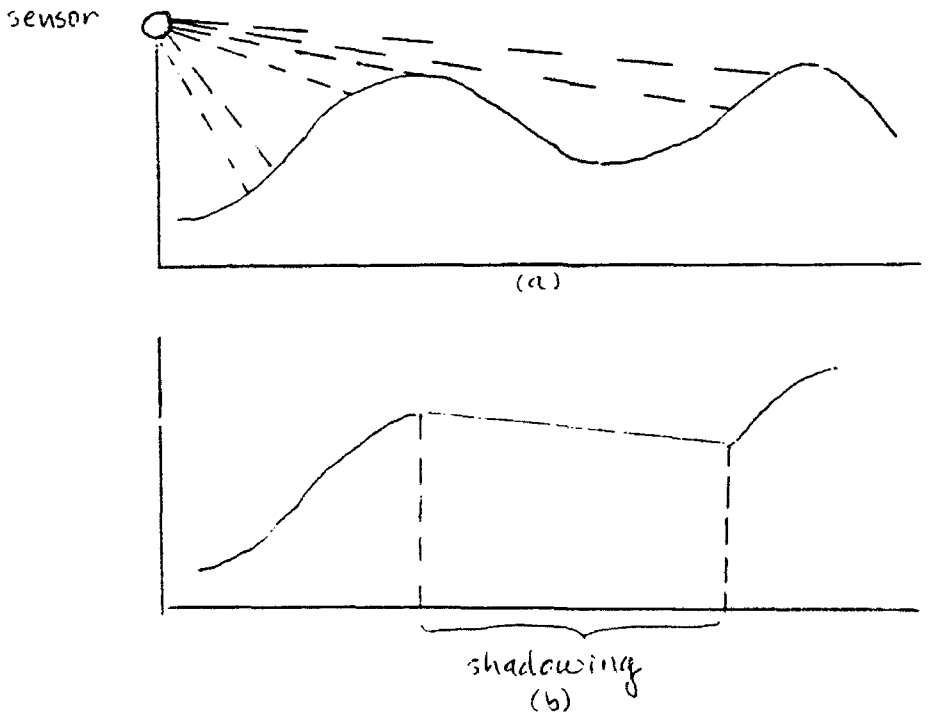


Fig. 2. Shadowing in Range Images. (a) Original Scene. (b) Height Image.

We next describe briefly neural nets, and then explain in more detail the particular model used in this research. For more information on neural nets, see [4].

Artificial neural networks have been introduced recently to provide alternative solutions to a class of mapping problems. One class of such problems is concerned with finding a function or mapping between an element $a \in A$ and $b \in B$, or, more generally, between a subset $\mathcal{A} \subset A$ and a subset $\mathcal{B} \subset B$. Classical neural network architectures, such as the multilayer perceptron, have a variety of iterative procedures associated with them, such as the backpropagation method, a gradient-descent method, and simulated annealing [2]; both of these are optimization processes that approximate the mapping that classifies a selection of training data. While these neural networks can approximate a wide range of linear and nonlinear mappings, all use the following basic algebraic scheme to combine weights w and nodal values a at each neuron or PE:

$$c_j = f\left(\sum_{i=1}^{N} a_i w_{ji}\right).$$

Typically the function f provides some type of nonlinear thresholding. A new type of neural network, called the fuzzy morphology neural network, has been introduced [1], which algebraically combines the information as

$$c_j = \left[\sum_{i \in S_0(w_j) \cap S_1(t_j)} w_j(i)[r(1 - t_j(i) - a(i))]^p\right]^{\frac{1}{p}}. \tag{1}$$

Here, w and t are templates (in the image algebra sense [7]), p is a parameter allowed to be any real number except 0; a represents the input vector; the sum is over the supports of w and t, and r is the following function:

$$r(z) = \begin{cases} z & 1 \geqslant z \geqslant 0 \\ 1 & z \geqslant 1 \\ 0 & z \leqslant 0 \end{cases}.$$

The support for a template is defined by

$$S_d(t_j) = \{i \in \{1, 2, \ldots, N\} : t_j(i) \neq d\}.$$

We omit a detailed discussion of image algebra; for further information consult [7]. The function r is the "fuzzy" part of the neural network, allowing a probabilistic classification of its argument to be considered.

This neural network differs markedly from the classical ones due to its distinctly different algebraic "flavor." In fact, it has been shown [1] that if p = 1 and $t_j(i) = 1$ for all i and j, and the values for a are between 0 and 1, then the output c is simply a linear convolution: $c = a \div w$. (We use the image algebra notation throughout the paper.) Also, if $w_j(i) = \frac{1}{N_j}$ for all i and j and p is allowed to approach the value minus infinity, then c approaches the mathematical morphology erosion $a \boxminus w$. Here, $N_j = |S_0(w_j) \cap S_1(t_j)|$. Thus, for varying values of the parameters p, w, and t, this equation can represent a transform varying between a linear one and a highly nonlinear one, the morphology erosion. This is the "morphology" part of the neural network.

Next, we present the particular architecture used. To simplify the model and show proof of concept, the following parameters were chosen: p = 1; $t_j(i) = 1$ for all i and j; and the values for a are allowed to vary between 0 and 255. This is basically a linear convolution, and in fact, is very similar to the backpropagation network. Thus, we used the multilayer perceptron with backpropagation to implement this first stage of the research. The input layer has 20 nodes representing the 20 points of convolution, the first hidden layer has 8 hidden nodes, the second hidden layer has 3 nodes, and there is one output node representing the output, a tank or not. Between layers the nodes are fully interconnected. See Fig. 3 for a picture of the architecture of this network. The software used to implement the neural network was the *NeuralWare Professional II* software package for the Sun workstation [5].

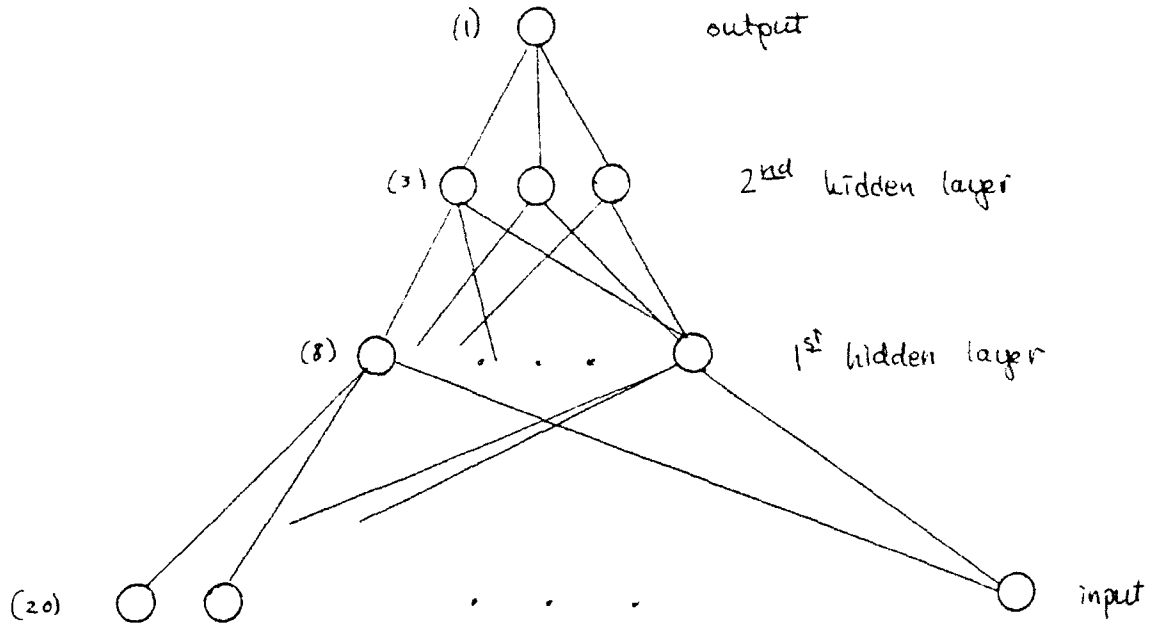


Fig. 3. Neural Net Architecture for Tank Detection.

4. Processing of the Data and a Flowchart of the Algorithm

It has been shown [6] that image algebra encompasses an environment for expressing neural network algorithms. This is why we chose to use image algebra to represent the pseudocode in, as well as the IAF Preprocessor, an image algebra preprocessor in FORTRAN, to implement other parts of the code besides the neural network part. Using the IAF preprocessor, the data was histogram equalized. Then a linear convolution was performed using the template $t \in (\mathbb{R}^{\mathbf{X}})^{\mathbf{Y}}$ as defined in Fig. 4. Here, $\mathbf{X}$ = 19 x 13 array, and $\mathbf{Y}$ = 5 x 4 array. The template used for this convolution was symmetric with respect to the vertical axis of the array $\mathbf{Y}$. In Fig. 4 only half of the template images are defined. For each of the 10 labeled points in Fig. 1, $y_1 - y_{10}$, the remaining template's images, $y_{11} - y_{20}$, are defined symmetrically with respect to the vertical line dividing $\mathbf{Y}$. Thus, $y_{11} = y_1, y_{12} = y_2$, etc. The input to this step was the image on the 19 x 13 array, and the output was the 20 convolved points, an image on the 5 x 4 array $\mathbf{Y}$.

Next we present a global description of the entire algorithm, from input of the original data to the output, a classification. The flow chart is given in Fig. 5.

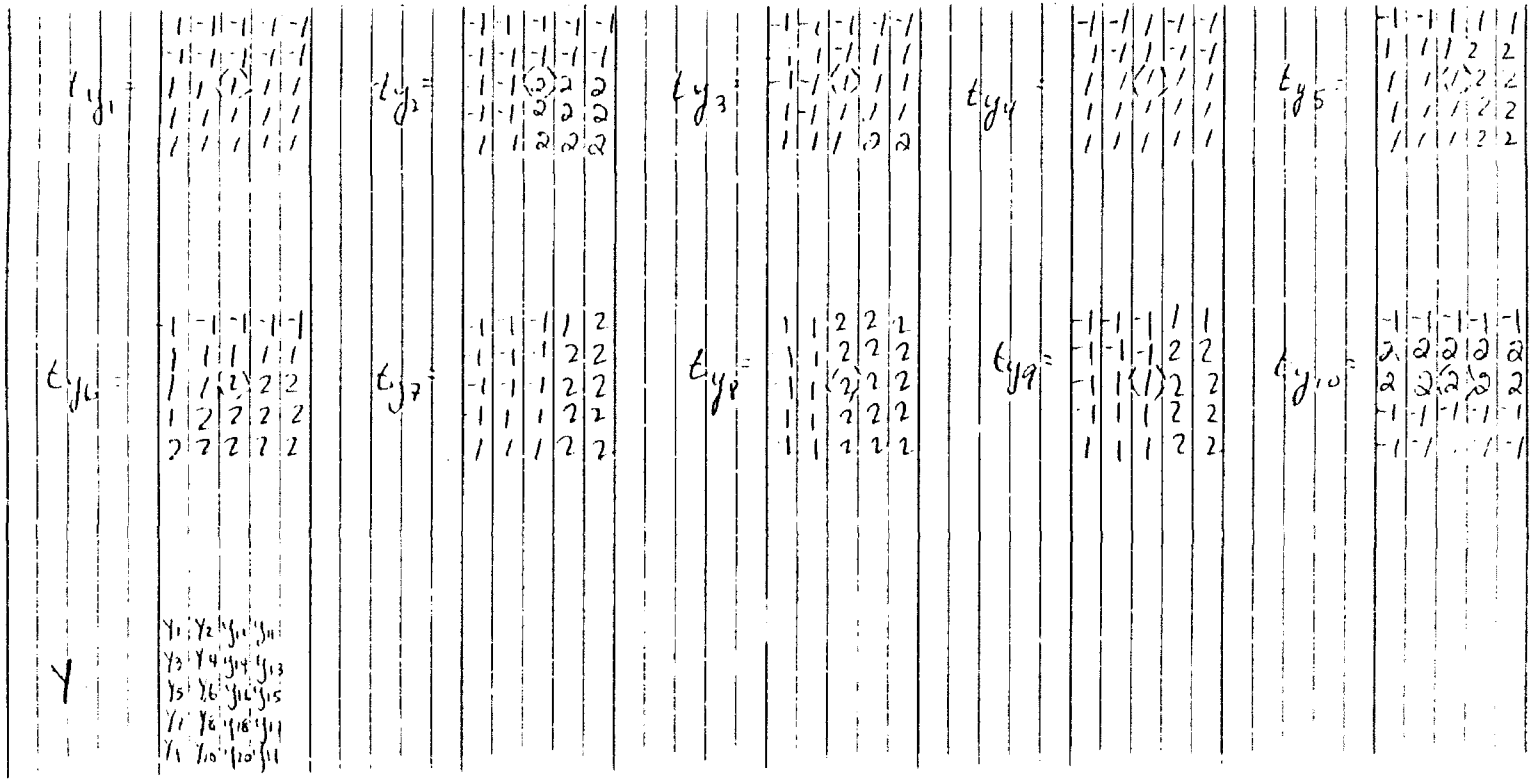


Figure 5. Template $t \in (\mathbb{R}^X)^Y$, used to extract 20 Convolution Points

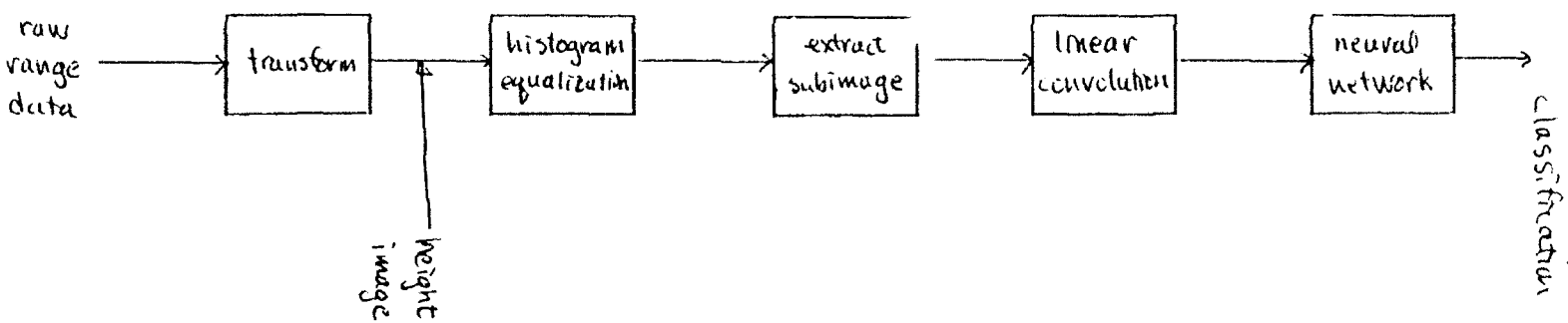


Figure 5. Flowchart of Algorithm

5. Results

The backpropagation network gave various performances on the data. When trained on all 40 data (20 tank and 20 nontank) then recalled on all 40, the classification was not 100% correct after 5000 iterations, but was 100% correct after 10000 iterations. When training was performed on a subset of the training data, the results were similar. When 5 tanks and 5 nontanks were used as the training data, after 10000 iterations, the net classified only 6/20 or 30% of the nontanks as nontanks, with the remaining 14/20 or 70% of the nontanks classified as tanks. While it classified all tanks as tanks, the false alarm rate, that is, the identification of a nontank as a tank, was very high, 70%. When the training set was increased to 7 tanks and 7 nontanks, after 5000 iterations, the network again recalled all tanks correctly but classified 14 out of the 20 nontanks as tanks (false alarm rate of 70%), and only 6/20 or 30% got classified correctly as nontanks. When the iterations were increased to 10000, all tanks got classified correctly again, but only 9/20 nontanks got classified as tanks, a false alarm rate of 45%. After 15000 iterations, all tanks were again classified correctly with only one of the 20 nontanks misclassified as a tank, a false alarm rate of 5%. Since only one tank subimage was used and from that, the 20 noisy tank images were generated, it is possible that the tank data was very similar, and perhaps this was why all tank data was correctly classified in the neural net during these training sessions. These results are tabulated in the following table. These results show it is possible that the backpropagation network can distinguish nontank from tank data. Remembering that the backpropagation network is a simple approximation to the more complicated fuzzy morphology neural network, these promising results warrant a more detailed investigation into this methodology.

Training Data		No. of Iterations	Recall Data		% Classified as Tanks (supposed to be tanks)	% Classified as Nontanks (supposed to be nontanks)	False Alarm Rate
no. Tanks	no. of Nontanks		no. Tanks	no. of Nontanks			
20	20	10000	20	20	100%	100%	0%
5	5	10000	20	20	100%	30%	70%
7	7	5000	20	20	100%	30%	70%
7	7	10000	20	20	100%	55%	45%
7	7	15000	20	20	100%	95%	5%

Table 1. Results of 5 Training sessions.

6 Conclusions

The research presented here was intended as a proof-of-concept for using a new neural network for target classification. We believe that this goal has been met, showing indeed that a simple version of the fuzzy morphological neural net, essentially a backpropagation network, can classify tank data accurately with a low false alarm rate. A high number of iterations was needed to perform good classification, and this is one issue that will be investigated in future research. A more varied selection of input data needs to be considered, with different objects such as different types of tanks, vehicles, and other targets and nontargets. A more detailed investigation of system parameters and more efficient training rules is one of the first areas to be investigated. The full power of this network should come, of course, when a network model is established which allows training of all parameters w and t as well as p. This is the goal of the minigrant proposal which is submitted as follow-on research. Future research includes a learning algorithm for the convolution parameters, tailoring the algorithm to classify partially obscured data, and using the intensity data in conjunction with the range data as an additional source of information available. Once developed for this application, it is possible that this approach could be applied with

success to other classification problems including automatic target recognition, as well as other types of data such as MMW.

7. Acknowledgments

Several personnel were key to the success of this project. Capt. Norm Champigny and Dr. Tom Davis were instrumental in getting LTV range data down loaded to a machine which I used for my data processing, as well as for providing a clear explanation of the basic principles underlying range data. Ginny Wolfe and other secretaries provided competent and necessary infrastructure support in a pleasant way including such things as mail and fax deliveries. Emily Martinee provided the Schwartz height data in a format easily accessible by my machine, which was a real time savor for me. Captain Eric Augustus and Trevor Meyer spent much time in getting the computers up and running in the new room that I was assigned to occupy, and Pat Cofield was very generous in allowing me to use his Sun workstation. Dr. Dennis Goldstein was very helpful in administering the logistic details of the research program and did it all before our arrival on base. And last but not least, Ms. Karen Norris provided the opportunity, facilities and research support for me to accomplish my goals this summer. In my opinion, the summer program was a great success for me as well as for many others.

References

[1] P.D. Gader. Template generation for pattern classification. In *SPIE - Proc. Soc. of Photo-Optical Instr. Eng.*, volume 1769, pages 72–81, San Diego, CA, Jul. 1992.

[2] R. Hecht-Nielsen. *Neurocomputing*. Addison-Wesley, New York, 1989.

[3] IVS Inc. *Image Algebra FORTRAN Version 3.0 Language Description and Implementation notes*. IVS Inc., Gainesville, FL., 1989.

[4] R. P. Lippmann. An introduction to computing with neural nets. *IEEE Magazine on Acoust., Speech, Signal Proc.*, ASSP-4:4–22, 1987.

[5] NeuralWare, Inc. *NeuralWorks Professional II, A User's Guide*. Pittsburgh, PA, 1989.

[6] G.X. Ritter, D. Li, and J.N. Wilson. Image algebra and its relationship to neural networks. In *Proc. of the 1989 SPIE Tech. Symp. Southeast on Optics, Elec.-Opt., and Sensors*, Orlando, March 1989.

[7] G.X. Ritter, J.N. Wilson, and J.L. Davidson. Image algebra: An overview. *Comp. Vis., Graphics, and Image Proc.*, 49:297–331, Mar. 1990.

[8] P.K. Simpson. *Artificial Neural Systems*. Pergamon Press, New York, 1990.

VLSI SYNTHESIS GUIDING TECHNIQUES USING THE SOAR ARTIFICIAL INTELLIGENCE ARCHITECTURE

Joanne E. DeGroat
Assistant Professor
Department of Electrical Engineering

Lindy Fung
Graduate Student (Master's Degree)
Department of Electrical Engineer

The Ohio State University
205 Dreese Laboratory
2015 Neil Avenue
Columbus, OH 43210

Final Report for:
Summer Research Program
Wright Laboratories

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

VLSI SYNTHESIS GUIDING TECHNIQUES USING THE SOAR ARTIFICIAL INTELLIGENCE ARCHITECTURE

Joanne E. DeGroat
Assistant Professor
Department of Electrical Engineering
The Ohio State University

Lindy Fung
Graduate Student (Master's Degree)
Department of Electrical Engineer
The Ohio State University

The design of VLSI circuits is a very complex problem. As such, there are numerous computer aided design (CAD) tools to assist the designer in the generation of VLSI circuits. These tools range from layout editors, physical design tools, to high level synthesis tools. In all of these tools the design style and architectural structure are very constrained to limit the complexity to a manageable level. These constraints result in circuits that require more area and are slower than circuits produced manually by a skilled VLSI designer. The application of SOAR to the problem of VLSI creates the possibility of relaxing some of these constraints. The design of VLSI circuits is a design process. Existing tools and methodologies are incapable of capturing any of the essence of this design process and therefore must severely limit the design methodology. This research is a first step in the application of the SOAR Artificial Intelligence Knowledge Based Architecture developed at Carnegie-Mellon University to the synthesis of VLSI circuits. This first step is the application of SOAR to the VLSI placement problem. Findings, limitations, and recommendations for future research are presented.

VLSI SYNTHESIS GUIDING TECHNIQUES USING THE SOAR ARTIFICIAL INTELLIGENCE ARCHITECTURE

Joanne E. DeGroat

Lindy Fung

INTRODUCTION

The design of VLSI circuits is a very complex problem. As such, there are numerous computer aided design (CAD) tools to assist the designer in the generation of very large scale integrated circuits (VLSI) [1]. These tools range from layout editors, physical design tools, to high level synthesis tools. High level synthesis tools generate VLSI circuit layouts from hardware description language code. Physical design is the automatic generation of macro cell circuit layouts that are then used in high level synthesis. In the past few years significant advances have been made in both these areas, and to some degree, both high level synthesis and physical design are considered commonplace. However, the design style and architectural structure used in high level synthesis is very constrained to limit the complexity of automating the VLSI design process to a manageable level. Physical design is likewise constrained to a few implementation styles. These constraints result in circuits that require more area and are typically slower than circuits produced manually by skilled VLSI designers.

The application of SOAR[2] to the problem of VLSI synthesis creates the possibility of relaxing some of these constraints. The design of VLSI circuits is a "design process." Existing tools and methodologies implement only a small aspect of this design process and are incapable of capturing the true essence of this design process, therefore severely limiting the design methodology. The research highlighted in this report takes the first step in the application of SOAR to VLSI synthesis.

BACKGROUND

The design of a VLSI circuit begins when the specification of the circuit is written. This specification describes the functionality of the circuit and can be done in a modern hardware description language. Once the specification is completed, the design of the circuit begins. This design is an iterative process of breaking the circuit into major components, then subcomponents and finally down to logic primitives or gates. The gates are arranged or *placed* and the signals between the gates connected, i.e., *routed.* The subcomponents that make up a component are then placed and routed. This process is repeated recursively until the circuit is completed.

Consider the following example. The designer desires to create the Roulette Game Chip shown in figure 1. At this level, the inputs, outputs, and functional and temporal behavior can be

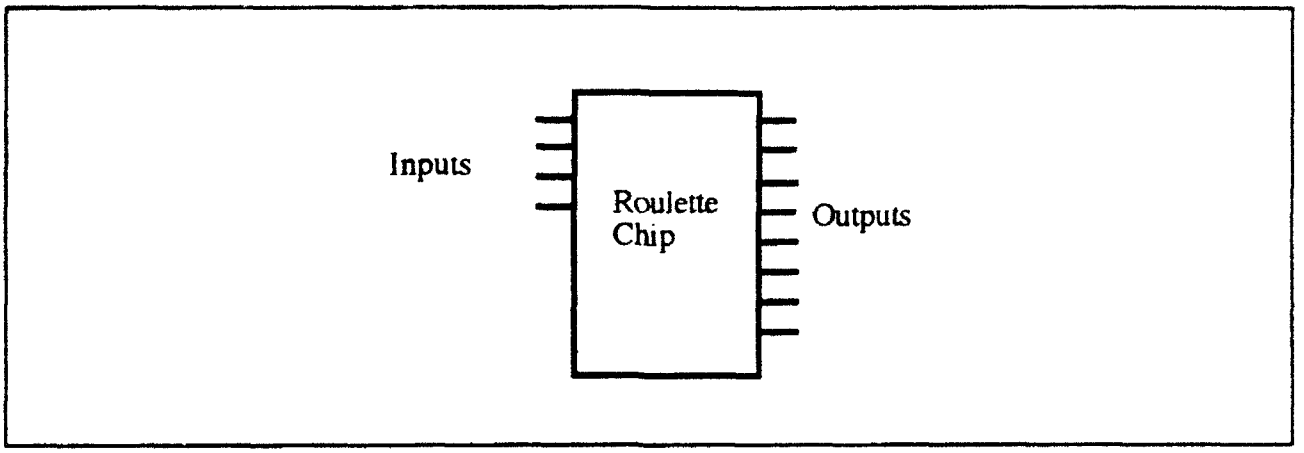


Figure 1. Roulette Chip

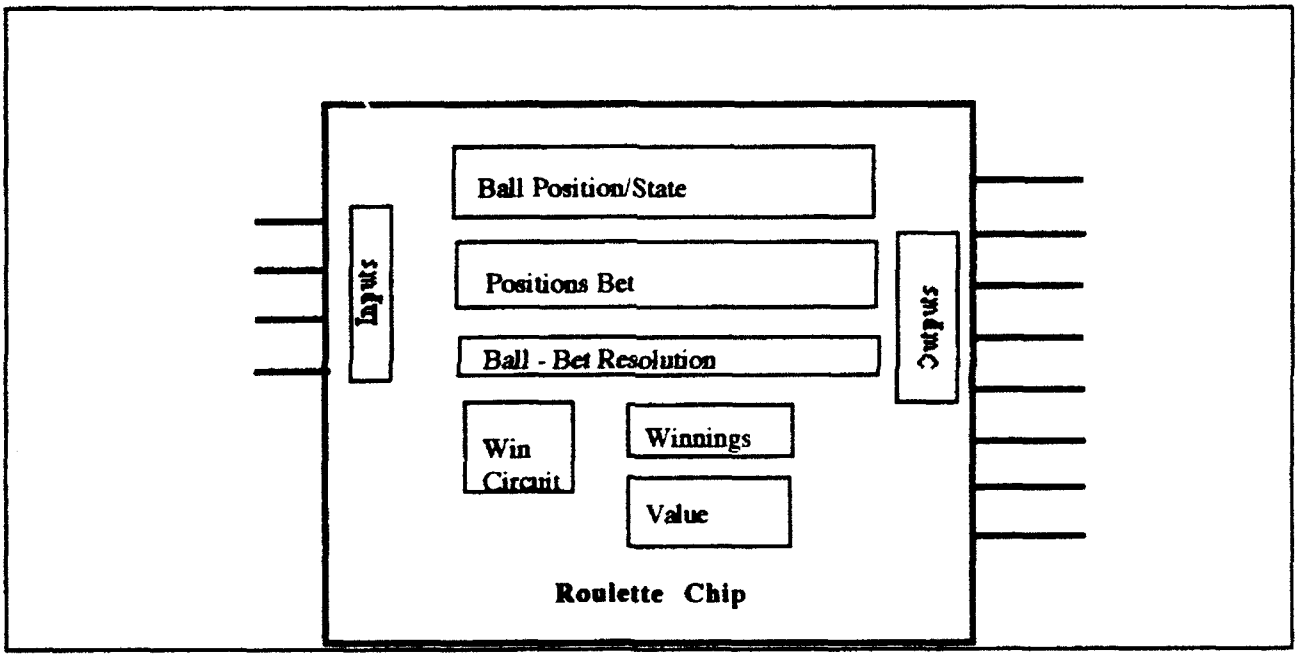


Figure 2. Major Components of the Roulette Chip

specified. Once this is complete the design can continue with partitioning into major components that make up the chip as shown in figure 2. Each of these components is then designed individually. When partitioning the design into major components, the components are initially placed reflecting their relationship and interconnections to the other components.

The major components shown in figure 2 may or may not be directly implementable. The level of complexity of these components can still be quite significant, requiring further subdivision of the component into more manageable subcomponents. One such component is the value component that consists of a four bit adder and four bit register as shown in figure 3. The signals that cross the dashed line around the value component are the external interface of

the component. These are the signals that will be routed and connected to the other components of the chip or to the inputs or outputs. The four bit adder and the four bit register are the two internal subcomponents of the value component and would each be generated individually. The adder would be designed and implemented using an adder architecture that provides the functional and temporal behavior required to meet the specifications. One possible implementation is a four bit ripple carry adder constructed by placing four full adders together. For that adder architecture, the four bit adder can be divided into four subcomponents, each of which is designed independently, or as in the case here, designed once and replicated.

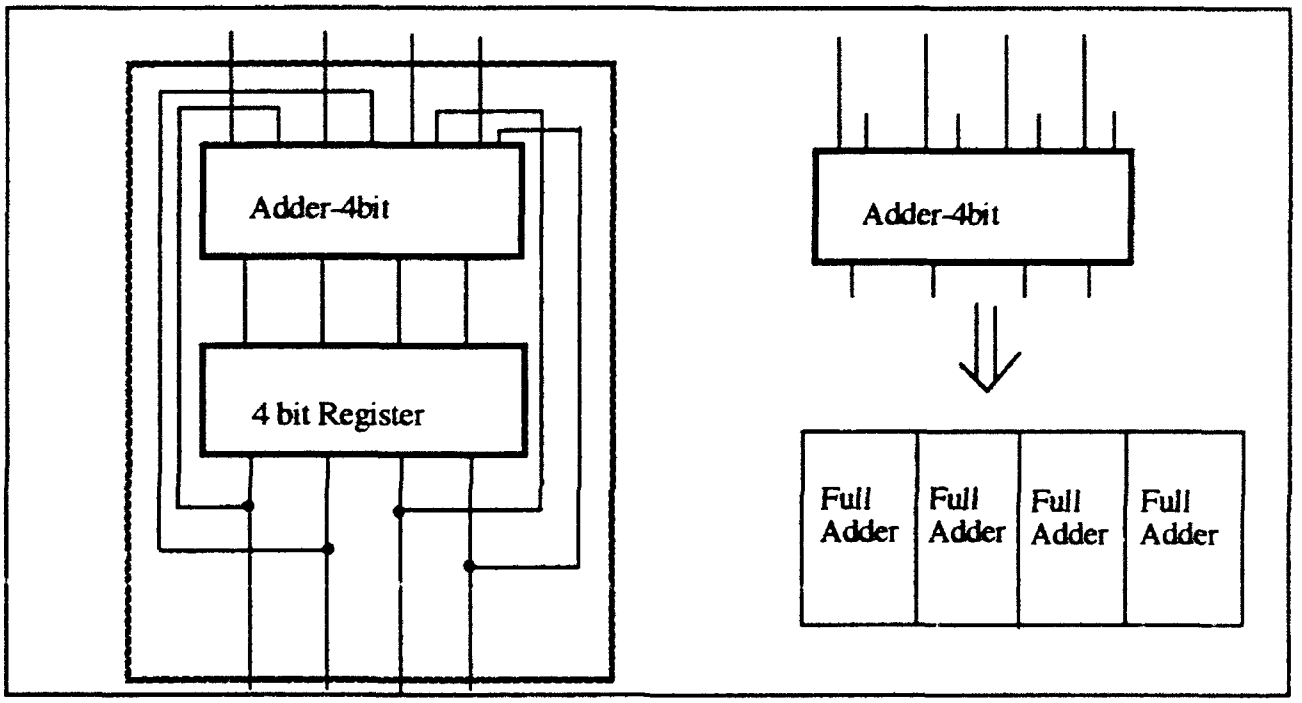


Figure 3. Value Component and Subcomponent Hierarchy of Value Component

The full adder sub component implements the full adder function. The logic equations to be implemented are given in figure 4 along with an implementation of the logic equations. The equations are implemented by six logic gates that must be placed and routed to construct the full adder component. Note that in the example shown, it is desirable to have the signals A and B at the top, Cin at the right, Cout on the left, and the Sum on the bottom boundary. These constrains provide boundary conditions for the placement of the gates within the full adder.

In each component, the elements of which it is comprised must be placed and routed within the boundaries specified. This is a constrained placement problem and, as illustrated by the example above, the placement of components is required at all levels in the design process. Thus the first application of SOAR to the task of synthesizing VLSI circuits will be the application of SOAR to the placement problem.

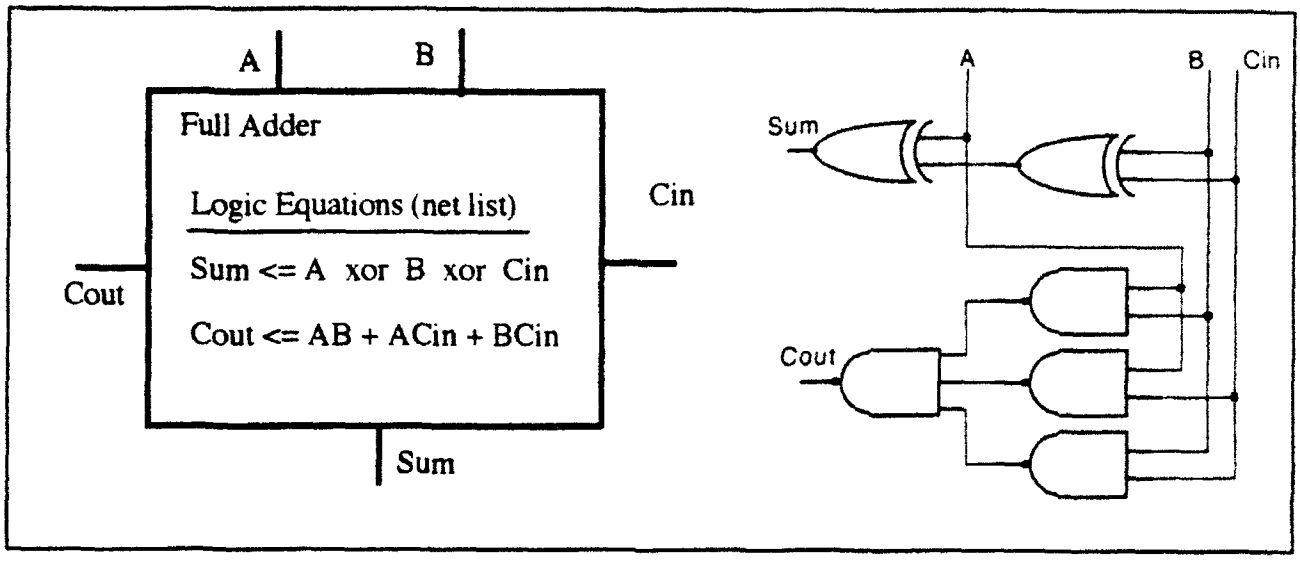


Figure 4. Full Adder Logic Equations

The general placement problem can be stated as "Locate a set of components in a given area such that they do not overlap." The VLSI placement problem has additional considerations beyond the basic constraint that the components must not overlap. These constrains arise from the relationship of components to each other through signal interconnections. A common placement criteria figure of merit is the total wire length required for the placement. This figure of merit is based on the fact that the length of interconnect between components is inversely proportional to the speed of the circuit – long signal interconnections results in slower circuits. Thus it is desirable to place components next components with which they have signals in common. Additionally, the inputs and outputs to the module or component under design are at fixed locations on the boundary of the module.

There are three classes of algorithms commonly utilized for the solution of the VLSI placement problem. These are constructive initial placement, iterative placement improvement, and branch and bound. Within these classes are various placement algorithms such as the cluster development method, the Steinberg assignment method, relaxation, pairwise interchange, and simulated annealing. In almost all algorithms the first step is to arrive at an initial placement. This pre-placement of components is very important as "during pre-placement a good configuration is found to aid the final placement algorithm in searching for a suitable final configuration. The quality of solutions of the pre-placement algorithm is of great importance for obtaining a good final result."[3]

Development of new algorithms in other fields has in some sense re-opened the VLSI placement problem. Recent papers have seen the use of neural nets for placement[3], the use of fuzzy logic[4], the use of the genetic algorithm[5], and improvement on existing algorithms

[1][6][7][8][9]. The research reported here follows this research in looking for new and novel approaches to the VLSI placement problem. As the placement problem spans all levels of automating the VLSI design process, it is an excellent choice for evaluating the use of SOAR in VLSI synthesis. Within the research, algorithms for initial placement of VLSI components will be mapped to SOAR and evaluated.

USE OF SOAR FOR VLSI PLACEMENT

Within the research two algorithms for initial placement were developed. These will be discussed in detail in this section. A third algorithm that is a hybrid of the these two algorithms will also be discussed and is a possible direction for future research. Prior to presenting the algorithms the general architecture of the systems must be discussed. The algorithms being developed are general in nature, i.e., they are designed to handle any valid VLSI placement problem. This means that there are

1) an arbitrary number of components to place,
2) an arbitrary number of signal connections between components and to the boundary,
3) an arbitrary area for placement,
4) and an arbitrary number and arbitrary placement of edge signal connections.

By arbitrary it is meant that in one problem instance there may be 4 components, 22 signal connections, and an area of 100λ[†] by 85λ. In the next problem instance there may be 20 components, 235 signal connections, the area is 1500λ by 350λ. The fact that the algorithm must be general in nature increases the complexity of mapping an algorithm to SOAR. Additionally, the I/O capabilities of SOAR must be used to input the initial design constrains. These inputs are obtained from files that contain the relevant information for each design.

The first of these files is the *DEFAULT_SIZES* file shown in figure 5. This file is used to specify the size and number of inputs and control signals for the primitive components (or gates) to be used in the design of the module. It is assumed that all primitive components will

DEFAULT_SIZES FILE FORMAT	Example			
#primitives	7			
TYPE Height Width #Inputs	INV_1_1	37	13	1
TYPE Height Width #Inputs	NAND_2_1	36	26	2
•	NAND_3_1	36	36	3
•	NOR_2_1	36	20	2
•	XOR_2_1	45	40	4
TYPE Height Width #Inputs	XNOR_2_1	45	40	4
	T-GATE_2_1	28	18	3

Figure 5. *DEFAULT_SIZES* File Format and Example

[†] Lambda, λ, is a unit of measure in scalable CMOS design.

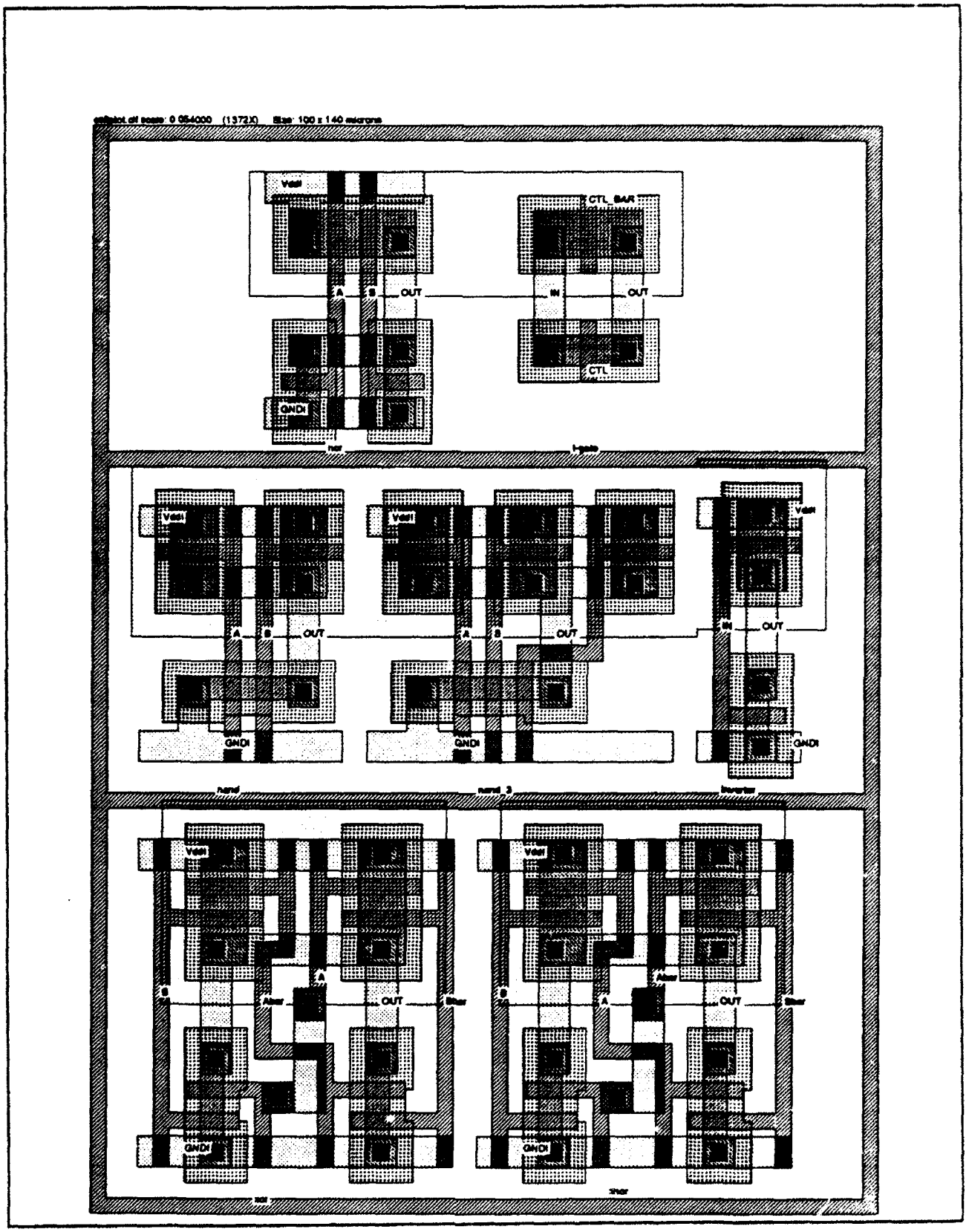


Figure 6. VLSI Primitives

have one output. Figure 5 shows both the file format and an example file. The height and width value are in λs and based upon the minimum size scalable CMOS using the MOSIS rev 6 design rules. The Type field is the type of the component. INV_1_1 is an inverter that requires an area of 37λ by 13λ on an integrated circuit and has a single input and a single output. The NAND_2_1 primitive represents a 2 input NAND gate. The VLSI layouts of the primitives are shown in figure 6.

The second file is the problem specification file, i.e., the file that contains the information as to the VLSI placement problem to be solved. This file contains the information shown in figure 7. The first item of information is the number of components in the problem. This is followed by the components, with the unique name, type, output name, and input names given for each. Following the specification of the components is the width and height of the area in which the components are to be placed. The number of edge signal connections is then specified and the information given for each edge signal. Figure 8 gives an example of a specification file and the boundary that is specified is shown in figure 9. Note that two of the signals are on the top boundary, one on the right boundary, and one on the bottom. It is obvious that the components using these signals should be placed as close as possible while allowing for a solution.

```
#components_to_be_placed
NAME        TYPE Output        Input   [Input]
•••
NAME        TYPE Output        Input   [Input]
Area_width  Area_Height
#edge_signal_connections
Name Mode xloc   yloc   width edge   material
•••
Name Mode xloc   yloc   width edge   material
```

Figure 7. Input File Specification

```
6
Ga1     INV_1_1      ABAR  A
Ga2     INV_1_1      BBAR  B
Ga3     INV_1_1      CBAR  C
Ga4     XOR_2_1      AXORB  A  ABAR  B  BBAR
G5      INV_1_1      AXORBBAR  AXORB
G61     XOR_2_1      SUM  AXORB  AXORBBAR  C  CBAR
100     160
4
A   in   30   160   2   n   p
B   in   60   160   2   n   p
C   in   100  100   2   3   p
SUM out  50   0   4   s   m1
```

Figure 8. Example Input Specification File for a Part of a Full Adder

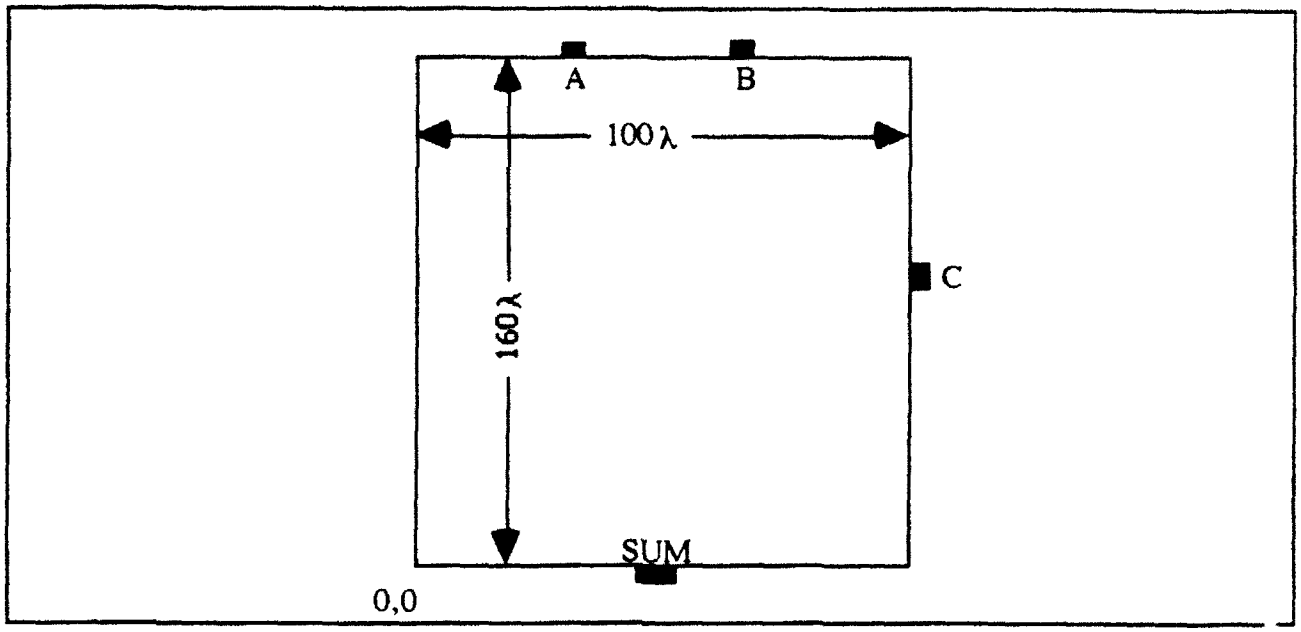


Figure 9. Boundary for Example Input File (reference figure 8)

The input file specification is read into SOAR using the file I/O capabilities of the SOAR system. This gives rise to the state structure shown in figure 10 that is the initial state. The gate augmentation is a multi-attribute augmentation to the state and an occurrence of this

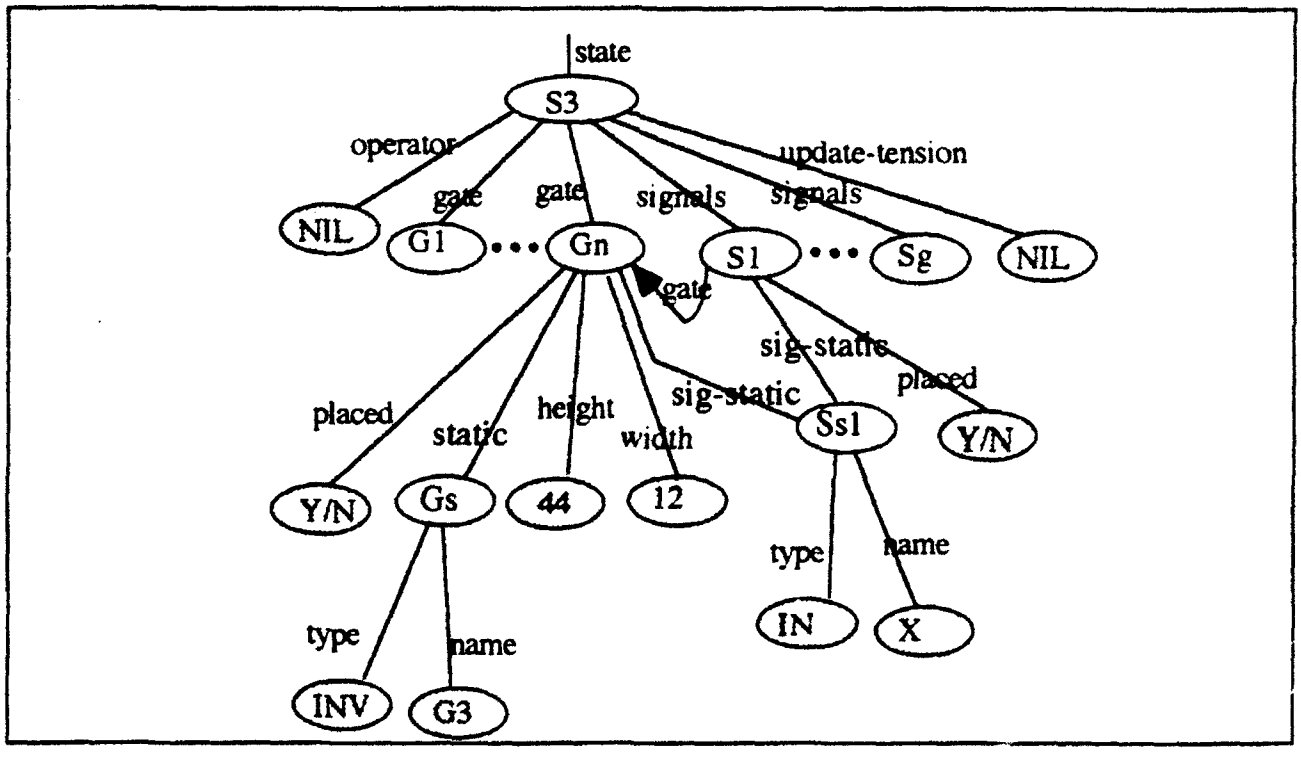


Figure 10. Initial State Structure

substructure exists for each gate in the problem. Similarly, the signals augmentation is a multi-attribute augmentation to the state. The operator and update-tension augmentation are used in the calculation of the wire length metric, called the tension of the system. The remainder of the information shown and present in the system is augmented to the gate or signal augmentations.

The other internal SOAR data structure derived from the input data file is the desired augmentation. This will be augmented to the goal. Figure 11 shown the structure of this augmentation. This structure is built from the boundary specification and is the internal representation of the boundary conditions. This information is made part of the desired augmentation as this information is static. The boundary conditions do not change as the solution is derived.

Using this internal information the initial placement algorithms work to place the components. The first algorithm is a building block algorithm and the second algorithm is a constructive initial placement algorithm.

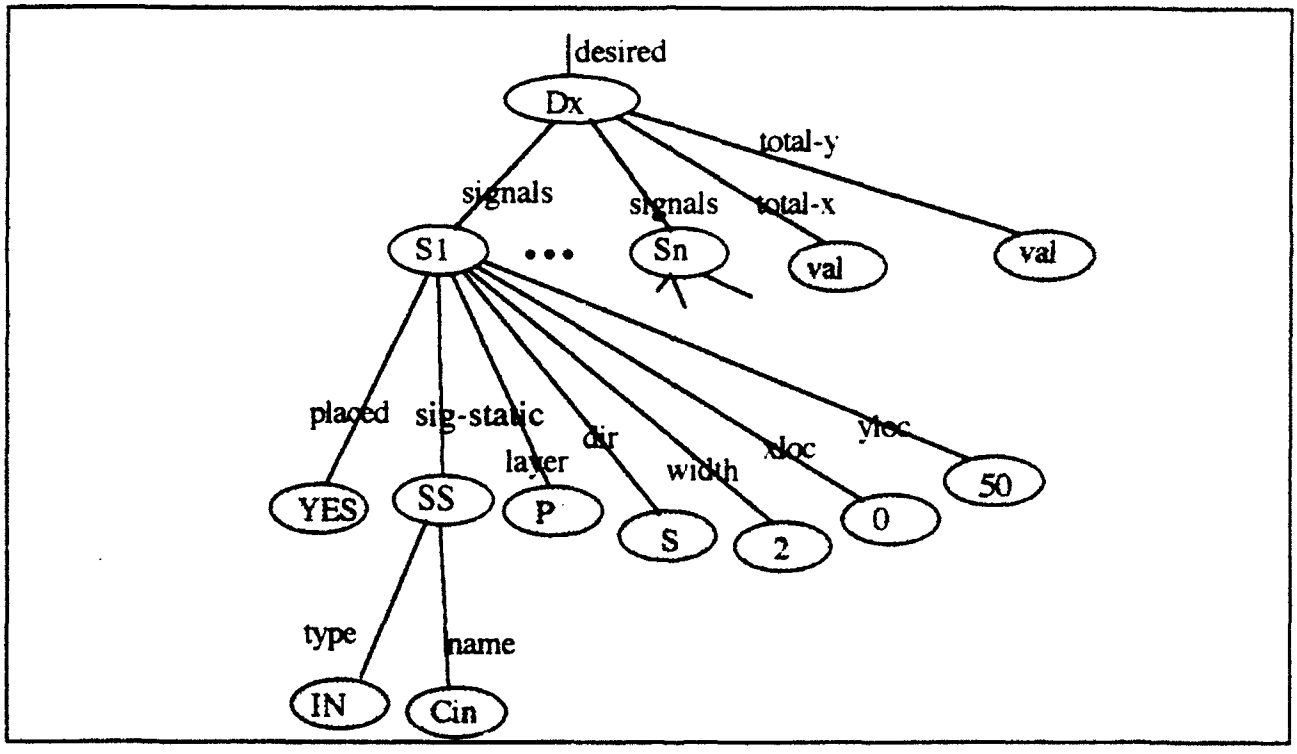


Figure 11. The Desired Augmentation

Algorithm 1

Algorithm 1 is a building block algorithm with few operator preferences. The algorithm is:

1) Propose all components for location 0,0
2) Place the component that has a signal common with any signals on the edge along

the block. If no component has a signal common with the edge use lookahead search to see if placement results in success.

3) Propose locations of upper left and lower right for all remaining components in relation to the component just placed. Go back to step 2 until all components are place and all signals are placed.

Figure 12 gives an example of the operation of the algorithm. Both components are proposed for location 0,0. The component with signals A,B, and X is chosen and found to result in success. It is placed in location 0,0 and then new locations at its lower right and upper left are proposed for all remaining components. The component with signals A and B is chosen and placed as shown. The new locations at its lower right and upper left are proposed for any remaining components. Note that the first placed component has signal A and matches signal A on the boundary along the edge of the component. Also note that the second placed component has signals A and B in common with the first placed block.

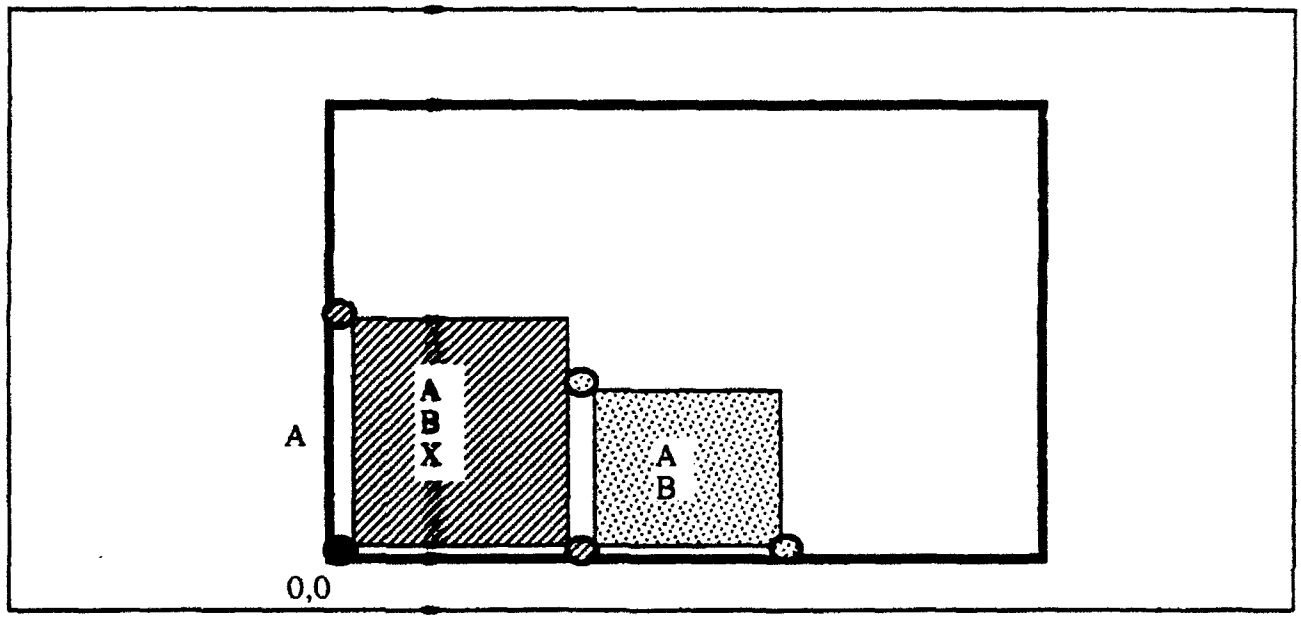


Figure 12. Example of Algorithm 1.

The problem space computational model (PSCM) of Algorithm 1 is shown in figure 13. It consists of a top problem space and a second problem space in which the actual placement of components is done. Within the PSCM there are two operators. These are *place-block-up* and *place-block-right.* Examining the algorithm it can be seen that as the problem is solved the number of operators present rises and then falls. The maximum number of operators present in a state at any given time is $\lfloor \left(\frac{n+1}{2}\right)^2 \rfloor$ where n is the number of components to be placed. This number of operators is reduced by those operators that are prohibited.

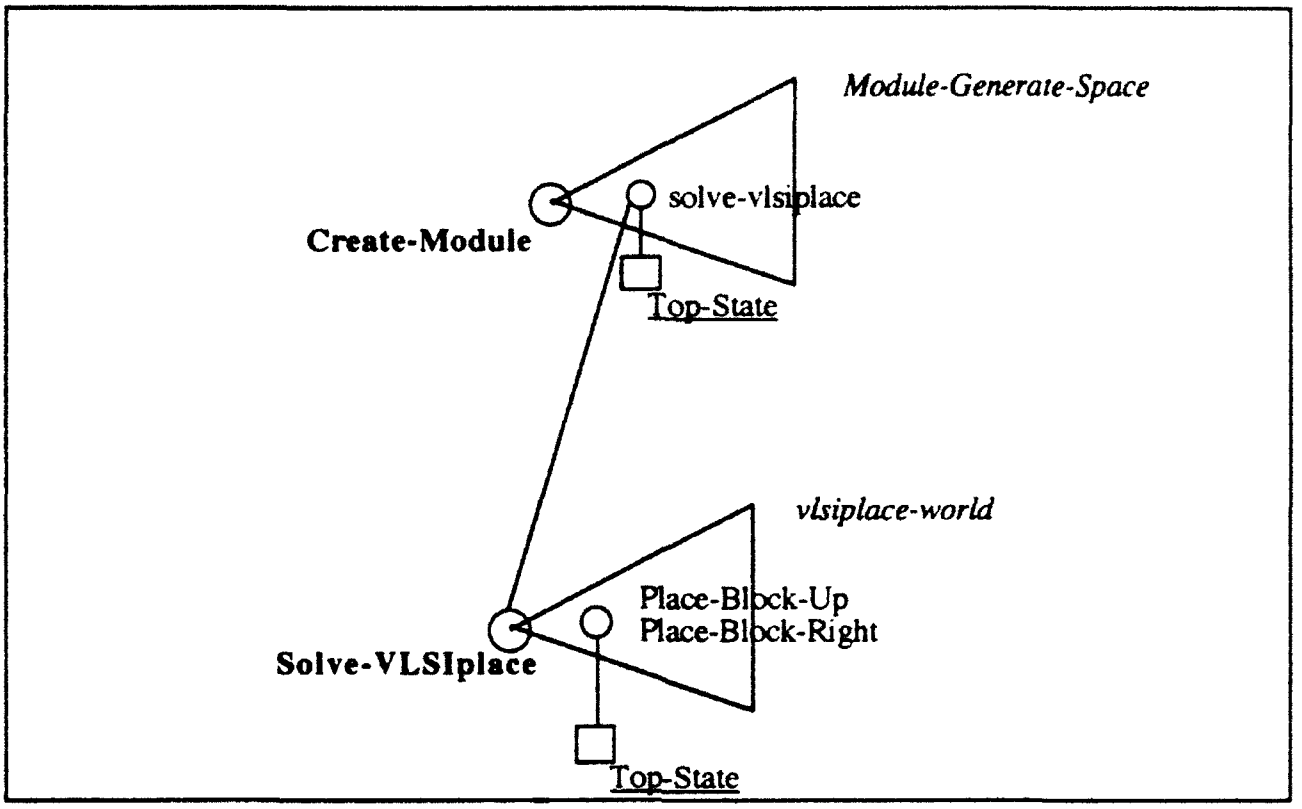


Figure 13. The PSCM of Algorithm 1

The prohibited operators in algorithm 1 are illustrated in figure 14. In the first case the placement of the component would result in the component extending outside the boundary. As this is not allowed in the placement problem this operator must be prohibited and thus removed from any further consideration. The second case shows the case of the proposed location resulting the component being placed such that it overlaps an already placed component. This is also not allowed and this operator must also be prohibited.

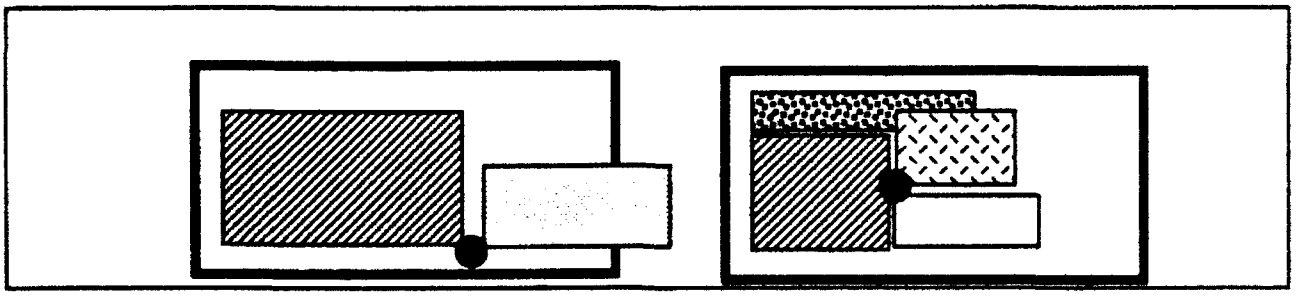


Figure 14. Prohibited Operators

The prohibited operators reduce the number of operators that must be considered in finding a solution. Algorithm 1 also has a preference rule that assists in choosing the best operator. Figure 15 illustrates the preference rule of algorithm 1. If the component is being

placed adjacent to the boundary or another component such that a signal on the component matches a signal on the boundary, then that operator is given a best preference. If no signal is adjacent, then the operator has an indifferent preference, i.e., it is no better or worse to choose this operator over any other indifferent operator.

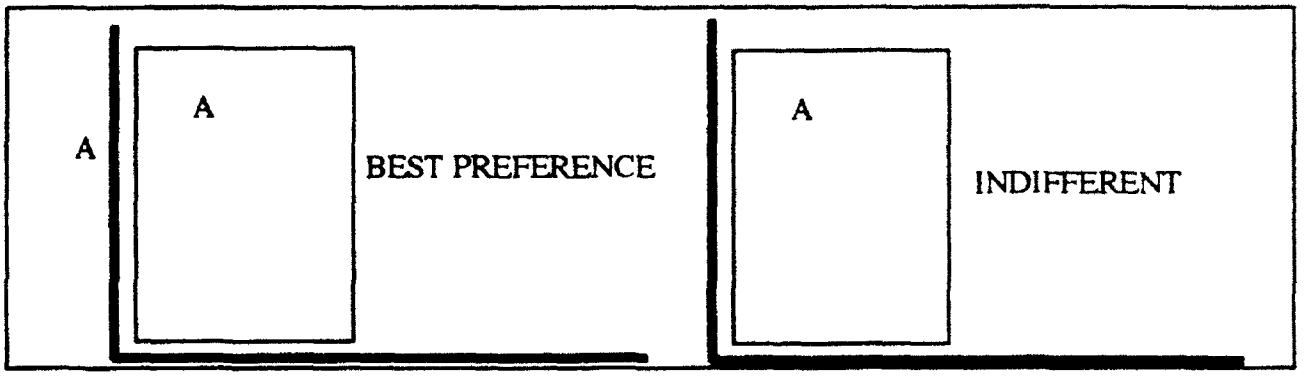


Figure 15. Algorithm 1 Preferences.

The final discussion of algorithm 1 is its test for success. Algorithm 1 continues to work to place the components until all blocks are placed and all signals are placed. A signal on a component is placed when the component is placed adjacent to another component that also has the signal, or when the component is placed adjacent to the boundary and the signal is on the boundary along the edge of the component. If all signals are not placed then the result is failure.

Algorithm 2

Algorithm 2 uses the distance between signal connections for the preferences of operators and proposes operators based upon signal connections. The algorithm is as follows.

1) For each signal on an unplaced component, propose an operator for placement of the component in relation to all other components or all locations on the boundary that have the signal in common
2) Calculate the tension of all proposed operators
3) Choose the operator with the lowest tension and place the component
4) Repeat steps 1, 2, and 3 until all blocks are placed.

In this algorithm five locations are proposed for each signal a component has in common with the boundary. This is illustrated in figure 16. These are the locations that would most likely be chosen by an experienced VLSI designer as an initial attempt in placement of the component. Similarly, four locations are proposed for placement of a component next to a placed component with which it has signals in common. This is shown in figure 17.

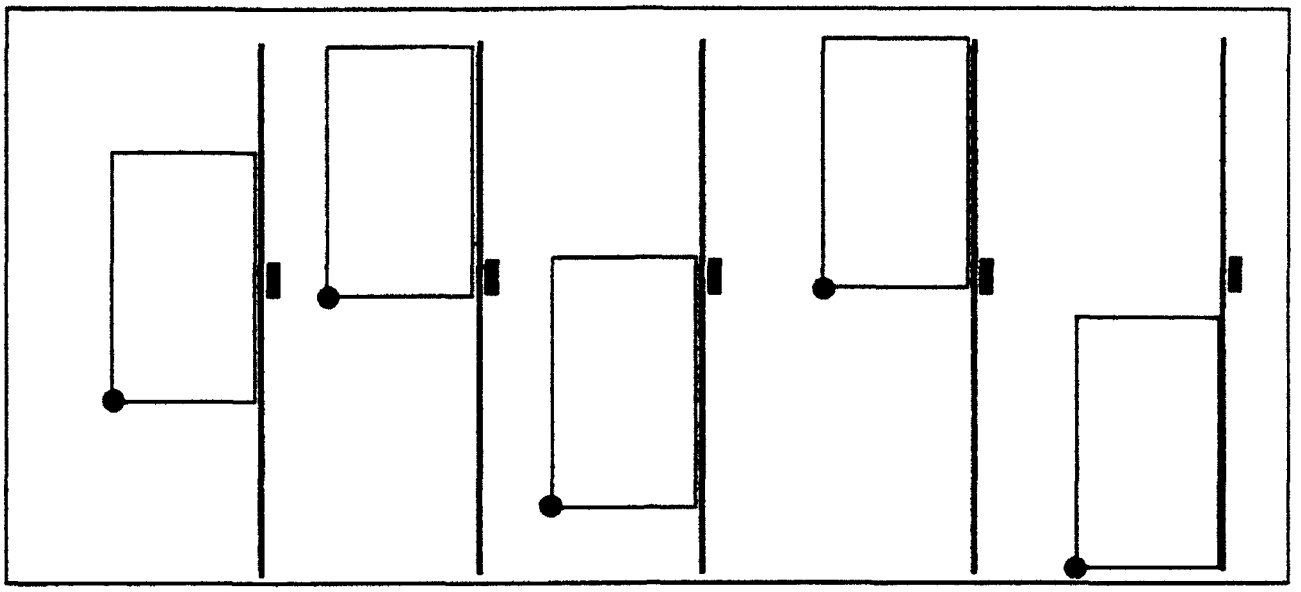


Figure 16. Operators Proposed for Boundary Signals in Algorithm 2

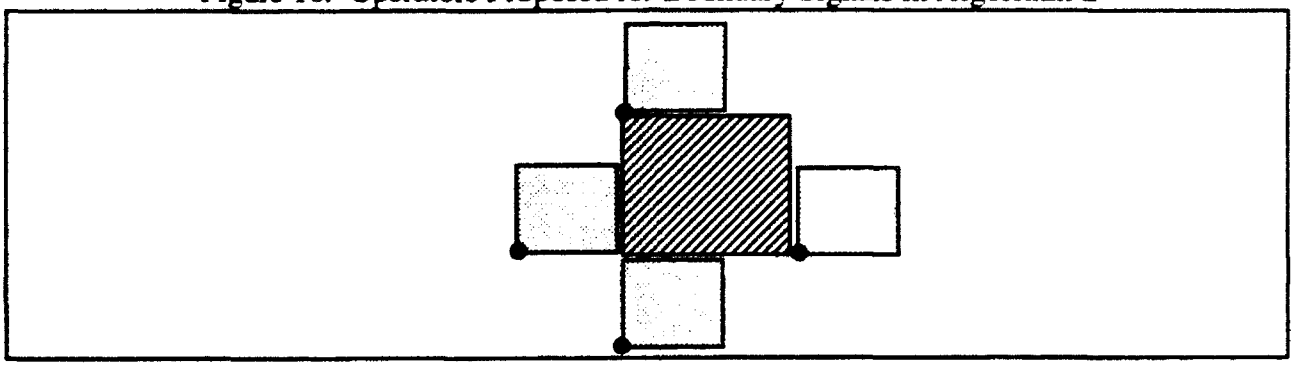


Figure 17. Operators for Placement Next to a Placed Component in Algorithm 2

The PSCM for algorithm 2 is shown in figure 18 and is very comparable to the PSCM of algorithm 1. There are however, significant differences. The first is the operators proposed in the *place-components-space*. Although there are 5 SOAR productions that propose operators under 5 unique conditions, there is a single operator type within the system. Thus, the operator for placement of a component based on a signal in common with the boundary is the same as the operator for placement next to a placed component with which it has a signal in common. The only difference is the placement location (as to be expected). The maximum number of operators present in algorithm 2 is based on the number of signals connections rather than the number of components. Also, as five operators are proposed for each signal in common with the signals on the boundary the maximum number of signals is 5*signal_interconnections. All components have at least one input and one output and most components have 2 inputs. Thus the number of signal interconnection is greater than approximately three times the number of components. Signals that are connected to many components increase the number operators dramatically.

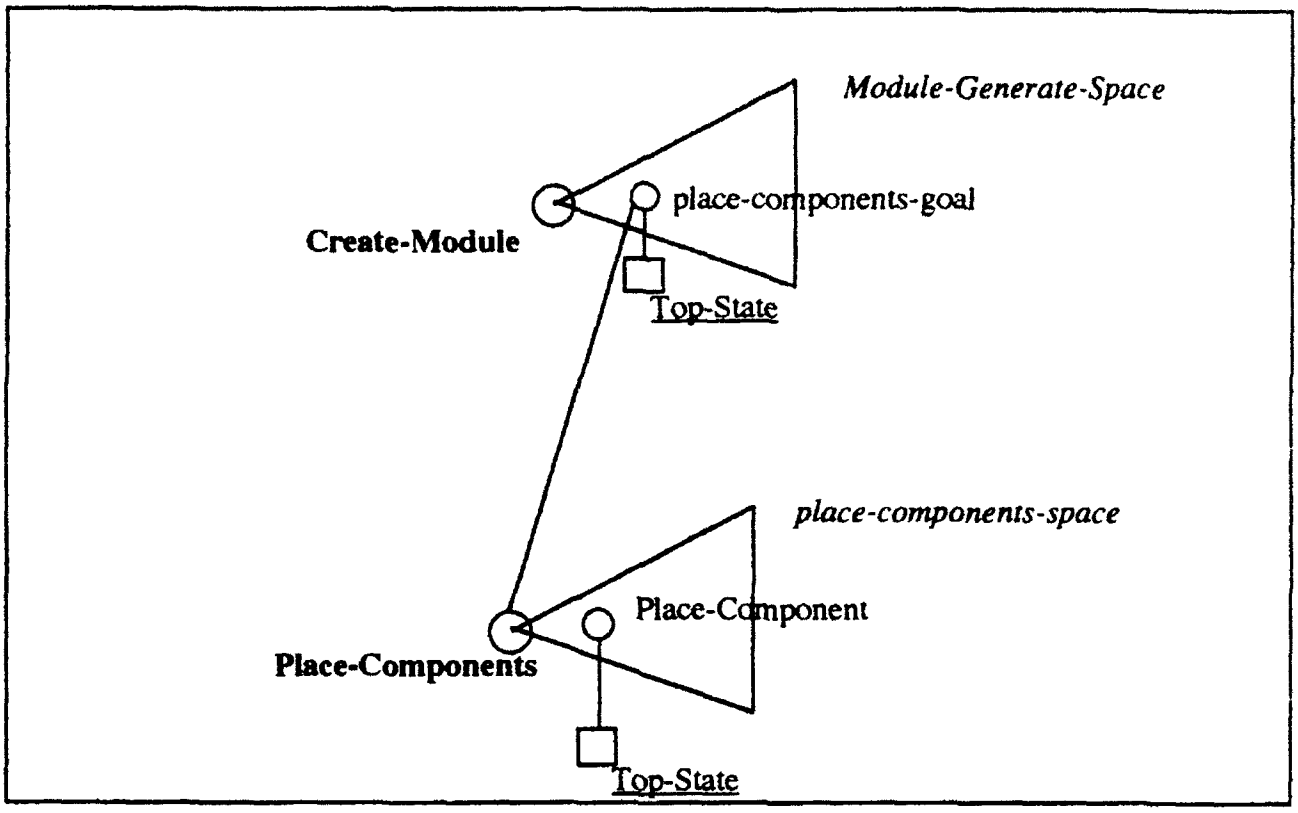


Figure 18. Algorithm 2 PSCM

As with algorithm 1 the number of operators is reduced as some of the operators are prohibited. The prohibited operators in algorithm 2 are the same as the prohibited operators in algorithm 1. In contrast, however, algorithm 2 has a more advanced preference mechanism. Figure 19 shows two components to be placed within a boundary. The operator to be preferred is the operator with the lowest tension. The tension calculation is also shown. It consists of

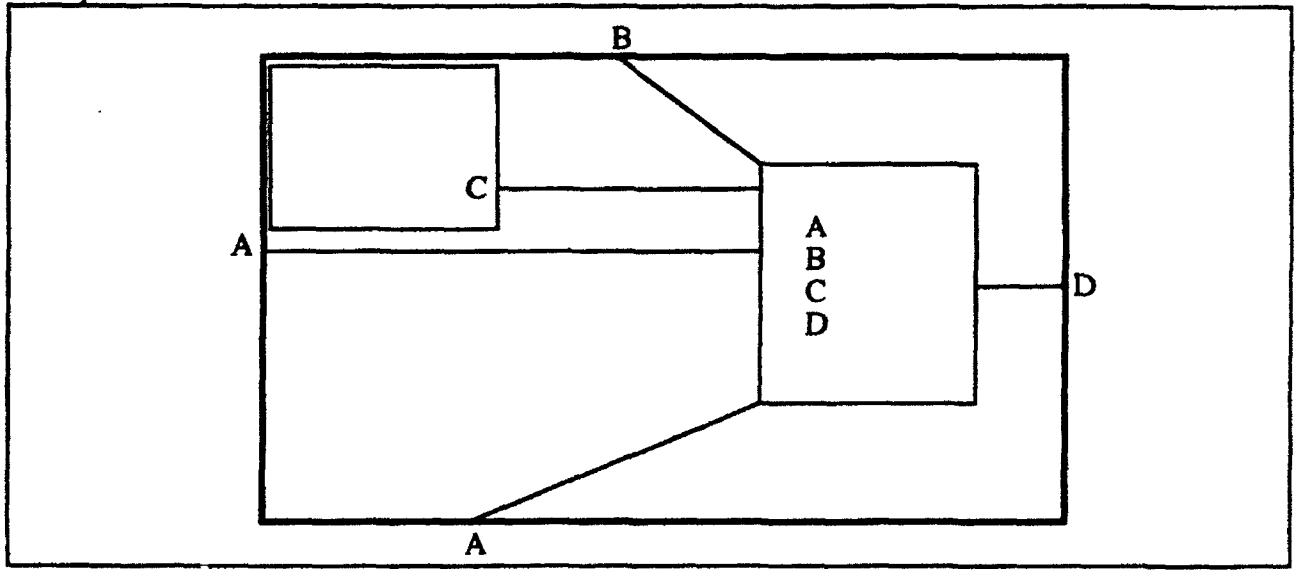


Figure 19. Tension Calculation Example

summing the signal interconnection distance in the x direction and the y direction. The total tension metric is the magnitude of the vectored tension.

Total Tension in x direction $T_X = \sum_{s \in S} (prx_s - plx_s)$

Total Tension in y direction $T_y = \sum_{s \in S} (pry_s - ply_s)$

Total Tension $T = \sqrt{(T_x^2 + T_y^2)}$

where S is the set of all signals.

The test for success in algorithm 2 is similar to algorithm 1 except that algorithm 2 is considered to have succeeded when all blocks are placed. Failure will occur when all blocks are not placed and there are no non-prohibited operators. When failure occurs a new operator can be chosen for placement of the first component until a successful solution is reached.

RESULTS

The two algorithms were run on numerous problem specifications. Algorithm 1 was run successfully on several problems including a problem involving the exact fit of 5 components. The final placement and the signals present in each component are shown in figure 20. In this problem specification there are 5 components named A, B, C, E, and F. The algorithm correctly places the components within the boundary matching external signals A1 and F1. Examination of algorithm 1 shown that it works for many problems but is overly restrictive for general use. This is due to the fact that in the algorithm requires all signals to be placed. If any other

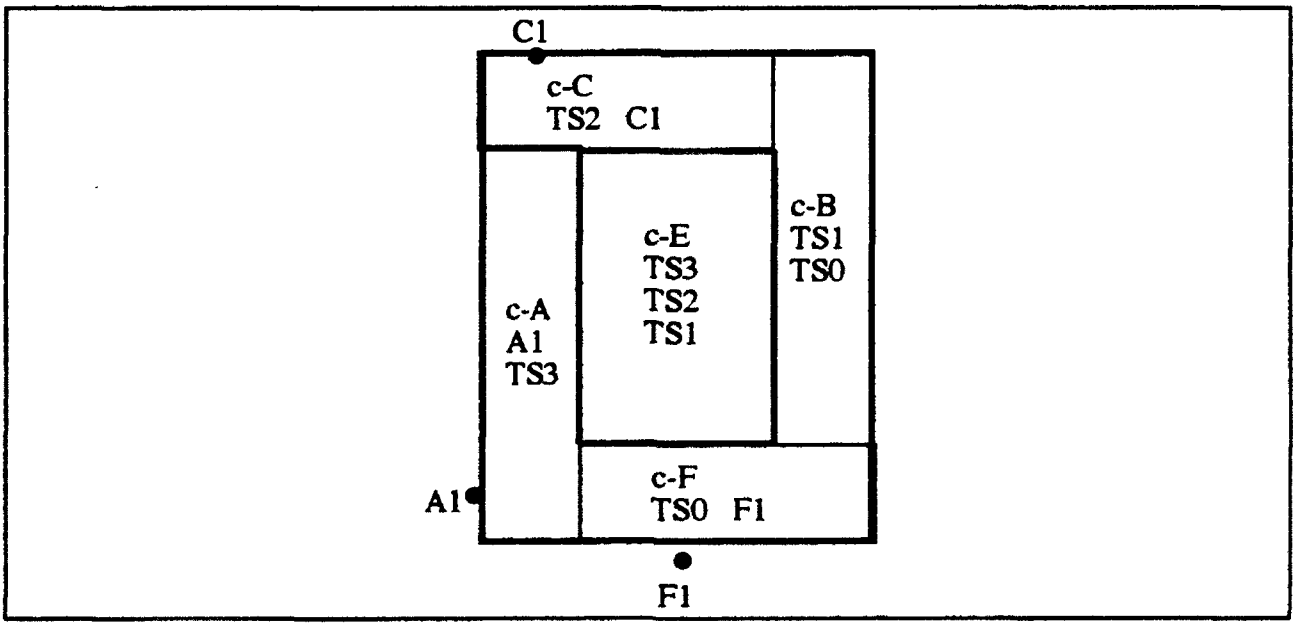


Figure 20. Results of Place of 5 Components by Algorithm 1

component contained signal A1 the algorithm would be unable to successfully place the components.

Algorithm 2 uses the tension of a component for determination of preferences, but only requires that the components be placed to achieve success. Thus many components may contain the edge signals and signals on other components. This is the general problem that must be solved. Figure 21 shows an example of a problem specification in which the components cannot be adjacent to all signals that they contain. Algorithm 2 places the components as shown in the figure. The statistics for the SOAR execution for this problem are also shown in the figure. During execution of algorithm 2, thirty operators are proposed initially, one for each of the signals on the components common with the edge signals. Recall that for each of these 5 possible locations are proposed. Once these operators are proposed, the operator with the lowest tension is chosen. In this case there are 3 operators with '0' tension and lookahead search must be used to determine if the operator chosen will result in success.

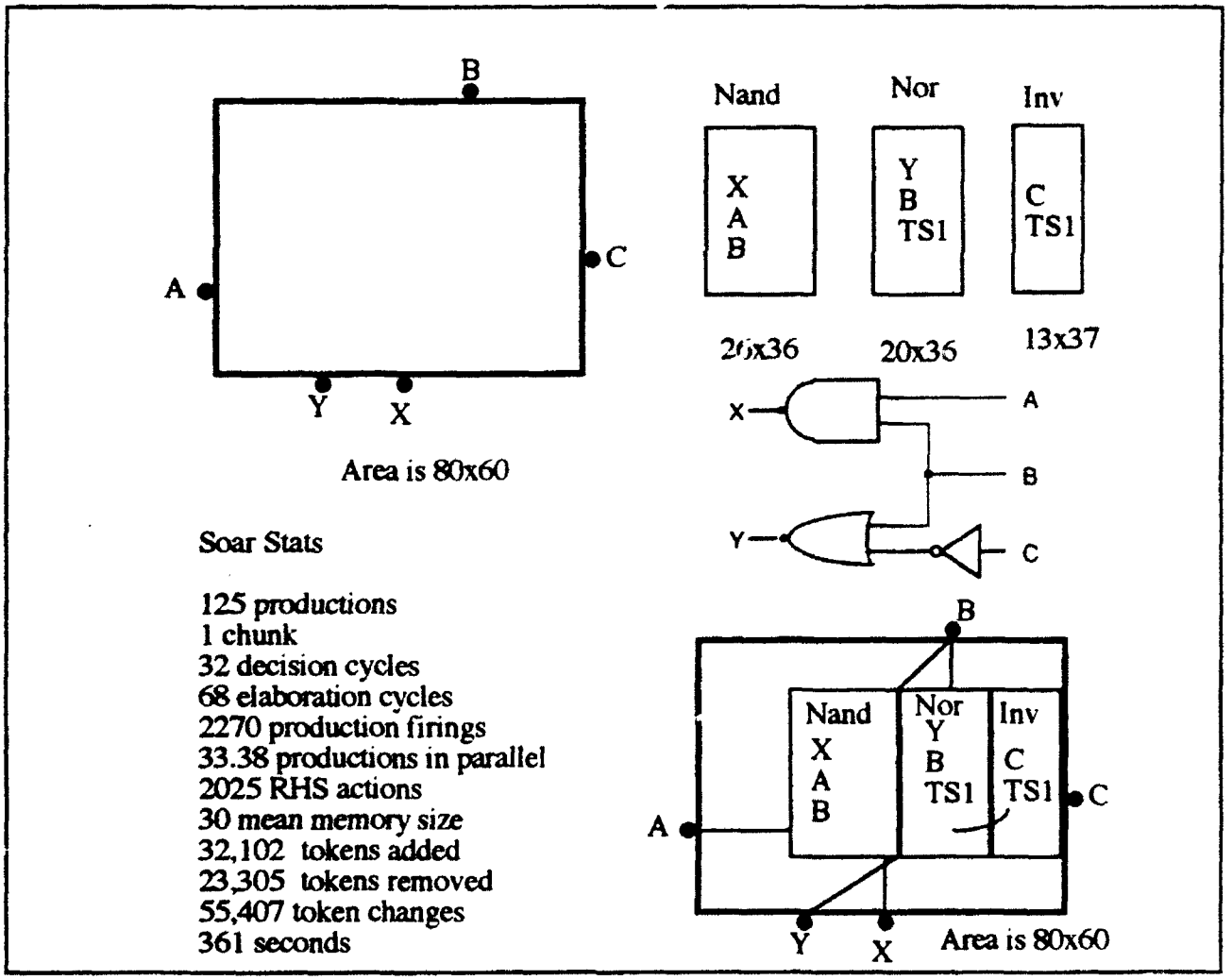


Figure 21. Results of Algorithm 2

CONCLUSIONS

There are several conclusions that can be derived from the results of the summer's research. The first is that SOAR does indeed have uses in the synthesis of VLSI circuits. It does, however, require reformulation of traditional algorithms to take advantage of the features of SOAR. During the research learning was not used in the execution of either algorithm 1 or algorithm 2. Learning was used on an example general SOAR program and found to be very beneficial. For learning to be beneficial required the SOAR operator action productions to be written in very specific forms to avoid the chunks being overgeneral. In fact the first two chunks formed were overgeneral. Once they were modified slightly and included in the original program, the additional chunks formed were no longer overgeneral. The results of the experiments with this general program indicate that if the SOAR action productions of the VLSI placement program are written specific enough, chunking will allow the system to learn new rules that aid in the execution of the program.

The results also show that the initial placements derived are *good* initial placements. This is to say that the placement arrived at is a good placement to hand over to a placement algorithm that is design to work on improving the placement. There are however several limitations on the programs as they currently exist.

The first limitation is that the algorithms cannot handle very large problems. Problems in the range of 20 to 50 components are typical for a module generator. The current code can only handle 5 or 6 components before running out of memory or requiring excessively long computation times. In the case of the first algorithm this arises from the large number of indifferent operators to choose from. In the case of the second algorithm this arises from a large number of operators in each state, of which several typically tie with zero tension. As each state in the lookahead space of algorithm may contain 90 or so operators for a problem specification with 10 components and 8 or 9 levels of lookahead are generated, memory is quickly consumed.

There are several approaches to remedy these problems. The first is an integration of algorithm 1 and algorithm 2. If the tension calculations of algorithm 2 were used in algorithm 1, the choice of which operator is best would be clearer and there would be a much smaller set of indifferent operators. If, in addition to the tension, the size of the component were also used as a selection criterion when the tension was tied, there would be a further reduction of the number of indifferent operators. A final selection criterion of choosing the operator with the lowest x and the lowest y coordinates for the proposed location of the component would further reduce the number of indifferent operators. With the number of indifferent operators reduced, the derivation of an initial placement will progress much more quickly and require significantly less memory. These additional preferences can also be added to algorithm 2.

These approaches and the completion of an X windows graphical interface to show the progress of the algorithm as it executes are the focus of continuing research. Also to be completed in continuing research is a complete placement algorithm that uses the initial placement result as a starting configuration.

Additional detail of the summer research and the complete SOAR code for this project will be published in a Wright Laboratories technical report.

ACKNOWLEDGMENT

The authors wish to express their sincere appreciation to Chahira Hopper for her support, encouragement, and guidance during this research.

REFERENCES

1. Walker, R.A. and R. Camposano, ed. *A Survey of High-Level Synthesis Systems*. 1991, Kluwer Academic: Boston, MA. 182.

2. Laird, J., *et al.*, *Soar User's Manual: Version 5.2*. 1990, Carnegie Mellon University:

3. Caviglia, D.D., *et al. Pre-placement of VLSI Blocks through Leaning Neural Networks*. in *European Design Automation Conference*. 1990. IEEE.

4. Razaz, M. and J. Gan. *Fuzzy Set Based Initial Placement for IC Layout*. in *European Design Automation Conference*. 1990. IEEE.

5. Esbensen, H. *A Genetic Algorithm for Macro Cell Placement*. in *Euro-DAC '92*. 1992. Hamburg, Germany: IEEE.

6. Hill, D., *et al.*, *Algorithms and Techniques for VLSI Layout Synthesis*. 1989, Boston, MA: Kluwer Academic. 212.

7. Sechen, C., *VLSI Placement and Global Routing Using Simulated Annealing*. 1988, Boston, MA: Kluwer Academic. 278.

8. Stauffer, A. and R. Nair. *Optimal CMOS Cell Transistor Placement: A Relaxation Approach*. in *International Conference on Computer Aided Design*. 1988. Santa Clara, CA: IEEE.

9. Swartz, W. and C. Sechen. *New Algorithms for the Placement and Routing of Macro Cells*. in *International Conference on Computer Aided Design*. 1990. Santa Clara, CA: IEEE.

MODELING OF PULSATING JET IN CROSSFLOW USING VORTEX ELEMENT METHODS

Mark A. Dietenberger
Associate Research Physicist
Environmental Science Group
Materials Engineering Division

University of Dayton Research Institute
300 College Park Ave.
Dayton, OH 45469

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

MODELING OF PULSATING JET IN CROSSFLOW USING VORTEX ELEMENT METHODS

Mark A. Dietenberger
Associate Research Physicist
Environmental Science Group
Materials Engineering Division
University of Dayton Research Institute

Abstract

The vortex element methods were investigated for predicting the behavior of a pulsating turbulent jet injected through a circular orifice into a turbulent crossflow. Published experimental results showed that modulating the diluting jets in a combustor or the cooling jets on a turbine blade can potentially increase the combustor's operating temperatures for better jet engine efficiency. In order to solve for the transient development of the jet's curved trajectory, cross-section shape, and mixing in the crossflow, the time-dependent Navier Stokes equations were descretized using the vortex element methods. The resulting system of first order ODE's described the changing motions and strengths of many discrete vortices. In devising a practical numerical scheme, the problem was analyzed in two stages. The first stage was to solve for the pulsating circular free jet. The results could then be used to calibrate the axial integral momentum equation for unsteady and quasi-irrotational flow in the axial direction. For the tranverse direction of the jet, the vorticity equation was solved for a quasi-two-dimensional transient flow around a pulsating and bending jet. Thus the response of toriodal and bound vortices to the effects of forcing and baroclinity can be calculated and can predict the jet's motion and mixing in crossflow.

MODELING OF PULSATING JET IN CROSSFLOW USING VORTEX ELEMENT METHODS

Mark A. Dietenberger

1. INTRODUCTION

In recent years, researchers have been investigating active flow concepts for improving the turbine engine performace [1,2]. Most investigators have focused on the active control of surge and stall of compressors and of combustion instabilities and mixing. There is a tendency to view any flow control in the turbine hot sections in a passive sense. Any moving parts in these sections would not last long. Indeed, no material for use in these sections has been developed to withstand stiochiometric temperatures without some air flow cooling. Estimates are that such materials are at least a couple decades in the future. The design methodology in the turbine engine industry for the air flow cooling of critical hot components are based on steady flow concepts. Recent experimental and theoretical results demonstrate such concepts can lead to large errors because of the effects of the flow unsteadiness.

One such example is the under-prediction by a factor of two for the heat loads on rotating turbine blades with steady state solutions of Navier Stokes equations. It is now recognized that large scale eddies [3] and periodic transonic flows [4] on turbine blades affect the heat transfer rates significantly; even causing variations of air transpiration cooling. Thus the magnitude of large scale turbulence must be minimized in addition to obtaining the correct temperature pattern at the combustor exit. This problem of dilution cooling is further excerbated by the less availibility of cooled air due to the increased air flow needed for combustion at higher operating temperatures. Some researchers have suggested the modulation (or pulsation) of the diluting/cooling jets to ultimately reduce the cooling loads required by the turbine blades. Indeed, Vermeulen, et. al. [5] demonstrated that by acoustically driving the jet at a critical Strouhal frequency, the air jet penetrated further into the crossflow, the mixing level increased, and the turbulence level decreased more rapidly downstream. This concept is promising also because the moving parts for driving the jets can be remote from the hot sections. However their results bring up a number of questions. One question is how would pulsating an air jet from the turbine blade affect the critical blowing ratio for effective cooling. Another question is how would the density ratios between the cooling jet and the crossflow affect the entrainment process. To help answer these questions, a mathematical model for the pulsating jet in crossflows is needed. Such a model not only would provide physical insight to the phenomena involved, it would established certain important processes as well as provide guidance to experimentation.

Several models of jet mixing in crossflows have been published. The simpler ones involve integral momentum equations with empirical entrainment functions [6,7,8], while the most complicated solve the steady Navier Stokes equations (NSE) with a k-e turbulence model [9,10]. One model analytically solved for the steady bound vortices so that the kidney shaped cross-section could be predicted as the jet was deflected by the crossflow [11]. The model still relied on empirical growth functions in conjunction with steady integral momentum equations. None of the above models are capable of simulating the pulsating jet, and really provide no guidance in answering the questions posed earlier. In principle one could solve the transient NSE using adaptive grids to provide the physical insight. However, to obtain the required resolution at high Reynolds numbers within practical resources is a couple of decades in the future. The main numerical problem is the difficulty in capturing the vortex sheet along the jet boundaries within the crossflow [12].

Upon further literature search for related flow geometries, the solution of the transient vorticity equation with the vortex element methods appeared to provide the essential features that can be applied to the pulsating jet flow in crossflow. In support of this approach we quote Ashurst and Meiburg [13].

> Three-dimensional vortex dynamics simulations are possible on today's computers. Lagrangian tracking of the (vortex) filaments replaces difficulties associated with numerical diffusion in finite difference solutions of the Navier-Stokes Equations. Since only the vorticity field must be discretized instead of the complete velocity field, there is also the economy of an adaptive mesh. A discouraging feature is the ever increasing growth rate of the filament arclength in three-dimensional inviscid flows. Vortex stretching renders the calculations very costly or requires a very small computational domain. Even with these limitations, the numerical integration of this dynamical problem presents a promising way to gain insight into the evolution of a shear layer(s), since this approach allows us to specify precisely initial perturbations in a nominally two-dimensional flow and then track their growth well into the nonlinear regime.

Their assessment of the vortex dynamics method are encouraging, particularly in that NSE has been reduced to a system of first order ODE's to describe the transient motions and vectors of discrete vortices. Perturbations in the initial conditions can be input to the equations. See also the review articles in References 14 and 15. In problems involving a relatively smooth continuum of vorticity, or variable core structure within a single tube of vorticity, a dense and several thousands of vortex or filaments are utilized and usually require a supercomputer. Studies of 2-D shear layer instabilities leading to 3-D large scale eddies are examples of these problems [14,15,16,17]. However, according to Leonard [14], "experience has shown that some fairly complex three-dimensional flows may be simulated with a nundred or so vortex elements requiring only modest computer power." The focus in this paper is to develop the modest modeling approach, which provides good approximations to the experiments on pulsating variable-density jets in crossflows. Thus, the following paragraphs describe

examples of modeling the vortex dynamics for specific experiments.

The first example is a study of the two-dimensional motion of a cylinder of fluid released from rest into a flow that is uniform far upstream of the cylinder [18]. Initially, the crossflow flows around an effective solid cylinder of curtains. Upon release of the curtains, a discontinuity of tangential flow velocities exist between the crossflow and the rested cylindrical fluid. The normal velocities and the pressure is continuous across the interface between the initially 2-D irrotational fluids. A vortex sheet is forced "instantaneously" to exist at the cylindrical interface, which then causes the rested fluid to accelerate nonuniformly throughout its cross-section area. In the reference cited above, three different vortex element methods were utilized to approximate the acceleration of the initially rested fluid. In the vortex-sheet method, any distribution of vorticity on the interface describes a flow that satifies the irrotational potential flow off the interface. Since the vortex sheet is transported with the fluid, knowledge of the flow field as derived from the potential function for the Poisson equation of vorticity, determines the shape and position of the interface. The continuity of pressure across the interface determines the time evolution of the distribution of vorticity on the interface. However, only short time predictions could be made with this approach, because the solution does not converge beyond a critical time due to roundoff errors. The vortex-blob method overcomes this problem by desingularizing the evolution of vortex sheets, but only applies only when the density ratios between the two fluids are unity. In the third method the sheet is divided into discrete vortices of zero core radius that randomly walk outside the interface. The velocity field is calculated from the vorticity stream function on an Eulerian grid, while the density or vorticity field is calculated in a Lagrangian fashion from the discrete vortex elements for preserving the sharp interface. With this model, Rottman, et. al. [18] simulated the time dependent developments of the Karman-Bernard vortex street, drag and lift coefficients, and non-uniform acceleration of the rested fluid. This random walk/vortex-in-cell method was not extended to periodic forcing nor to variable density ratios.

The periodic forcing of round free jets with uniform densities has been studied both computationally and experimentally. Quoting from a review by Leonard [14], we note that:

> ...round jets, with only gentle perturbations applied, show interesting, erratic behavior of the jet fluid. By pulsing the jet flow axially at a frequency $f \approx 0.5\ U_{av}/D$, one obtains a steady stream of nearly circular vortex rings in the near-field. Farther downstream the rings collide, merge, and become three dimensional, forming a classical turbulent jet. But, if, in addition, the jet nozzle is subjected to a small, steady orbital excitation around the azimuth, the included angle of the jet increases dramatically from 25 degrees up to 80 degrees...
>
> The dynamics of this process for the bifurcating jet, are being studied computationally using the thin-filament approximation (of the

vortex element methods.) The results indicate that the axial spacing between rings, which is proportional to $1/f$, is a crucial parameter... At the present time the computational results require a significantly stronger azimuthal perturbation than that used in the experiment. The discrepency could be due to errors in the thin-filament approximation or to neglect of the additional potential-flow term required to maintain flow tangency at the nozzle wall.

Although the thin filaments have other limitations [14], they are successful in simulating 3-D pulsating jet flows using only modest computer power. Agui and Hesselink [19] described a similar vortex element menthod for simulating their detailed experimental results on the coflowing jets. Their vortex sheet at the interface between the coflowing jets was discretized by many filament rings, which captured the periodic forcing of the jets. Thus they could simulate the initial Kelvin-Helmholtz instability leading to formation of the single strong vortex ring. So far the examples given represent the extreme blowing ratios for a pulsating jet in a crossflow. All are simulated with the vortex element method and indicate the types of vortices generated from the shear layer.

One more example in what follows demonstrate the effect of density ratios (or baroclinity) in conjunction with forcing on inviscid vorticity layers [20]. Although the geometry is a planar shear layer, the computational results provide additional insights to what could happen for a pulsating cold jet injected into a hot crossflow. Their numerical scheme discretizes the vorticity and density-gradient into a finite number of transport elements. The evolution of the flow and density is obtained by transporting the elements along the particle trajectories. Vorticity varies according to the local acceleration in a heterogeneous flow, while the density gradient changes according to the deformation of the flow map. By assuming initial error functions for velocities and densities, the initial vorticity and density-gradients distributions are obtained by differentiation into gaussian functions. Thus transport elements are located across the shear layer as well as along the mean flow direction. This approach to the vortex dynamics simulation leads to predictions of the growth rate, volumetric entrainment ratio, and eddy convective speed for the developing shear layer. Quoting Soteriou, et. al. [20], these predictions show that,

> ...in abscence of external forcing, density variation affects the development of the Kelvin-Helmholtz instability of the vorticity layer, both in its linear and non-linear stages. At fixed velocity ratio, the layer spreading rate and entrainment ratio increase with increasing the density of the low speed stream. The phase speed of the Kelvin-Helmholtz wave, and the convective speed of the eddies which form following its maturation, decrease with increasing the low speed stream density. For the density and vorticity profiles considered, the eddy convective speed and phase speed of the linear instability waves almost coincide. The computed results are found to be in good agreement with experimentally obtained data.
>
> The dynamics of the vorticity field in the variable density layer are strongly affected by high-amplitude in-phase forcing at the fundamental and subharmonic frequencies. For short downstream

> distances, external forcing leads to an earlier growth and maturation of the Kelvin-Helmholtz waves... Pairing of the eddies yields composite structures whose shape and size differ according to baroclinic vorticity generation. These mechanisms occur in such a way as to enhance the spreading rate of layer at low or high momentum ratios, and to substantially decrease the latter when the momentum ratio is near unity. (See Figure 1 that shows this result.) On the other hand, the convective speed of the eddies and entrainment ratio within the layer are weakly affected by external forcing.

These numerical results could be qualitatively extrapolated to predict some behavior of the pulsating cold jet injected into a hot crossflow. The earlier and stronger development of toriodal vortices of the modulated jet exiting a wall orifice corresponds to external forcing of the Kelvin-Helmholtz waves. Figure 1 shows that a pulsating cold jet (ie. relatively less dense low speed layer) could enhance the process further. In the development of the bound vortices behind the jet due to the crossflow, the effect of baroclinity can have opposing effect on the growth rate for with or without forced modulation. To determine the effect of baroclinity and forcing on the development of the bound vortices requires analysis of the shear layer development at the jet/crossflow interface. It is apparent from above discussion that a single thin vortex sheet at the interface is insufficient for simulating these effects. Rather, the most promising approach is utilization of the transport/vortex element method by Soteriou, et al [20] as applied to the pulsating jet in crossflow. Thus in Section 2, the fundamentals of the vortex element method are presented. The discretization of the viscous term of NSE is included as suggested by Fishelov [21]. To avoid a complicated 3-D analysis we propose a quasi-two-dimensional, two-staged approach to analyze the transient jet development. The first stage is the free jet simulation with axisymmetric toriodal vortices. This is discuss in detail in Section 3. It is assumed that upon interaction with the crossflow at the boundary layer thickness, the ring vortices become very unstable and twist to merge with the developing bound vortices. The jet will begin to develop a kidney shaped cross-section and to deflect in the direction of the crossflow. Section 4 discusses in detail the quasi-two-dimensional approach for simulation of the developing, deflecting jet. The axial momentum equation is approximated as an unsteady, quasi-irrotational flow in the integral form. This equation should approximate the pulsating free jet as a limiting solution. The momentum equation in the cross-section directions is approximated by the two-dimensional vorticity dynamics. The conclusion and recommendation is in Section 5.

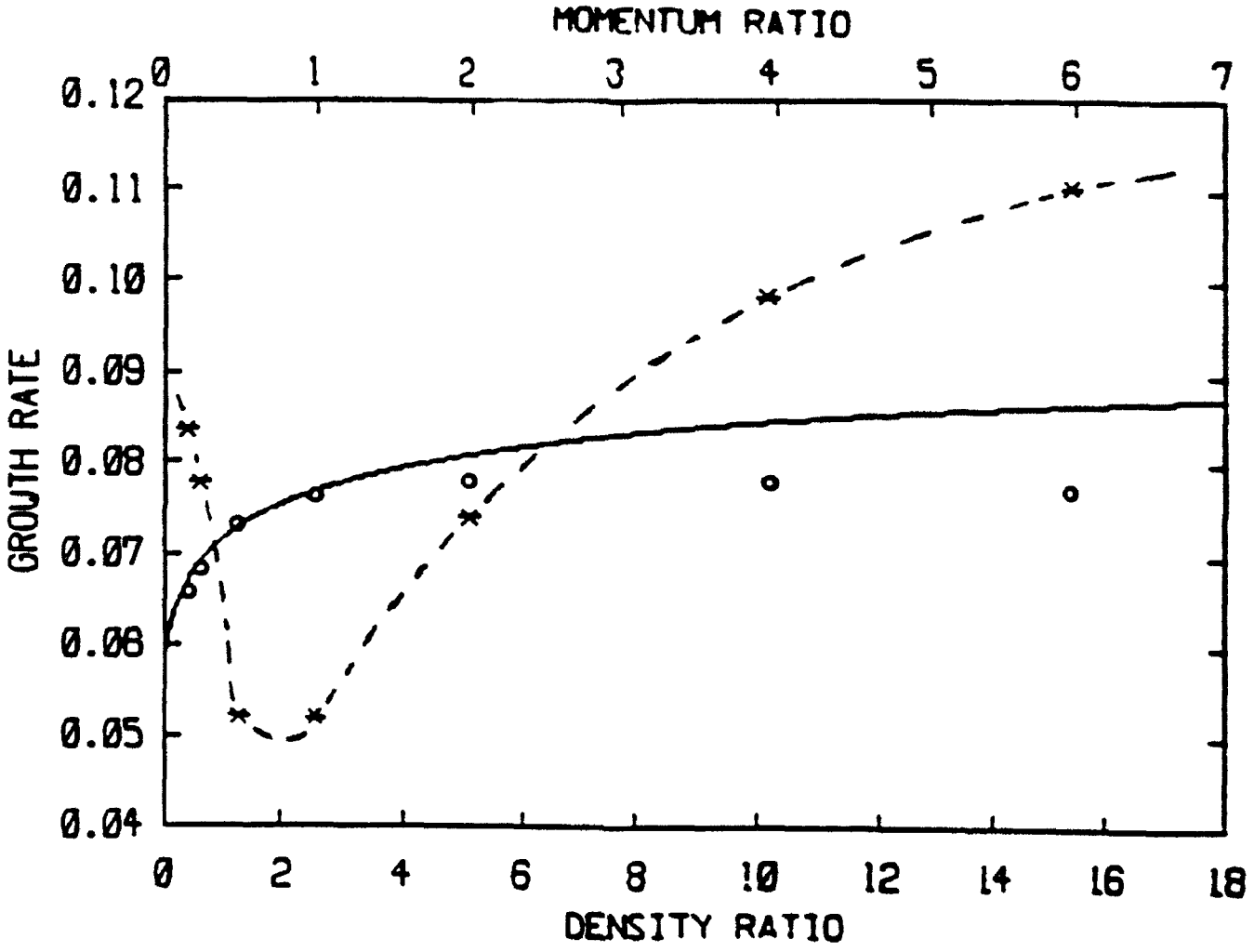


Figure 1. Predicted spatial growth rate of forced shear layers at velocity ratio of 0.625 is indicated by dash line. Results for unforced shear layer agrees with experimental data (circled points) and a formula (solid line). Figure is from [20].

2. FUNDAMENTALS OF THE VORTEX ELEMENT METHOD

Vortex Dynamics

The nondimensional form of the governing equations for an unsteady, isentropic, variable-density flow are given as:

$$\frac{D\rho}{Dt} = 0 \tag{1}$$

$$\nabla \cdot \boldsymbol{u} = 0 \tag{2}$$

$$\rho \frac{D\boldsymbol{u}}{Dt} = -\nabla p + \frac{1}{Re}\nabla^2 \boldsymbol{u} \tag{3}$$

where ρ is density, $\boldsymbol{u}$ is the velocity vector, p is pressure, $Re = UL/\nu$ is the Reynolds number, and t is the non-dimensional time. U and L are typical velocity and length, respectively and $\frac{D}{Dt}$ is the material derivative. The governing equations are recast into the vorticity form by taking the curl of Eq. (3) and into the density-gradient form by taking the gradient of Eq. (1). The result is:

$$\frac{D\boldsymbol{\omega}}{Dt} = \boldsymbol{\omega}\cdot\nabla\boldsymbol{u} + \frac{1}{Re}\nabla^2\boldsymbol{\omega} - \frac{\boldsymbol{g}}{\rho}\times\left(\frac{D\boldsymbol{u}}{Dt} - \frac{1}{Re}\nabla^2\boldsymbol{u}\right) \tag{4}$$

$$\frac{D\boldsymbol{g}}{Dt} = -\boldsymbol{g}\cdot\nabla\boldsymbol{u} - \boldsymbol{g}\times\boldsymbol{\omega} \tag{5}$$

where $\boldsymbol{\omega} \equiv \nabla\times\boldsymbol{u}$ and $\boldsymbol{g} \equiv \nabla\rho$ are the definitions for vorticity and density-gradient, respectively. The first term on the right side of Eq. (4) represents rate of change of vorticity due to stretching and tilting of vortex lines. The second term over is represents the rate of change of vorticity due to diffusion of vorticity. The last term is the density gradient crossed with the pressure gradient; it represents the baroclinity that generates the vorticity. Note that the pressure has been eliminated from Eq. (4) by using Eq. (3) directly for the pressure gradient. Eq. (4) can reduce to simpler forms for very high Reynolds number, or for constant density flows, or for two-dimensional flows.

If the vorticity distribution is known, (that is, from numerical integration of Eq. (4)), the velocity vector can be evaluated from the integration of Eq. (2) and the vorticity defining equation. Based on the

uniqueness of the decomposition of a vector field, the velocity can be split into two components,

$$\boldsymbol{u} = \boldsymbol{u}_{\omega} + \boldsymbol{u}_{p} \tag{6}$$

where $\boldsymbol{u}_{\omega}$ is a solenoidal field and $\boldsymbol{u}_{p}$ is a potential field satisfying the boundary conditions, such as the cross-flow velocity at infinity. The solenoidal field is determined from the well-known Biot-Savart law,

$$\boldsymbol{u}_{\omega}(\boldsymbol{x}) = \int \boldsymbol{K}(\boldsymbol{x}-\boldsymbol{x}^{*}) \times \boldsymbol{\omega}(\boldsymbol{x}^{*})\, d\boldsymbol{x}^{*} \quad , \tag{7}$$

where

$$\boldsymbol{K}(\boldsymbol{x}) = \frac{-\boldsymbol{x}}{4\pi |\boldsymbol{x}|^{3}} \quad , \tag{8}$$

is the kernel of the Poisson equation. The analogue of the Biot-Savart integral was derived by Knio and Ghoniem [16] to solve for the density from the integral,

$$\rho(\boldsymbol{x}) = \int \boldsymbol{g}(\boldsymbol{x}^{*}) \cdot \nabla G(\boldsymbol{x}-\boldsymbol{x}^{*})\, d\boldsymbol{x}^{*} + \rho_{p}(\boldsymbol{x}) \quad , \tag{9}$$

where,

$$G(\boldsymbol{x}) = \frac{1}{4\pi |\boldsymbol{x}|} \quad , \tag{10}$$

is the Green function of the Poisson equation in three-dimensions. Although a new spatial variable, $\boldsymbol{x}$, has been introduced, it is solved for by integrating along the characteristic line,

$$\frac{D\boldsymbol{x}}{Dt} = \boldsymbol{u} \quad , \tag{11}$$

along which the vorticity and the density-gradient are given by Eqs. (4) and (5), respectively. Thus a numerical scheme could follow a sequence of integrations. Starting with an initial distribution of vorticities and density-gradients, the first step is to evaluate the integrals in Eqs. (7) and (9) for each time step. The next step is to evaluate the material time derivative, gradient, and Laplacian of the velocity vector field by numerical approximations or summations for the time step. The Laplacian of the vorticity needed in Eq. (4) can be evaluated by using the technique by Fishelov [21]. The final step is to numerically integrate the coupled Eqs. (4), (5), and (11) over a time step. A

practical evaluation of the integral in Eqs. (7) and (9) requires a process of discretization in the spatial domain. The discretization of vorticities and density-gradients can be concentrated in the shear layer for sufficient accuracy. This is called the vortex/transport element method.

Vortex/transport Element Method

An initial uniform grid $\boldsymbol{x}_j(0)$, $j=1,\ldots,n$ is constructed with spacing h_1, h_2, h_3 for a three-dimensional problem and h_1, h_2 for a two-dimensional one. For simplicity, assume spacings in all directions are equal. In the vortex method, the vorticity field is discretized between elements centered at $\boldsymbol{x}_j(0)$, so that,

$$\boldsymbol{\omega}^h(\boldsymbol{x},t) = \sum_{j=1}^{n} f_\delta(\boldsymbol{x}-\boldsymbol{x}_j)\,\Gamma_j^h \quad , \tag{12}$$

where $\Gamma_j^h = h_j^N\,\boldsymbol{\omega}(\boldsymbol{x}_j,t)$ and $N=2,3$ is the dimension of the problem. The vorticity associated with each element is smoothed in a small neighborhood of $\boldsymbol{x}_j$ according to a core function f_δ with a core radius δ. The core function is chosen to satisfy the conditions: (1) $\int f_\delta\,d\boldsymbol{x} = 1$ and f_δ converges to the Dirac delta function as δ approaches 0; (2) the induced velocity field away from the core is the same as that induced by a concentrated vortex element, i.e. the core function decays at a fast rate; and (3) the velocity field of a finite vortex element is non-singular at its center. Thus the core function is,

$$f_\delta(\boldsymbol{x}) = \frac{f(|\boldsymbol{x}|/\delta)}{\delta^N} \tag{13}$$

where $f(r)$ is called the cutoff function. The accuracy of the discretization in Eq. (12) depends on: (1) the ratio δ/h (i.e. greater than one); (2) the choice of the core function; and (3) the scheme used to determine the values of ω_j. A second-order cutoff function for 2-D fluid flows has been suggested [20,21] as, $f(r) = (1/\pi)\exp(-r^2)$, while for 3-D fluid flows it has been suggested [16] as, $f(r) = (3/4\pi)\exp(-r^3)$. With these functions the stability of the vortex element method is ensured and is second-order discretization. The most accurate initialization of discrete vortices has been found by solving a system of linear equations resulting from the application of

Eq. (12) to the mesh points x_j, $j=1,\ldots,N$. The discretization of the initial density-gradients is analogous to that of vorticity. That is, replace ω by g in Eq. (12) and the transport elements are created [20]. Now that discrete vortices and density-gradients have been calculated, they can be used to discretized the integrals in Eqs. (7) and (9) and the terms in Eqs. (4) and (5).

The discrete desingularized version of the Biot-Savart law is derived as,

$$u^h_{\omega_i} = \sum_{j=1}^{n} K_\delta (x^h_i(t) - x^h_j(t)) \times \omega^h_j(t)\, h_j^N \quad . \tag{14}$$

The singular kernel $K(x)$ is approximated by a smoothed one $K_\delta(x)$, where $K_\delta = f_\delta * K$. The end result of this convolution using the cutoff functions from the previous paragragh is to multiply K by the scalar term, $\kappa = 1 - e^{-|(x_i - x_j)/\delta|^N}$. The discrete desingularization of Eq. (9) for the integral of density-gradients is derived to be [16,20],

$$\rho^h_i(x_i) = \sum_{j=1}^{n} (g^h_j(t)\, h_j^N) \cdot \nabla G_\delta (x^h_i(t) - x^h_j(t)) + \rho_p(x_i) \tag{15}$$

where the gradient of the Green's function is,

$$\nabla G_\delta(x) = \frac{\kappa x}{2^{N-1}\pi |x|^N} \quad . \tag{16}$$

The discretization of the terms in Eqs. (4) and (5) involving gradients and Laplacians are as follows. The gradient of the velocity vector can be approximated by taking the analytical gradient of Eq. (14). The gradient can be taken to the inside of the summation and then the gradient of the smoothed kernel is derived. Likewise the Laplacian of the velocity vector would involve the Laplacian of the smoothed kernel. Following Fishelov [21] procedure, the Laplacian of the vorticity in the viscous term can be approximated by taking the analytical Laplacian of Eq. (12). The Laplacian can be taken inside the summation and the Laplacian of the core function f_δ can then be performed. The material derivative of the velocity vector in Eq. (4) is approximated by a backwards euler's time differencing [20]. The discretization of the vortex dynamics equations are completed, so that the next steps are to specify the flow geometry and boundary conditions.

3. PROBLEM STATEMENT FOR PULSATING CIRCULAR FREE JET

Consider the circular free jet geometry shown in Figure 2. The jet will penetrate a still (or even a co-flowing) fluid at some mean velocity and density.

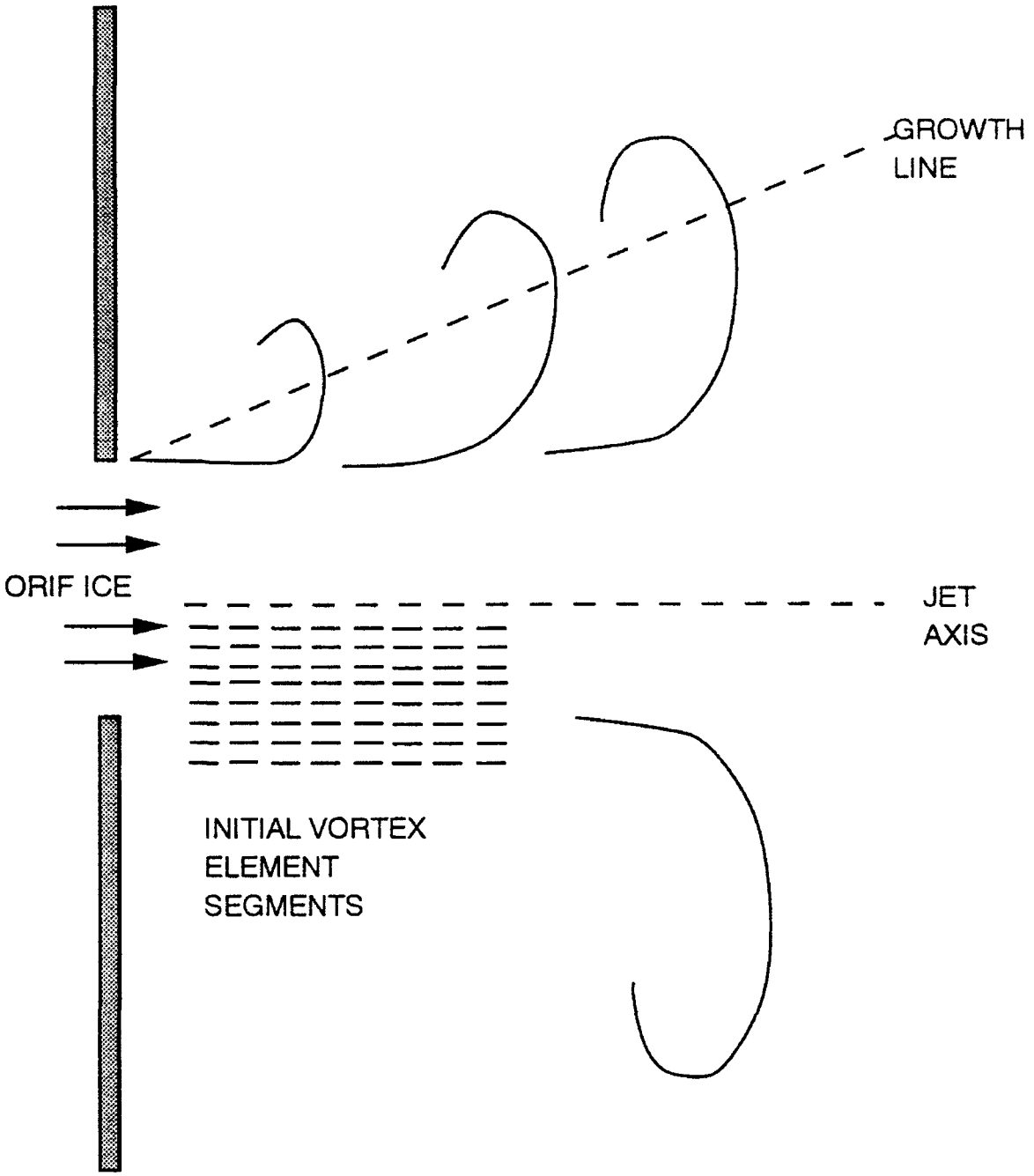


Fig. 2 Schematic of Circular Free Jet and Location for Vortex Elements in the Shear Layer

Although the vortex discretization is actually three dimensional, it can be made more computationally efficient by enforcing a symmetry in the azimuthal direction. Our main interest is to establish the initial growth rate of the shear layer as the toriodal vortices develop. Thus we ignore the instabilities of the vortex rings that cause kinking of the vortex lines. Since the shear layer is the greatest at the jet/free air interface, we will concentrate the discrete vortices around the interface, in both radial and axial direction. This is shown in Figure 2 in a manner analogous to Soteriou, et. al. [20] for their planar shear layer. Because of the boundary conditions of the wall, an appropriate image system of transport elements should be used to compute the potential components of velocities and densities. At the downstream section, a condition of vanishing vorticity and density-gradient is used as outflow boundary condition, and is applied by removing the transport elements which cross the maximum axial distance. On the other hand, the extreme stretching of vortex elements can lead to addition of transport elements to avoid losing accuracy of the numerical scheme. Also in the same manner as Soteriou, et.al., the initial velocity and density profiles are assumed to follow an error function curve as follows. The inlet fluid velocity has zero cross-stream component and a streamwise component is given by,

$$U(z=0, r, t) = \frac{U_j + U_f}{2} + \frac{U_j - U_f}{2} erf\left(\frac{r - \bar{r}}{\sigma}\right) \quad . \tag{17}$$

This is equivalent to the initial vorticity distribution as,

$$\omega(z=0, r, t) = - \frac{U_j - U_f}{\sqrt{\pi}\sigma} \exp(-(r - \bar{r})^2/\sigma^2) \quad . \tag{18}$$

This distribution is used to initialize the discrete vortices using the method described in Section 2 under the vortex/transport element method. The density and density-gradients profiles are assumed to have similar distributions. That is, replace U by ρ in Eqs. (17) and (18) and replace ω by $\partial\rho/\partial r$ in Eq. (18). The last set of initialization is in defining various parameters such as orifice diameter, initial shear layer dimensions, core radius, element spacings, imposition of periodic function on initial vorticity and density-gradients, fluid properties such as viscosity and densities, and the jet/fluid mean velocities.

The results are expected to be qualitatively similar to Figure 1. The conclusions should be similar to the discussion near the end of Section 1. The fact that toridal vortices deflect and merge rapidly into the bound vortices of a jet in crossflow, makes a full three dimensional analysis a challenging problem. By placing vortex elements in all three directions in the shear layer and redo potential flow analysis, the free jet simulation can be extended to the

jet in crossflow using a supercomputer.

4. PROBLEM STATEMENT FOR PULSATING COLD TURBULENT JET IN CROSSFLOW

A quasi-two-dimensional approach could be suggested to simulate the time development of the bound vortices. The schematic for the jet in crossflow is shown in Figure 3. The developing bound vortices defines how much the jet is spreading, deflecting and entraining into the crossflow. Near the jet entrance the sources for the bound vortices are due to combination of several effects. The first is the existence of a thin shear layer at the interface of the jet and crossflow. The initial toriodal vortices developed within the crossflow boundary layer will deflect into the bound vortices already developing in the shear layer. Further changes in the intensities of the bound vortices will result from baroclinity, viscosity, and external forcing. By treating the development of bound vortices as a two-dimensional vorticity problem, as shown by the vortex elements in Figure 3, the Eqs. (4), (5), (7), and (9) can be greatly simplified. The important processes are still retained, except for the term for calculating the stretching and tilting of vortex lines. An assumption could be made, however, that the circulation of the initial toriodal vortices is absorbed into the bound vortices within the crossflow boundary layer thickness.

While the cross-direction of the jet is being accelerated, the axial direction of the jet is also being decelerated. This deceleration is somewhat like that of the free jet. The axial deceleration of the jet will also affect the axial speed of the sprialing bound vortices. Perhaps it would be feasible to use an axial integral momentum equation for unsteady and quasi-irrotational flow that approximates the free jet in the limit of zero cross-flow velocity. Such an equation would input entrainment and growth rate from the solutions for the toriodal and bound vortices. The output would be axial velocities of jet's potential core and vortices. These velocities would apply a centrifugal force to affect the cross-direction acceleration.

5. CONCLUSIONS AND RECOMMENDATIONS

The review of the literature indicates that the modulation of the diluting jet in the combustion chamber or of cooling jets on the turbine blades are candidates for active flow control in the turbine hot sections. Vortex dynamics seems to provide theoretical explanations of recent experimental results. The enhancement in the jet spread rate at a critical forcing frequency has been observed and seems related to the development of toriodal and bound vortices of the jet. Further investigation of the vortex element methods indicate a solution procedure for the pulsating jet in crossflow. The discretization of the Navier-Stokes equations with the vortex elements method result in a system of first order ODE's. Thus active flow concepts can be introduced and an algorithm possibly could be implemented on a medium-sized computer. Although a full 3-D analysis was formulated, some ideas toward a quasi-two-dimensional approach were suggested in an attempt at practical calculations. It is recommended that better

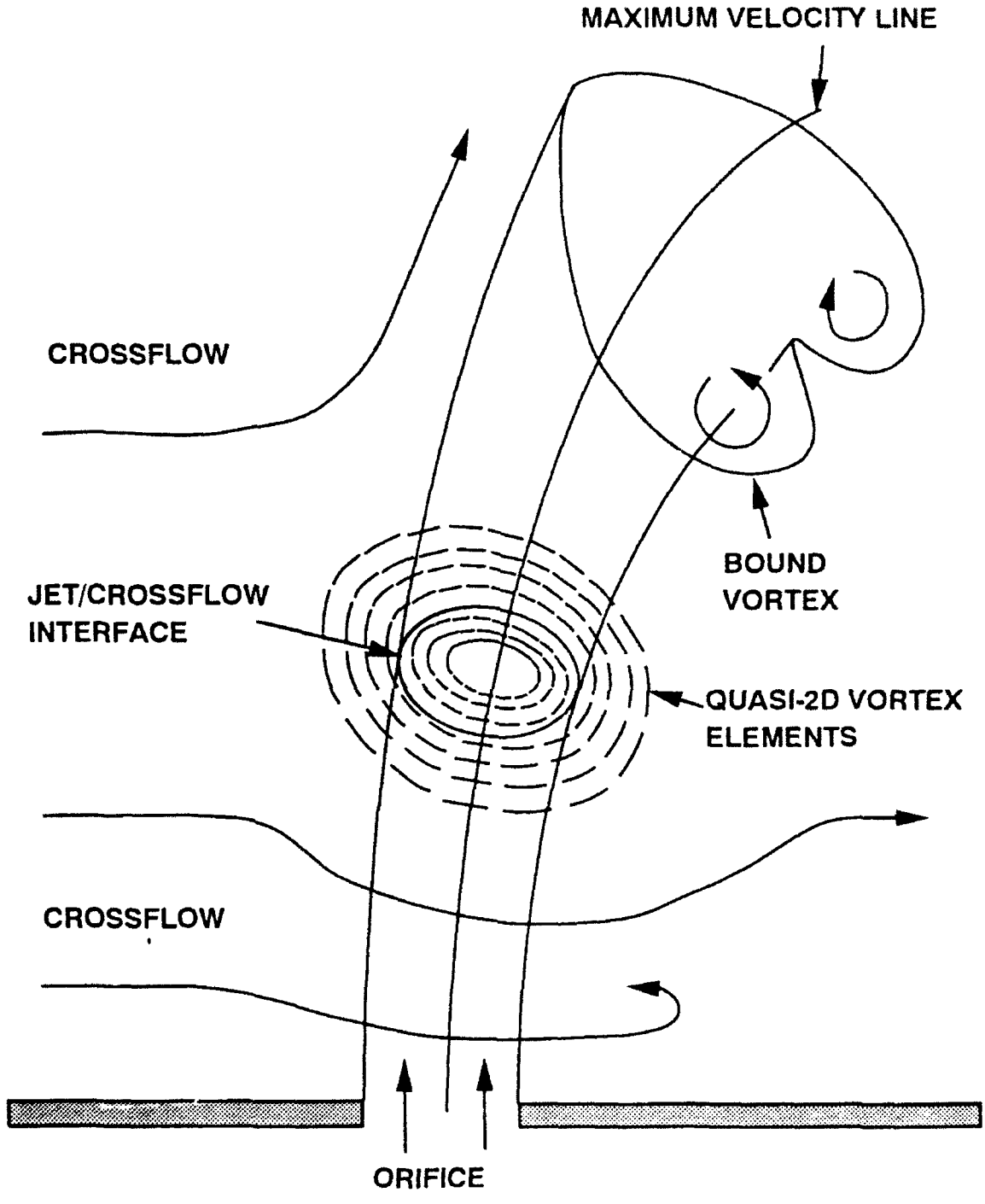


Fig. 3 Schematic of a deflecting jet in crossflow and of the cross-sectional vortex element segments in the shear layer

understanding be made of the 2-D approach and that codes be developed to demonstrate its effectiveness. To verify the kind of results shown in Figure 1, new experiments are recommended to investigate the combined baroclinity and external forcing on the diluting jet's trajectory and mixing in the confined crossflow.

REFERENCES

1. Hansen, R.J. and Hendricks, E.W., 1992, "Active Control of Complex Systems: An Overview", ASME paper 92-GT-35.

2. Hendricks, E.W., Ed. of "1991 Review of ONR Active Control Research Initiative", Conference at Penn State University, 9-10 April 1991.

3. Jumper, G.W., Elrod, W.C., and Rivir, R.B., July 1991, "Film Cooling Effectiveness in High-Turbulence Flow", Trans. of ASME, Journal of Turbomachinery, Vol. 113, pp. 479-483.

4. Shokrollah-Abhari, R., Jan. 1991, "An experimental Study Of The Unsteady Heat Transfer Process In A Film Cooled Fully Scaled Transonic Turbine Stage", Ph.D. Thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology.

5. Vermeulen, P.J., Chin, C-F., and Yu, W.K., Dec. 1990, "Mixing of an Acoustically Pulsed Air Jet with a Confined Crossflow", ASME Journal of Propulsion, Vol. 6, No. 6, pp. 777-783.

6. Novick, A.S., Arvin, J.R., and Quinn, R.E., Oct. 1980, "Development of a Gas Turbine Combustor Dilution Zone Design Analysis", Journal of Aircraft, Vol. 17, No. 10, pp. 712-718.

7. Sucec, J. and Bowley, W.W., Dec. 1976, "Prediction of the Trajectory of a Turbulent Jet Injected Into a Crossflowing Stream", Journal of Fluids Engineering, Trans. of ASME, pp. 667-673.

8. Stoy, R.L. and Ben-Haim, Y., Dec. 1973, "Turbulent Jets in a Confined Crossflow", Journal of Fluids Engineering, Trans. of ASME, pp. 551-556.

9. Schonung, B. and Rodi, W., Oct. 1987, "Prediction of Film Cooling by a Row of Holes With a Two-Dimensional Boundary-Layer Procedure", Vol. 109, pp. 579-587.

10. Fairweather, M., Jones, W.P., Lindstedt, R.P., and Marquis, A.J., 1991, "Predictions of a Turbulent Reacting Jet in a Cross-Flow", Combustion and Flame, Vol. 84, pp. 361-375.

11. Adler, D. and Baron, A., Feb. 1979, "Prediction of a Three-Dimensional Circular Turbulent Jet in Crossflow", AIAA Journal, Vol. 17, No. 2, pp.168-174.

12. Rizzi, A. and Engquist, B., 1987, "Review Article - Selected Topics in the Theory and Practice of Computational Fluid Dynamics", Journal of Computational Physics, Vol. 72, pp. 1-69.

13. Ashurst, W.T. and Meiburg, E., 1988, "Three-dimensional Shear Layers via Vortex Dynamics", Journal of Fluid Mechanics, Vol. 189, pp. 87-116.

14. Leonard, A., 1985, "Computing Three-Dimensional Incompressible Flows With Vortex Elements", Annual Review of Fluid Mechanics, Vol. 17, pp. 523-559.
15. Ashurst, W.T., 1983, "Large Eddy Simulation Via Vortex Dynamics", Paper No. AIAA 83-1879-CP.
16. Knio, O.M. and Ghoniem, A.F., 1990, "Numerical Study of a Three-Dimensional Vortex Method", Journal of Computational Physics, Vol. 86, pp. 75-106.
17. Knio, O.M. and Ghoniem, A.F., 1991, "Three-Dimensional Vortex Simulation of Rollup and Entrainment in a Shear Layer", Journal of Computational Physics, Vol. 97, pp. 172-223.
18. Rottman J.W., Simpson, J.E., and Stanby, P.K., 1987, "The Motion of a Cylinder of Fluid Released from Rest in a Cross-flow", Journal of Fluid Mechanics, Vol. 177, pp.307-337.
19. Agui, J.C. and Hesselink, L., 1988, "Flow Visualization and Numerical Analysis of a Coflowing Jet: a Three-Dimensional Approach", Journal of Fluid Mechanics, Vol. 191, pp. 19-45.
20. Soteriou, M.C., Knio, O.M., and Ghoniem, A.F., 1991, "Manipulation of the Growth Rate of a Variable Density, Spatially Developing Mixing Layer Via External Modulation", 29th. Aerospace Sciences Meeting, AIAA-91-0081.
21. Fishelov, D., 1990, "A New Vortex Scheme for Viscous Flows", Journal of Computational Physics, Vol. 86, pp. 211-224.

Laser Multiphoton Ionization Detection of Methyl Radicals in a Filament-Assisted Chemical Vapor Deposition Reactor

David A. Dolson
Assistant Professor
Department of Chemistry

Wright State University
3640 Col. Glenn Highway
Dayton, OH 45435

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored By:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

Laser Multiphoton Ionization Detection of Methyl Radicals in a Filament-Assisted Chemical Vapor Deposition Reactor

David A. Dolson
Assistant Professor
Department of Chemistry
Wright State University

Abstract

An apparatus and procedures were developed for the detection of methyl radicals in the discharge-assisted decomposition of methane/hydrogen gas mixtures flowing over a heated wire filament. The apparatus was designed to detect small transient changes in the discharge current as the focussed output of a tunable dye laser ionizes methyl radicals between a hot filament and a deposition substrate. This design does not require the use of an additional (intrusive) electrode for ion detection, and it allows for adjustment of the filament-to-substrate distance. Results from a conventional flow tube reactor suggests that the detection sensitivity for methyl radicals is better than 0.3 mtorr or 10^{13} radicals per cm^3. Preliminary data is presented for methyl radical detection in the filament reactor.

Laser Multiphoton Ionization Detection of Methyl Radicals in a Filament-Assisted Chemical Vapor Deposition Reactor

David A. Dolson

INTRODUCTION

The recent demonstrations of chemical vapor deposition (CVD) of diamond-like thin films has stimulated an overwhelming number of attempts to understand the principal chemical and physical processes that govern diamond deposition.[1] There is general agreement from modeling results and from experiments that the methyl radical ($CH_3\cdot$) is one of the "growth species" linked to the transport of carbon to the substrate where diamond growth occurs. Consequently, efforts have been directed toward quantitative measurements of [$CH_3\cdot$] using various techniques such as mass spectrometry, diode laser absorption, and resonance enhanced multiphoton ionization (REMPI).

REMPI allows for a degree of spatial resolution because ions are only created in a small focal volume. Celii and Butler[2] have used the REMPI technique to measure the relative $CH_3\cdot$ concentration in a filament-assisted diamond CVD reactor as a function of several process parameters. In their application an ion collector wire was fixed 4 mm above the substrate surface, the laser focus was approximately midway between the wire and the substrate, and the filament position was varied in order to probe the spatial [$CH_3\cdot$] profile.

This report describes the development of a REMPI detection experiment for methyl radicals that is designed for spatial profiling of [$CH_3\cdot$] in a fixed gap between the filament and

substrate. This design avoids the use of the potentially intrusive ion collector wire. Preliminary results are presented, and suggestions are made for continuing efforts.

APPARATUS and PROCEDURES

Methyl radicals were generated in a conventional fast flow tube reactor, initially, in order to develop the REMPI technique and to estimate the detection sensitivity. The reactor was constructed from 40 mm O.D. glass tubing, and o-ring compression fittings or knife-edged flange seals were used throughout. The flow reactor, previously used with laser-excited fluorescence detection, was modified by inserting two 0.040" tungsten rods into the reactor through the fluorescence observation port. A 20 cm focal length lens was used to focus the laser 3.6 cm above the end of the discharge tube, in the center of the ~10 mm electrode gap. This apparatus is depicted schematically in Fig. 1.

The upper electrode was biased at +140 volts relative to the lower electrode, which was connected in parallel to the inputs of a PAR boxcar averager (100 kΩ input impedance, 50 mV sensitivity) and an oscilloscope (1 MΩ input impedance). The transient REMPI current produced by the laser-generated electrons was detected as a small transient voltage signal across the effective 90 kΩ resistance. The signal input gate of the boxcar averager was typically delayed 5-40 μs from the laser firing and opened for ~15 μs. The excimer-pumped dye laser wavelength was scanned under computer control and the boxcar averager output was read with a digital voltmeter by the computer.

Gas flow rates were measured with electronic mass flow controllers, and total pressure was measured with a 10 torr full scale rating. CF_4 (2-3 sccm) and

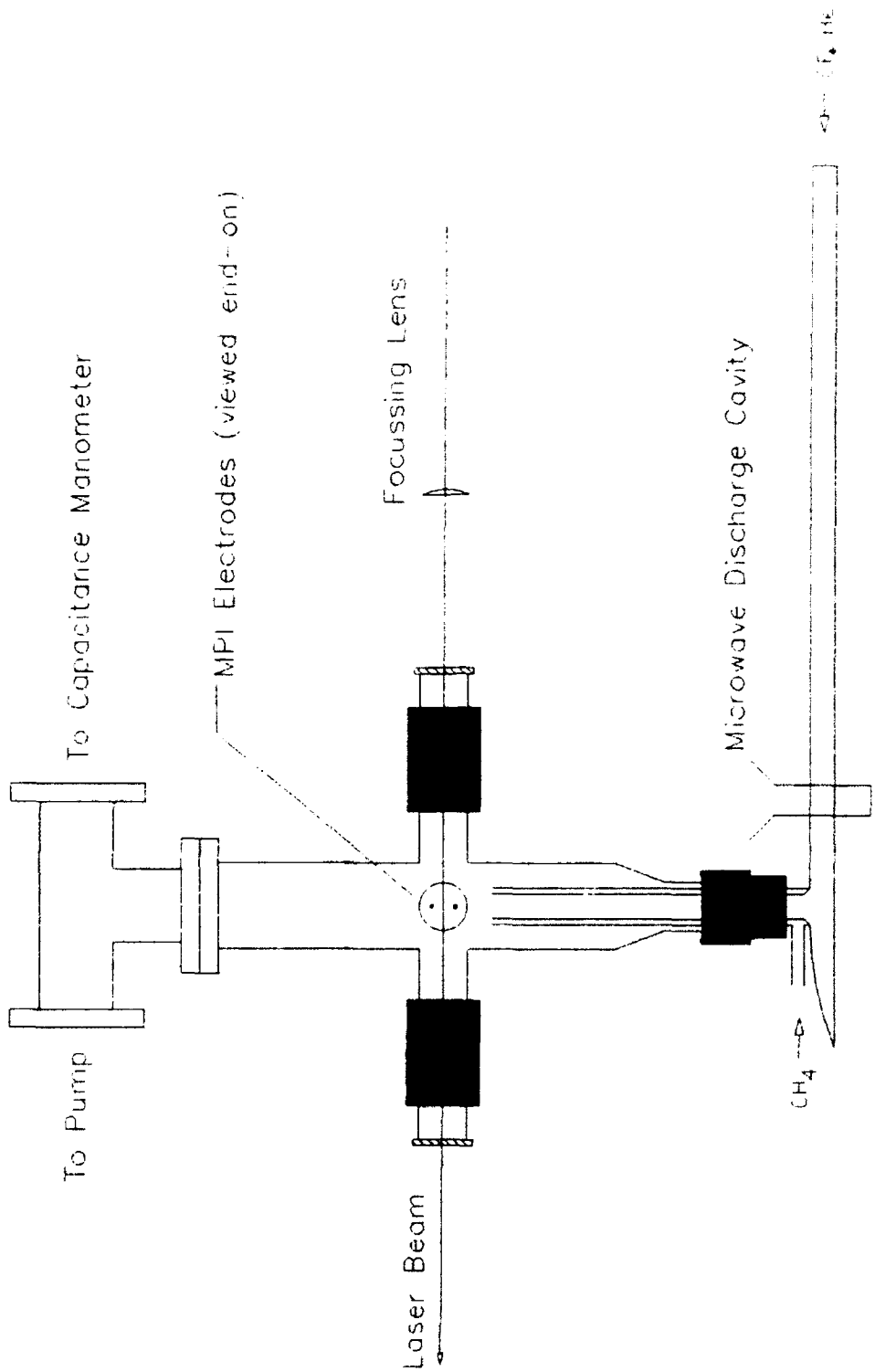


Fig. 1. Fast flow tube reactor for generation of CH_3 radicals in the reaction of methane with F atoms.

He (140 sccm) were flowed through an microwave generated plasma to generate F atoms in the flow. CH_4 (0-8.1 sccm) was mixed with this gas flow 3.6 cm below the laser focal volume. The F atoms were titrated in the gas flow by varying the CH_4 flow rate and measuring the relative $[CH_3\cdot]$ from the intensity of the REMPI signal. At the titration endpoint, $[F\cdot] = [CH_4]$ is given from the gas flow rates and total pressure. From the observed signal-to-noise ratio (S/N) of the REMPI signal and the knowledge that each F· atom produced a methyl radical, the detection limit was calculated for methyl with S/N=1.

A filament-assisted CVD reactor was constructed in order to use the REMPI probe for methyl radicals in an apparatus more closely resembling those used for diamond deposition. The CVD reactor was constructed from heavy-walled 4.5 inch O.D. glass pipe. The reactor was constructed in two sections, connected with an o-ring seal. Its overall length was ~18". Kovar rods, sealed through one section of the reactor, were used to power the filament and give it support. A 1" dia. port in this section allowed a 0.5" dia. pyrex tube to introduce a mixed gas flow at selected distances from the filament.

The length of the rods was such that the filament was located ~5 cm inside the other section when the reactor was assembled. This section (the main body of the reactor) held the windows for passing the laser, the pressure measurement and vacuum ports, and the introduction port for a linear translator for the substrate. The vacuum port was mounted to an aluminum plate for connection to a pressure control throttle valve. Flexible tubing connected this valve to the 400 $m^3\ hr^{-1}$ booster pump, which was backed by a 60 $m^3\ hr^{-1}$ two-stage rotary pump.

Several filament materials, geometries, and mounting methods were evaluated in numerous attempt to obtain a stable filament for the REMPI experiments. At the end of the project, the

most stable filament had been wound from 0.008" dia. rhenium wire, but a successful methyl probe was not possible because of a laser malfunction. REMPI signals were observed using a 0.010" dia. tungsten (G.E. "218") wire filament which was attached to the Kovar rods by winding the wire around the rod a few times and securing it with an alligator clip. The filament was heated by adjusting the current slowly and monitoring the temperature with a hot filament pyrometer. A 200 sccm gas flow of 2% CH_4 in H_2 was caused to flow over the filament toward the substrate 1-4 cm away and on to the pumping port at the far end of the reactor. The substrate was a 1" dia x 0.1" thick molybdenum disc in the current work. The filament was biased at a negative potential relative to the substrate. Electrons emitted from the hot filament are accelerated toward the substrate. This discharge current generated a voltage across a 1KΩ (1%) resistor in the circuit of Fig. 2. Transient REMPI signals were detected as a transient voltage across this resistor. The discharge-generated DC voltage was blocked by AC-coupling the inputs to a different gated integrator/boxcar averager (1 MΩ input impedance, 5 mV sensitivity) and the oscilloscope.

REMPI signals were obtained by scanning the tunable output of an excimer-pumped dye laser (Lambda Physik EMG-103MSC/FL2002) near the 333.5 nm transition in the methyl radical. Carbon atom detection was also accomplished at 340.8 nm. The dye laser delivered up to 9mJ/pulse at 334nm and 14 mJ/pulse at 341 nm using a p-terphenyl/p-dioxane dye solution. The laser FWHM linewidth was previously determined at 430nm to be 0.4 cm^{-1}. Most of the current work was done with the laser operating at 30Hz.

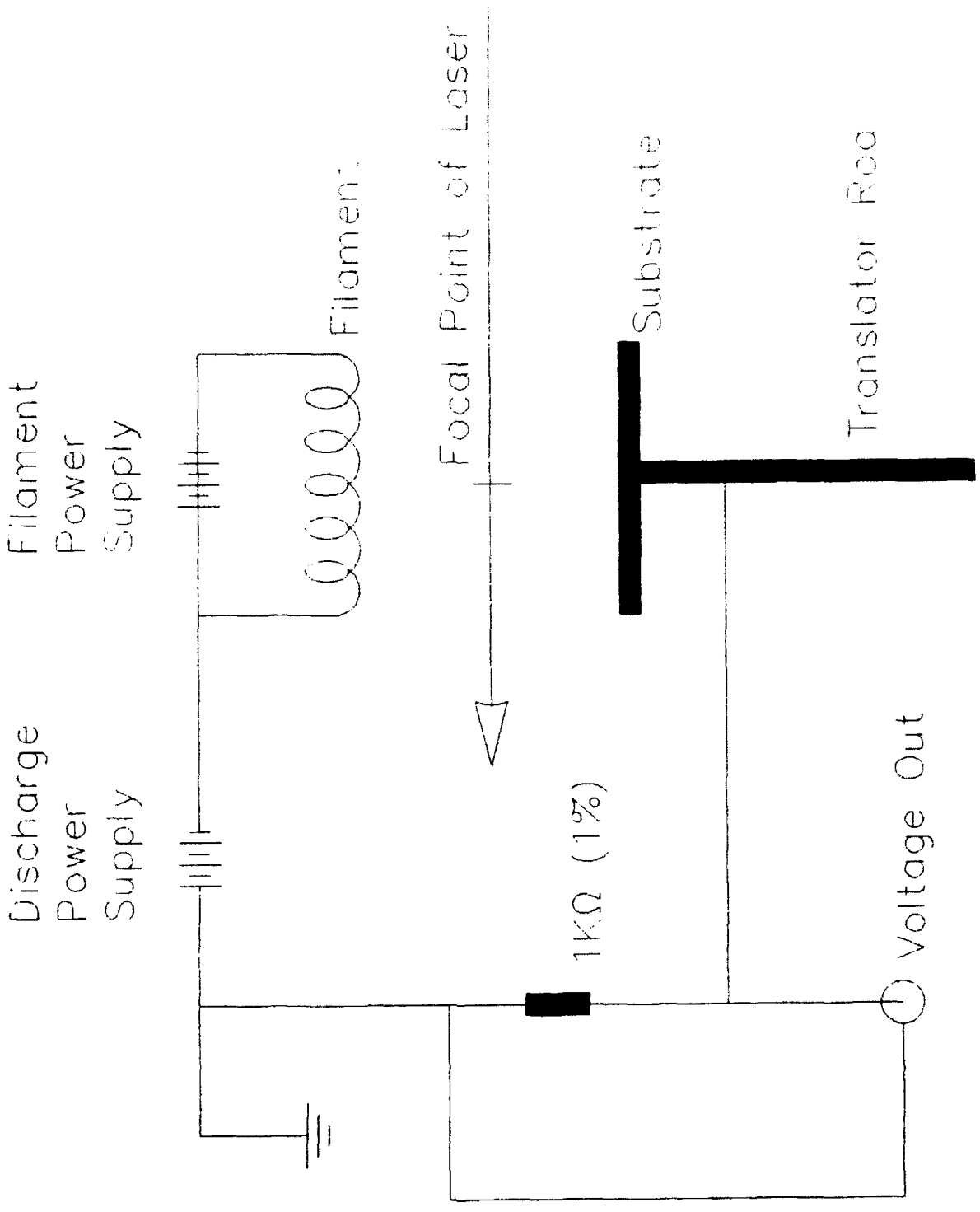
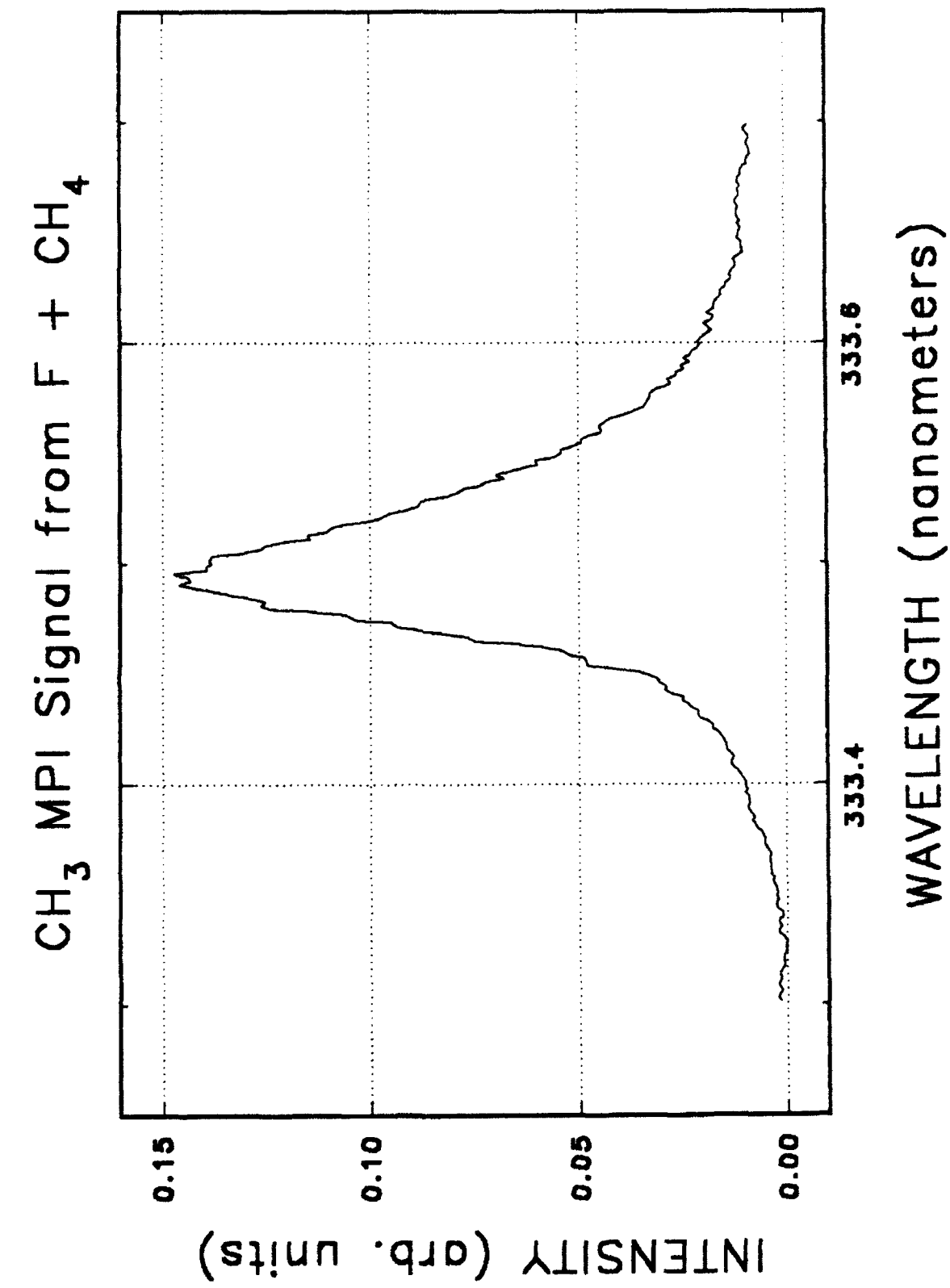


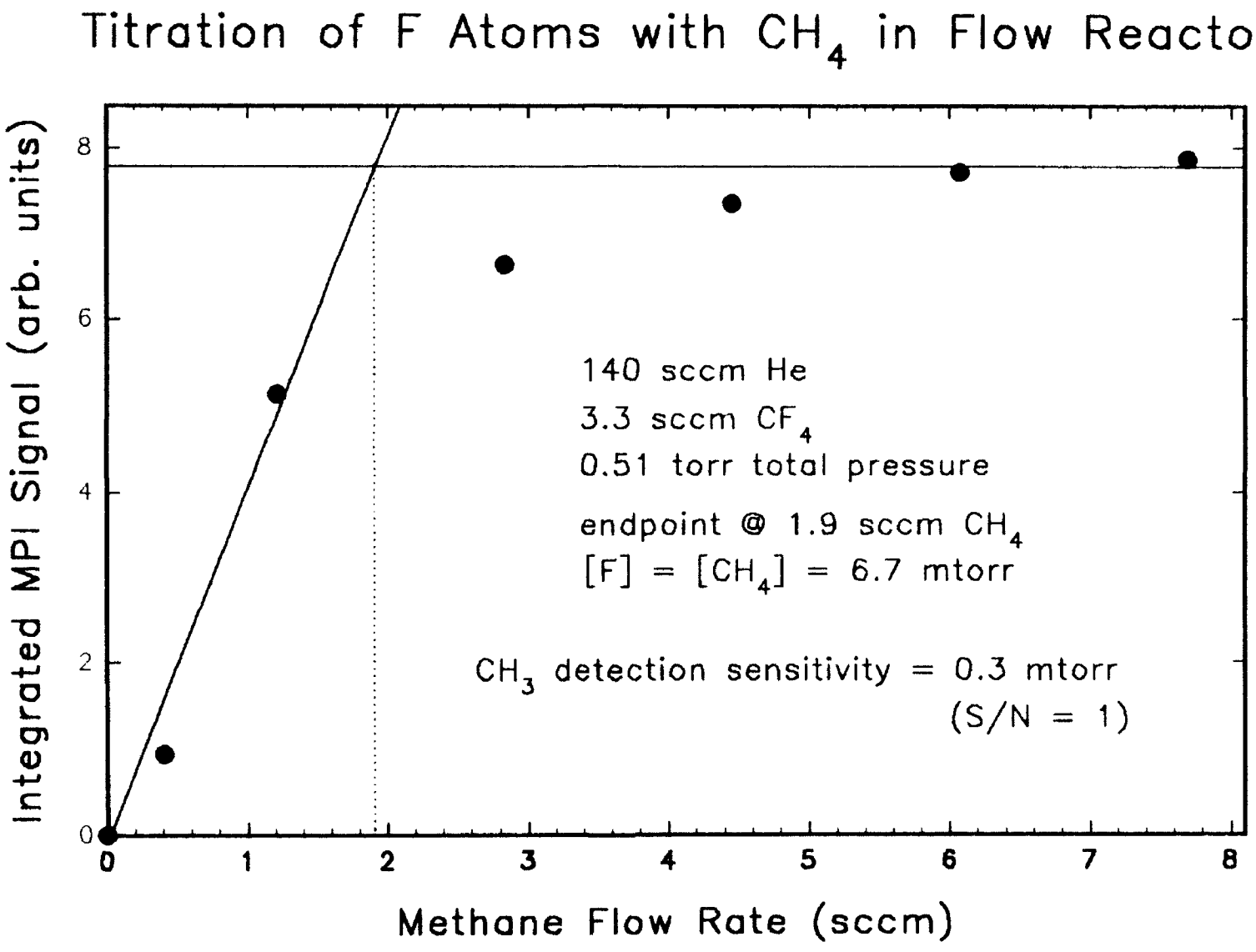
Fig. 2. Electrical Schematic for the Filament-Assisted CVD Reactor.

RESULTS and DISCUSSION

The methyl REMPI signal observed in the flow tube reaction, $F\cdot + CH_4 \rightarrow HF + CH_3\cdot$, is presented in Fig. 3. Flow conditions for obtaining this signal were: 140 sccm He, 3.3 sccm CF_4, 7.7 sccm CH_4, and 0.51 torr total pressure. The peak is centered near 333.5 nm[3] and has a FWHM bandwidth of 0.07nm (6 cm^{-1}) at 300K. Individual rotational lines were not observed in this band. Taking the integrated intensity of this band to be proportional to $[CH_3\cdot]$, F· atoms were titrated with methane in order to determine the detection sensitivity for methyl in this reactor. Increasing the methane concentration resulted in an increase in the REMPI signal intensity up to a plateau level. Intersecting lines were used to determine the endpoint of the titration, which was found to be equivalent to a pressure of 6.7 mtorr CH_4. This procedure is illustrated in the titration results, which are presented in Fig. 4. Since the reaction stoichiometry is one-to-one from methane to methyl, the REMPI signals in the plateau region of the titration curve are indicative of a methyl radical pressure of 6.7 mtorr. The signal of Fig. 3 was obtained as the last point (7.7 sccm CH_4 flow) in the titration curve of Fig. 4. Taking the signal-to-noise (S/N) ratio to be 20:1 for the REMPI signal, the $CH_3\cdot$ detection sensitivity was calculated to be 0.3 mtorr for S/N=1 in this apparatus. This value corresponds to a methyl radical number density of 1 x 10^{13} per cm^3 at 300K. The titration experiments were repeated two more times with similar results. The detection limit found in the flow reactor was considered sufficient to attempt the REMPI probe in a CVD reactor.

The primary difficulty in developing the methyl probe in the CVD reactor has proven to be the successful preparation and operation of a stable filament at temperatures of 2000 to 2300 °C. Near the end of the summer term it appeared as if rhenium would be the most stable

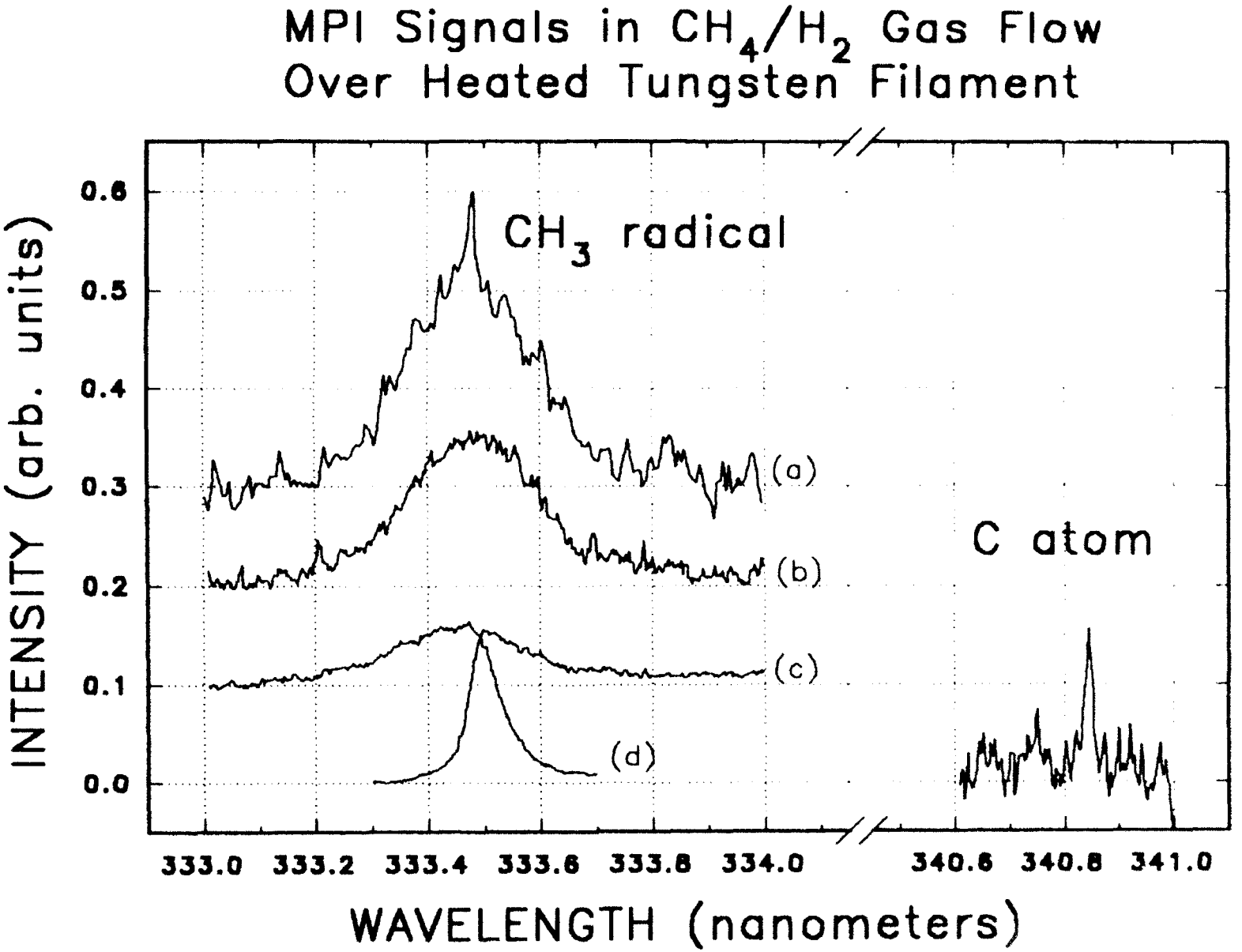

CH_3 MPI Signal from F + CH_4
INTENSITY (arb. units)
0.15
0.10
0.05
0.00
333.4
333.8
WAVELENGTH (nanometers)


Titration of F Atoms with CH_4 in Flow Reactor
Integrated MPI Signal (arb. units)
Methane Flow Rate (sccm)
0
2
4
6
8
0
1
2
3
4
5
6
7
8
140 sccm He
3.3 sccm CF_4
0.51 torr total pressure
endpoint @ 1.9 sccm CH_4
[F] = [CH_4] = 6.7 mtorr
CH_3 detection sensitivity = 0.3 mtorr
(S/N = 1)

filament, and work will continue in that area. Prior to that time we experienced moderate success in obtaining REMPI signals from methyl generated near tungsten wire filaments, but none of these filaments proved to be stable enough for a complete set of spatial measurements in the filament-to-substrate gap.

The REMPI signals obtained with a tungsten filament are presented in Fig. 5. The methyl signals in traces (a) through (c) were obtained with the same 0.010" dia tungsten (G.E. "218") filament. The filament was wound ten turns on a ~3 mm dia. mandrel. The individual coils expanded to ~5 mm dia. after winding. Approximately one inch of free wire was left at either end for winding around the Kovar rods. The wire was held to the rods with an alligator clip, and the axis of the filament windings was oriented vertically (perpendicular to the direction of the laser traversal). The filament was heated slowly in a 200 sccm gas flow of 2% CH_4/H_2 and maintained at a temperature of 2200-2300 °C. The filament power supply, was operated in current-limiting mode, supplying 8.0 amps at 33.6 volts to generate this temperature. The filament-to-substrate distance was ~40 mm, and the laser focal point was fixed at the center of this gap while the REMPI signals of Fig. 5. were obtained. Under these conditions an initial filament discharge current of 19.3 mA was generated with a filament-to-substrate potential of 200 volts. The 8.0 amp filament current and 200 volt discharge potential were fixed for the scans (except for trace (d)) presented in Fig. 5. The dye laser power was at 6 mJ/pulse in the 333.5 nm region and 11 mJ/pulse near 341 nm.

The first data obtained under the filament conditions just described was obtained by scanning the laser from 333.7 nm to 333.3 nm, a range which had proven to be sufficient for room temperature observations. This first scan (not shown) revealed the increased width of the


MPI Signals in CH_4/H_2 Gas Flow
Over Heated Tungsten Filament
CH_3 radical
C atom
(a)
(b)
(c)
(d)
INTENSITY (arb. units)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
333.0
333.2
333.4
333.6
333.8
334.0
340.6
340.8
341.0
WAVELENGTH (nanometers)

REMPI signal at higher temperature. Trace (a) was obtained immediately thereafter, with a laser repetition rate of 10 Hz and a 30 sample average of the gated integrator output. As time passed the filament moved away from the substrate, and the discharge current decreased. Trace (b) was obtained approximately 12 minutes later with a laser repetition rate of 30 Hz, a 100 sample average of the gated integrator output, and a discharge current of 18.5 mA. The S/N ratio is improved, but the absolute signal intensity had decreased. Attention was turned briefly to scan the region, 340.6 nm to 341.0 nm, where a carbon atom REMPI signal has been reported[2]. The discharge current dropped further to 16.7 mA. We obtained a weak signal for carbon atoms, also with a 30 Hz repetition rate and 100 sample average. This trace has been multiplied by a factor of three for display. Trace (c) was rescanned when the discharge current had dropped to 15.2 mA. The signal has decreased to almost a third of the intensity of trace (b).

The methyl REMPI signals in the CVD reactor may be compared with trace (d), the room temperature methyl signal of Fig. 3. The CVD REMPI signals are noticeably broader than the 300K signals due to a broader Boltzmann population of rotational levels at the hotter temperature. The width of our methyl signal is consistent with that obtained by Celii and Butler in a similar study[2]. The CVD REMPI signal intensities cannot be directly compared with trace (d) to estimate the methyl concentration because a different gated integrator was used; however, calibration of the signals is possible by titration in the CVD reactor.

Subsequent work with spatial resolution has indicated that the intensity changes due to filament instability currently prevents a spatially dependent methyl probe by scanning the methyl peak. We are hopeful that we may be able to obtain a spatial methyl profile by fixing the laser wavelength first at the peak and more quickly scanning the focal point between the filament and

the substrate. This scan would be followed by one in which the laser wavelength was tuned off resonance to obtain a baseline signal level. The intensity differences would then be related to the spatially-resolved methyl number density. Other attempts will continue to correct the filament instabilities.

CONCLUSIONS

The summer project demonstrated that spatially-resolved methyl detection is possible in the CVD reactor. Titration results in a flow tube reactor suggests that the detection sensitivity may be of the order of 10^{13} cm^{-3}. The REMPI signal in the CVD reactor was found to be very sensitive to the movement of the filament and/or the discharge current. The largest impediment to progress has proven to be the preparation and operation of a stable filament. Success in this area will require more experience and knowledge to learn how filament preparation and conditioning affects the operational stability of the filament.

As the work progresses it is expected that a thermocouple will be placed on the substrate and that other substrates such as silicon may be mounted on the molybdenum disc. Goals for the continued work on this project include the determination of the spatial CH_3· density as function of: (1) methane concentration, (2) total pressure, (3) filament material and temperature, (4) substrate material and temperature, (5) discharge current, and (6) filament-to-substrate distance.

ACKNOWLEDGEMENTS

The assistance of Michelle Klingshirn in the design, execution, and analysis of these experiments is gratefully acknowledged. Ideas and assistance from Dr. Peter Bletzinger were invaluable as were conversations with Dr. Bish Ganguly and Dennis Grossjean regarding high temperature filaments. The technical expertise and cheerful assistance of Mike Ray and Bob Knight made this work possible.

REFERENCES

1. Two reviews of the diamond CVD processes are: P. K. Bachmann; R. Messier *Chem. Eng. News*, pp 24-38, May 15, 1989. and F. G. Celii; J. E. Butler *Ann. Rev. Phys. Chem.* **(1991)**, *42*, 643-84.

2. F. G. Celii; J. E. Butler *J. Appl. Phys.* **(1992)**, *71*, 2877-83.

3. J. W. Hudgens; T. G. DiGuiseppe; M. C. Lin *J. Chem. Phys.* **(1983)**, *79*, 571.

THIS PAGE INTENTIONALLY LEFT BLANK

PERFORMANCE ANALYSIS OF A HETERODYNE LIDAR SYSTEM INCORPORATING A MULTIMODE OPTICAL WAVEGUIDE RECEIVER

Bradley D. Duncan, Ph.D.
Assistant Professor

University of Dayton
Department of Electrical Engineering
Center for Electro-Optics
300 College Park
Dayton, OH 45469

Final Report For:
AFOSR Summer Faculty Research Program
Wright Laboratory
(WL/AARI-2)

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 14, 1992

PERFORMANCE ANALYSIS OF A HETERODYNE LIDAR SYSTEM INCORPORATING A MULTIMODE OPTICAL WAVEGUIDE RECEIVER

Bradley D. Duncan, Ph.D.
Assistant Professor

Department of Electrical Engineering
Center for Electro-Optics
University of Dayton

ABSTRACT

A theoretical/numerical performance analysis of a heterodyne lidar system incorporating a multimode optical waveguide receiver has been performed. For our purposes, the performance parameters of interest are the coupling and mixing efficiencies of the lidar receiver, as they relate to the overall system carrier to noise ratio (CNR). The coupling efficiency relates the ability of the lidar receiver optics to couple light returning from a distant object into the multimode optical waveguide receiver. The energy thus received is then mixed with a local oscillator (LO) field in an evanescent wave coupler to generate the desired intermediate frequency (IF) signal current. The efficiency with which this mixing occurs, however, is very strongly influenced by the receiver waveguide geometry, and thus the constituent modal structures of the guided LO and signal beam fields. Initial results indicate that for a symmetric square waveguide supporting seventy-five distinct modes, the coupling and mixing efficiencies can approach 78% and 23%, respectively. These results rival the performance characteristics of many free space and single mode waveguide lidar receiver designs, while higher efficiencies are expected for more realistic multimode waveguides supporting many more modes, and for lidar receivers constructed from circular (possibly graded index) multimode optical fiber waveguides.

PERFORMANCE ANALYSIS OF A HETERODYNE LIDAR SYSTEM INCORPORATING A MULTIMODE OPTICAL WAVEGUIDE RECEIVER

Bradley D. Duncan, Ph.D.

I. INTRODUCTION

Heterodyne lidar systems invoke the coherent detection concept in which a laser echo, scattered from a distant target, is mixed with a frequency shifted reference laser radiation field called the local oscillator (LO). The mixing of these two coherent laser emissions then induces an intermediate frequency (IF) signal current in an optical detector which carries the target information at a frequency equal to the frequency difference between the signal (echo) and LO radiations. Conventionally, the mixing of the LO and the signal fields takes place in free space. For such a mixing scheme, though, not all the optical power collected by the receiving aperture contributes to the electrical power in the intermediate frequency signal. This results because only that portion of the returning echo signal which is in the same temporal and spatial mode of the LO will contribute to the IF signal [1].

Efforts to increase the efficiency by which the optical to electrical power conversion takes place have thus far included LO beam truncation methods [2] and the incorporation of single mode fiber optics into the LO and return signal paths [3]. In a typical free space lidar mixer design, a broadened Gaussian LO beam is mixed with a randomly distributed laser echo field, assuming the worst case of a purely speckle type target. For such a scheme, only about a twenty-five percent mixing efficiency can be achieved between the LO and signal fields [2]. That is, only about twenty-five percent of the optical return power collected by the receiving optics contributes to the desired IF signal. By truncating (or apodizing) the LO field, however, higher mixing efficiencies can be achieved by causing the LO field to diffract in such a way as to cause a stronger spatial matching between the LO and signal fields. In fact, mixing efficiencies approaching forty-four percent can be achieved for properly truncated LO fields [2].

By contrast, mixing efficiencies approaching 100 percent can be achieved in heterodyne detection systems employing a single mode fiber mixer. By causing ideal spatial matching between the LO and signal fields, all the optical signal power propagating in the fiber will contribute to the IF signal. However, even if the numerical aperture (NA) of the lidar receiver optics is perfectly matched to the NA of the fibers used in the single mode mixer, not all the power collected by the receiving optics can be

coupled into the fiber's single propagating mode. Due to inherently low coupling efficiencies then, it has been shown that under optimum conditions the single mode fiber mixer will perform no better than the free space mixer [3].

In an attempt to increase both the mixing and coupling efficiencies simultaneously in a heterodyne lidar system we have proposed a multimode optical waveguide mixing arrangement. By matching the NA of the receiving optics to that of the waveguide employed in the mixer, it can be shown that nearly all the optical power passing through the receiver aperture can be coupled into the waveguide, provided that the waveguide supports sufficiently many modes [4]; for mixers employing "few mode" waveguides, the coupling efficiency will be somewhat less than 100%, as will be shown later. Furthermore, the mixing efficiency of multimode receivers can be increased since the LO is not confined to a single spatial mode, thereby allowing a stronger correlation between the LO and signal fields propagating in their respective waveguides. As we will see, though, the efficiency with which this mixing occurs is very strongly influenced by the receiver waveguide geometry, and thus the constituent modal structures of the LO and signal fields.

To begin our development, we shall derive the carrier-to-noise ratio (CNR) relationship for a monostatic heterodyne lidar system employing a general multimode optical waveguide receiver. From this development we will see more directly the influence the coupling and mixing efficiencies have on lidar performance. We will then introduce the essential characteristics of the lidar system we assumed for simulation purposes. The mixing and coupling efficiency simulation results themselves will then be presented and discussed, after which we will conclude with a discussion of our systems' advantages (and possible improvements), and the likely directions of future work.

II. GENERAL CARRIER-TO-NOISE RATIO ANALYSIS

A simplified block diagram of the monostatic (i.e., common transmitter/receiver axis) heterodyne lidar system we wish to investigate is shown in Figure 1. We see that a single laser source is used to generate both the transmitter and LO signals, and as we are considering a monostatic system, a common aperture is used for both the transmitted and received beams. Lens ℓ_1, shown schematically in the plane of the monostatic aperture, is used to provide a collimated transmit beam for target illumination. Lens ℓ_2, is then chosen so that, in conjunction with lens ℓ_1, a receiver lensing arrangement is created with a numerical aperture (NA) which matches the NA of the waveguide used in the multimode mixer design.

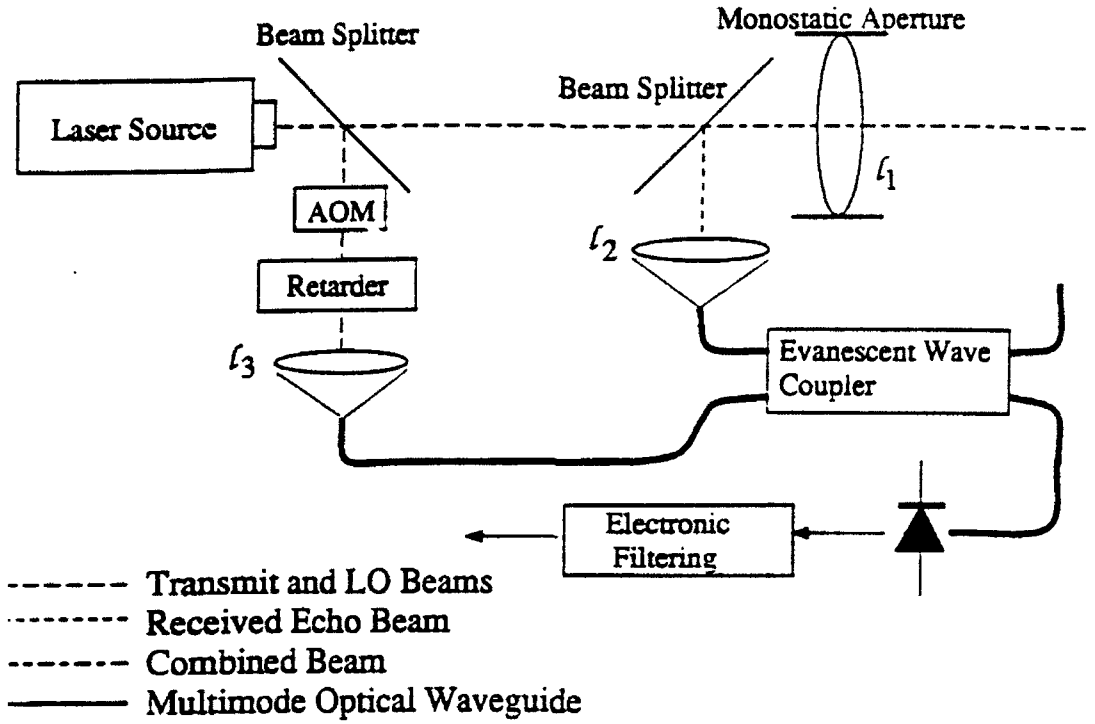


FIGURE 1
Heterodyne Lidar Block Diagram

This is done in order to assure that optical power returning from a distant target is optimally coupled into the receiveing arm of the waveguide mixer [4]. The third lens ℓ_3 is similarly chosen for optimum coupling of LO power into the LO arm of the waveguide mixer, while the indicated acousto-optic modulator (AOM) is used to upshift the LO frequency by an amount ω_{IF}, the desired intermediate frequency. The variable retarder we have indicated is simply used to ensure a zero initial relative phase state between the LO and signal beams as they are focused into their respective waveguide mixer arms. The LO and receiver fields are then mixed in an evanescent wave coupler, with the resultant field being incident upon a photodetector. Through electronic filtering then, the IF signal current is extracted for further signal processing and analysis.

In this section we will develop a relationship for the carrier-to-noise ratio (i.e., we assume no refractive atmospheric turbulence between the lidar system and its target) at the output of the electronic filter of Figure 1. For this analysis we will base our "signal" calculations only on the average optical return power P_r passing through the monostatic aperture. We will also assume that the LO level is adjusted so that the CNR is LO shot noise limited. Furthermore, we assume that the target of interest is an infinite (i.e., resolved) planar speckle target normal to the transmit/receive axis. With these system constraints identified we now model the received signal and LO fields in their respective arms of the multimode

waveguide mixer as,

$$\underline{\tilde{U}}_{fi}(\overline{\rho},t) = \sum_{\ell=1}^{M} \underline{\tilde{b}}_{\ell}\,\xi_{\ell}(\overline{\rho})\,e^{j\omega t} \quad \textit{and} \quad \underline{U}_{LO}(\overline{\rho},t) = \sum_{k=1}^{M} \underline{a}_{k}\,\xi_{k}(\overline{\rho})\,e^{j(\omega+\omega_{IF})t} \quad , \tag{1}$$

respectively, where $\{\xi_{k,\ell}(\overline{\rho}) \mid k,\ell = 1,2,\ldots,M\}$ is the allowed set of orthonormal guided wave mode solutions, $\overline{\rho}$ is a 2-D spatial frequency variable, ω is the laser source frequency and ω_{IF} is the intermediate frequency. Also, to clarify our notation, we mention that an underscore is used here to indicate a complex quantity, an overscore is used to indicate a vector and an "overtilde" indicates a random quantity. The LO and received powers coupled into their respective arms of the waveguide mixer can then be written as,

$$P_{LO} = \sum_{k=1}^{M} |\underline{a}_{k}|^2 \quad \textit{and} \quad \tilde{P}_{fi} = \sum_{\ell=1}^{M} |\underline{\tilde{b}}_{\ell}|^2 \quad , \tag{2}$$

respectively, where the "a" and "b" coefficients are identified as complex mode weighting factors. Note that only the b's are indicated as both complex and random. This is due to the random nature of the light scattered by the assumed speckle target.

After filtering, the IF signal portion $\underline{\tilde{y}}(t)$ of the photodetector output is

$$\underline{\tilde{y}}(t) = \Re \int_{-\infty}^{\infty}\int \underline{\tilde{U}}^{*}_{fi}(\overline{\rho},t)\,\underline{U}_{LO}(\overline{\rho},t)\,d\overline{\rho} \quad , \tag{3}$$

where $\Re$ is the photodetector responsivity. Inserting equations (1) into equation (3) and performing the required simplifications then yields

$$\underline{\tilde{y}}(t) = \Re e^{j\omega_{IF}t} \sum_{k=1}^{M} \underline{a}_{k}\underline{\tilde{b}}^{*}_{k} \quad . \tag{4}$$

Note now, that, for the purposes of this analysis, we shall assume that the guided propagation paths which both the LO and signal fields travel are relatively short (say, a meter or so) so that intermodal phase dispersion effects can be neglected. This, combined with the additional phase equalization effects of the variable retarder shown in Figure 1, indicates that the complex nature of the photodetector output current $\underline{\tilde{y}}(t)$ can only be due to intramodal phase variations. For most commonly interesting waveguide geometries,

however, modal solutions manifest themselves as non-uniform plane waves (i.e., varying amplitude and uniform phase in a plane transverse to the direction of propagation). If we assume a sufficiently narrow line width source, then, we see that intramodal phase effects are negligible as well; thus causing the coefficients $\underline{a}_k$ and $\underline{b}_k$, as well as the signal current $\underline{y}(t)$ to be strictly real.

To continue our analysis, the signal power P_s is then written as the time averaged expectation of the IF signal current, squared. That is,

$$\begin{aligned} P_s &= \left\langle E\left[|\underline{y}(t)|^2 \right] \right\rangle \\ &= \Re^2 E\left[\sum_k \underline{a}_k \underline{b}_k^* \cdot \sum_\ell \underline{a}_\ell^* \underline{b}_\ell \right] \\ &= \Re^2 \sum_k \sum_\ell \underline{a}_\ell^* E[\underline{b}_\ell \underline{b}_k^*] \underline{a}_k \end{aligned} \quad . \tag{5}$$

Simplifying our notation we write

$$P_s = \Re^2 \underline{\overline{a}}^\dagger \underline{\overline{B}} \underline{\overline{a}} \quad , \tag{6}$$

where $\underline{\overline{a}}$ is a complex vector containing the elements a_k, $\underline{\overline{a}}^\dagger$ is the complex transpose of $\underline{\overline{a}}$ and $\underline{\overline{B}}$ is a complex Hermetian matrix (i.e., $\underline{\overline{B}} = \underline{\overline{B}}^\dagger$) containing the elements $B_{k\ell} = E[\underline{b}_\ell \underline{b}_k^*]$. Note that the Hermetian nature of $\underline{\overline{B}}$ indicates that $B_{k\ell} = B_{\ell k}^*$. This is a very important result which leads to a considerable reduction in the computational complexity of the numerical simulations which we shall present later.

As mentioned previously, to determine the detector noise power we assume the LO level is adjusted so that LO shot noise dominates. The detector noise P_n is then written as

$$\begin{aligned} P_n &= q\Re B_n \int_{-\infty}^{\infty}\int \left| \underline{U}_{LO}(\overline{\rho}, t) \right|^2 d\overline{\rho} \\ &= q\Re B_n \sum_k |\underline{a}_k|^2 \\ &= q\Re B_n P_{LO} \end{aligned} \quad , \tag{7}$$

where B_n is the post detection bandwidth centered at ω_{IF} and q is the basic electronic charge (i.e., 1.6×10^{-19} C). The CNR for our lidar system is then found by taking the ratio of equations (6) and

(7) to yield

$$CNR = \frac{P_s}{P_n} = \frac{\Re \underline{\bar{a}}^{\dagger} \underline{\bar{B}} \underline{\bar{a}}}{qB_n P_{LO}} \quad . \tag{8}$$

For the purpose of this analysis, we now assume that the LO waveguide field is launched such that there is zero initial dispersion between the constituent LO field modes - thus allowing us to take the "a" coefficients to be real. We furthermore assume that the LO modes are equally excited such that $\left\{ a_i = \sqrt{(P_{LO}/M)} \mid i=1,2,\ldots,M \right\}$, where M is the total number of allowed guided modes. This is a reasonable assumption, provided we consider the LO mode distribution sufficiently far from the source. With these assumptions we then write the CNR in "standard" form [3] as

$$CNR = \frac{\Re P_r}{qB_n} \left\{ \frac{1}{MP_r} \sum_k \sum_l B_{kl} \right\} \quad , \tag{9}$$

where the quantity in braces is recognized as the overall efficiency η_{mm} of the lidar system, which, in the case of the multimode mixer, incorporates both the coupling and mixing efficiencies. More specifically, using equation (2) the coupling efficiency η_{couple} is written as

$$\eta_{couple} = \frac{E\left[\tilde{P}_{fi}\right]}{P_r} = \frac{1}{P_r} \sum_k E\left[|\tilde{b}_k|^2 \right] = \frac{1}{P_r} \sum_k B_{kk} \quad , \tag{10}$$

where the summation is simply the trace of the $\underline{\bar{B}}$ matrix. The mixing efficiency η_{mix} is then found by dividing η_{mm} by η_{couple}. That is,

$$\eta_{mix} = \frac{\eta_{mm}}{\eta_{couple}} = \frac{1}{M} \frac{\sum_k \sum_l B_{kl}}{\sum_k B_{kk}} \quad . \tag{11}$$

The focus, now, for the rest of this paper will be to determine these efficiency quantities for a proposed multimode waveguide mixer. Before continuing, however, two important comments/observations must be made.

First, in equation (10) it is assumed that the beam splitter in the path of the received target signal (i.e., the beam splitter shown nearest the monostatic aperture in Figure 1) does not play a role in reducing the amount of collected signal power, P_r, being coupled into the waveguide mixer. In practice this assumption is addressed by judiciously incorporating a polarized source, polarizing beam splitters and optical rotators into the lidar design. One may then, in essence, consider the coupling to be a quantification of the multimode waveguide mixer's ability to "sample" the many random spatial signal modes collected by the monostatic aperture. Secondly, even though the guided modes of a multimode waveguide are indeed orthogonal, the mechanisms by which each mode is excited are not necessarily independent. This result in essence arises in a way similar to the one by which the correlation time of a time domain signal is increased by low pass filtering [5]. In our case, though, the signal of interest is the random back scatter form the speckle target, and the low pass spatial filtering is performed by the finite NA receiving optics. The quantities $B_{k\ell} = E[\underline{\tilde{b}}_\ell \underline{\tilde{b}}_k^*]$ thus cannot be written in a simpler form as the product of the expected values of $\underline{\tilde{b}}_\ell$ and $\underline{\tilde{b}}_k^*$, separately. As a result, the determination of the actual $B_{k\ell}$ values, for a given waveguide, involves some rather laborious, though straight forward, theoretical manipulations. An abbreviated account of the required analysis will be provided in the following section.

III. MULTIMODE WAVEGUIDE MIXING EFFICIENCY: A CLOSER LOOK

In this section we will look more closely at the overall efficiency η_{mm} in terms of the transmitted beam field, the target's random complex reflectivity and the geometry of the monostatic transmit/receive aperture. We begin by stating that Snyder [6] has shown that it is possible to approximate the coupling of a free space field and a waveguide mode by an overlay integral. Using this formalism we may then write each of the constants $\underline{\tilde{b}}_i$ as

$$\underline{\tilde{b}}_i = \int_{-\infty}^{\infty}\int \underline{\tilde{U}}_f(\bar{\rho})\,\xi_i^*(\bar{\rho})\,d\bar{\rho} \quad , \tag{12}$$

where $\underline{\tilde{U}}_f$ is the field focused on the end face of the optical waveguide and where the 2-D coordinate

system $\overline{\rho}$ is taken in the focal plane and is normal to the waveguide propagation axis. The matrix elements $B_{k\ell}$ are then written as

$$B_{k\ell} = E[\underline{\tilde{b}}_k \underline{\tilde{b}}_\ell^*] = \int_{\overline{\rho}} \int_{\overline{\rho}'} E\left[\underline{\tilde{U}}_f(\overline{\rho})\underline{\tilde{U}}_f^*(\overline{\rho}')\right] \tilde{\xi}_k^*(\overline{\rho})\tilde{\xi}_\ell(\overline{\rho}')\, d\overline{\rho}\, d\overline{\rho}' \quad . \tag{13}$$

To now represent the received field $\underline{\tilde{U}}_f$ in terms of the target and system characteristics, we propagate the transmitted field U_T to the target, multiply by the target's complex random reflectivity $\underline{\tilde{T}}$, and then back propagate the resultant field through the receiving aperture to the receiving optics' focal plane. We, of course, assume the waveguide end face to be positioned at the focal point of the receiver optics. The mathematical details of this analysis can be found in reference [3]; the result, however, is

$$\underline{\tilde{U}}_f(\overline{\rho}) = \frac{1}{(j\lambda f)(\lambda L)^2} \int_{A_T} \int \underline{\tilde{T}}(\overline{\rho}_T)\, \hat{W}\left(\frac{\overline{\rho}}{\lambda f} - \frac{\overline{\rho}_T}{\lambda L}\right) \hat{U}_T\left(\frac{\overline{\rho}_T}{\lambda L}\right) d\overline{\rho}_T \quad , \tag{14}$$

where L is the distance to the target (assumed to be large enough that the target is in the Fraunhoffer zone of the lidar transmit/receive optics), f is the focal length of the receiving optics, $\overline{\rho}_T = (x_T, y_T)$ is the coordinate system at the target, A_T is the target extent (area), and $\hat{W}$ and $\hat{U}_T$ represent, respectively, the Fourier transforms of the receiver aperture function W and the transmitted beam field U_T. By careful examination we see that equation (14) simply represents the image of the target in the focal plane of the receiving optics, with illumination given by the far field pattern of the transmitted beam, and with the image blurred by the Fourier transform of the receiving aperture.

We now insert equation (14) into equation (13) and perform the required simplifications to obtain, for an assumed resolved target,

$$B_{k\ell} = \frac{\tau_s}{(\lambda L)^4 f^2} \int_{-\infty}^{\infty} \int d\overline{\rho}_T \left| \hat{U}_T\left(\frac{\overline{\rho}_T}{\lambda L}\right)\right|^2 \cdot \int_{-\infty}^{\infty} \int d\overline{\rho}\, \hat{W}\left(\frac{\overline{\rho}}{\lambda f} - \frac{\overline{\rho}_T}{\lambda L}\right) \xi_k(\overline{\rho}) \cdot \int_{-\infty}^{\infty} \int d\overline{\rho}'\, \hat{W}^*\left(\frac{\overline{\rho}'}{\lambda f} - \frac{\overline{\rho}_T}{\lambda L}\right) \xi_\ell^*(\overline{\rho}') \quad , \tag{15}$$

where τ_s is the average target reflectivity and the intermediate result

$$E[\underline{\tilde{I}}(\overline{\rho}_T)\underline{\tilde{I}}^*(\overline{\rho}_T)] = \lambda^2 \tau_s \delta(\overline{\rho}_T - \overline{\rho}_T) \tag{16}$$

has been employed [7]. By defining the transmitted field, the geometry of the geometry of the transmit/receive aperture and the allowed waveguide mode set we are now in a position to determine (primarily through numerical simulations) the matrix elements B_{kl} and thus the mixing and coupling efficiencies for the multimode waveguide mixer. We will look first at the monostatic aperture geometry and the transmitted beam field, while deferring a detailed discussion of the waveguide and its allowed guided mode set till the next section.

For simplicity, in section IV we will present a square optical waveguide mixer geometry. We will thus also assume that the monostatic aperture is square with the aperture function W being written as

$$W(\overline{\rho}_a) = W(x_a, y_a) = rect\left(\frac{x_a}{a}, \frac{y_a}{a}\right) = \begin{cases} 1 \; ; \left(-\frac{a}{2} \le x_a \le \frac{a}{2}\right) \; and \; \left(-\frac{a}{2} \le y_a \le \frac{a}{2}\right) \\ 0 \; ; \; elsewhere \end{cases} , \tag{17}$$

where $\overline{\rho}_a = (x_a, y_a)$ is a 2-D variable in the aperture plane and "a" is the width of the aperture in both the x_a and y_a dimensions. The Fourier transform $\hat{W}$ of this aperture function, in the focal plane of the receiving optics, is then written as

$$\hat{W}\left(\frac{\overline{\rho}}{\lambda f}\right) = \mathcal{F}\{W(\overline{\rho}_a)\}\Big|_{\frac{\overline{\rho}}{\lambda f}} = a^2 Sa\left(\pi a \frac{\overline{\rho}}{\lambda f}\right) = a^2 Sa\left(\pi a \frac{x}{\lambda f}\right) Sa\left(\pi a \frac{y}{\lambda f}\right) , \tag{18}$$

where $\overline{\rho}/\lambda f$ is the 2-D spatial frequency variable, $\overline{\rho} = (x, y)$ is, again, the coordinate system in the focal plane of the receiving optics, f is the common focal length of the receiving optics in both the x and y dimensions, $\mathcal{F}$ indicates the Fourier transform operation and Sa indicates the separable sine-over-argument function [5]. Note that equation (18) represents a strictly real even function. This indicates that $\hat{W} = \hat{W}^*$ and that the last two integrals in equation (15) may be interpreted as convolutions - a result which will also reduce the computational complexity of our numerical simulations, the results of which will be presented later.

The transmitted beam U_T, in the aperture plane, is similarly taken to be a collimated Gaussian plane wave apodized by the square aperture function. We will also assume that the transmitted beam is

normalized is such a way that

$$\int_{-\infty}^{\infty}\int \left|U_T(\bar{\rho}_a)\right|^2 d\bar{\rho}_a = P_T \quad , \tag{19}$$

where P_T is the transmitted power. Without proof, then, the transmitted beam field is described by

$$U_T(\bar{\rho}_a) = U_T(x_a, y_a) = \frac{\sqrt{P_T}}{erf(R/\sqrt{8})}\, rect\left(\frac{x_a}{a}, \frac{y_a}{a}\right) \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x_a^2 + y_a^2)/4\sigma^2} \quad , \tag{20}$$

where σ is the rms width of the Gaussian intensity distribution incident on the square aperture, $R = a/\sigma$ is the transmitted beam truncation ratio and the error function (erf) arises due to the normalization described in equation (19). The Fourier transform $\hat{U}_T$ of the transmitted field, in the target plane, is thus

$$\hat{U}_T\left(\frac{\bar{\rho}_T}{\lambda L}\right) = \mathcal{F}\{U_T(\bar{\rho}_a)\}\Big|_{\frac{\bar{\rho}_T}{\lambda L}} = \frac{\sqrt{8\pi P_T}}{erf(R/\sqrt{8})}\, a^2\sigma\, Sa\left(\pi a \frac{x_T}{\lambda L}\,,\, \pi a \frac{y_T}{\lambda L}\right) \bullet e^{-\left(\frac{2\pi\sigma}{\lambda L}\right)^2 (x_T^2 + y_T^2)} \quad , \tag{21}$$

where $\bar{\rho}_T = (x_T, y_T)$ is the 2-D coordinate system in the target plane, $\bar{\rho}_T/\lambda L$ is a 2-D spatial frequency variable and $\bullet$ indicates the 2-D spatial frequency domain convolution of the Sa and exponential functions.

To end this section, we need to made one further comment, pertaining to the relationship between the transmitted power P_T and the average received power P_r. That is, for a resolved speckle target, we may conveniently write

$$P_r = P_T \tau_s \frac{a^2}{L^2} \quad . \tag{22}$$

Simply, this equation states that the average received power is equal to the transmitted power P_T times the average target reflectivity τ_s and the area of the receiving aperture a^2, divided then by the square of the distance to the target L^2. By now incorporating equation (22) into equation (21), equations (21) and (18) into equation (15), and equation (15) into equations (10) and (11) we see that the mixing and coupling efficiencies are independent of the average received power P_r. Furthermore, though not as obvious, it can be shown that the coupling and mixing efficiencies are also independent of the distance

L, as long as L is large enough to place the target in the Fraunhoffer region of the transmit/receive optics.

IV. SQUARE WAVEGUIDE MODE SOLUTIONS

As an extensive analysis of the guided modes allowed in a square optical waveguide structure would be quite lengthy, we will present here only the results necessary to perform the desired simulations of the coupling and mixing efficiencies for a multimode mixer constructed from square optical waveguides. A more detailed analysis of the square optical waveguide structure can be found in reference [8]. We have chosen to base our simulations on a square waveguide structure simply because the modal solutions are relatively simple and are based on the superposition of sine and cosine functions, rather than the Bessel-type transcendental functions encountered in the more common circular optical waveguide structures. As we will see, though, quite a bit can be learned from the analysis of the square waveguide mixer.

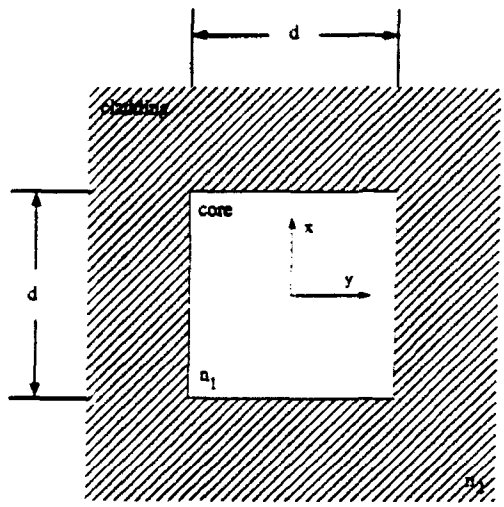


FIGURE 2
SQUARE WAVEGUIDE GEOMETRY

The symmetric square waveguide structure we have chosen to model is illustrated above in Figure 2. We see that the core material, of index n_1, is surrounded on all sides by a cladding material of index n_2. The dimension d is simply the width of one of the core sides. In addition to considering a symmetric waveguide stricture, we will in all cases assume that our waveguide is weakly guiding. That is, the indices of refraction n_1 and n_2 are very similar, such that the index difference parameter

$\Delta = (n_1 - n_2)/n_1$ is very much less than unity. We furthermore assume that all modes are linearly polarized normal to one of the core faces, say in the x direction. We make this assumption, first, because in practice we will use a linearly polarized source and, second, because due to the degenerate, though orthogonal (i.e., no IF signal current is produced by the interaction of orthogonally polarized modes), nature of oppositely polarized square waveguide mode sets, all the information we require regarding the action of our waveguide mixer can be determined from the consideration of modes in only one linear polarization state.

The guided mode shapes $\xi(\bar{\rho})$, now, which are allowed in such a square waveguide structure are quite simple and can be shown to be the superpositions, in the core region, of sinusoidal amplitude variations in both the x and y directions. In the cladding there is simply an exponential modal amplitude decay [8]. More specifically, we may write $\xi(\bar{\rho}) = \xi_x(x)\,\xi_y(y)$ as a real function where

$$\xi_t(t) = \frac{1}{C_{eN}} \begin{cases} \cos(V\sqrt{1-b_e}\,/2) \cdot \exp\{-V\sqrt{b_e}\,(t - d/2)/2\} & ;\ (d/2) \le t \\ \cos(V\sqrt{1-b_e}\,t/2) & ;\ (-d/2) \le t \le (d/2) \\ \cos(V\sqrt{1-b_e}\,/2) \cdot \exp\{+V\sqrt{b_e}\,(t + d/2)/2\} & ;\ t \le (d/2) \end{cases} \tag{23}$$

or

$$\xi_t(t) = \frac{1}{C_{oN}} \begin{cases} \sin(V\sqrt{1-b_o}\,/2) \cdot \exp\{-V\sqrt{b_o}\,(t - d/2)/2\} & ;\ (d/2) \le t \\ \sin(V\sqrt{1-b_o}\,t/2) & ;\ (-d/2) \le t \le (d/2) \\ -\sin(V\sqrt{1-b_o}\,/2) \cdot \exp\{+V\sqrt{b_o}\,(t + d/2)/2\} & ;\ t \le (d/2) \end{cases} \tag{24}$$

with the coordinate origin taken to be at the center of the waveguide core and where t is a dummy variable representing either the x or y spatial coordinates. In addition, we see that the modal variations depend on the normalized frequency $V = k_o d n_1 \sqrt{2\Delta}$, where $k_o = 2\pi/\lambda$ is the free space wave number, and the normalized propagation constants b_e and b_o. (In the above equations the e and o subscripts indicate the even and odd variations of ξ_t, respectively.) The normalized frequency and propagation constants are in turn related by the following eigenvalue equations:

$$\tan\left\{\frac{V}{2}\sqrt{1 - b_e}\right\} = \sqrt{\frac{b_e}{1 - b_e}} \quad \textit{and} \quad \tan\left\{\frac{V}{2}\sqrt{1 - b_o}\right\} = -\sqrt{\frac{1 - b_o}{b_o}} \ . \tag{25}$$

Equations (25) are used to determine the multiple solutions for b_e and b_o, given V, for both the $\xi_x(x)$ and $\xi_y(y)$ solutions. (We mention that due to the symmetric, square nature of the waveguide we are studying, V is common to both the $\xi_x(x)$ and $\xi_y(y)$ modal variations.) The last parameters of interest in equations (23) and (24) are the amplitude normalization constants C_{eN} and C_{oN} for the even and odd modal solutions, respectively. These constants are chosen for normalized modal power such that

$$\iint_{\bar{\rho}} |\xi(\bar{\rho})|^2 d\bar{\rho} = 1 \quad . \tag{26}$$

TABLE 1
(Normalized Propagation and Amplitude Constants)

b	C_N	even/odd
$b_1 = b_{e1} = 0.9904$	0.5335	even
$b_2 = b_{o1} = 0.9615$	0.5339	odd
$b_3 = b_{e2} = 0.9134$	0.5349	even
$b_4 = b_{o2} = 0.8463$	0.5362	odd
$b_5 = b_{e3} = 0.7604$	0.5382	even
$b_6 = b_{o3} = 0.6560$	0.5412	odd
$b_7 = b_{e4} = 0.5336$	0.5456	even
$b_8 = b_{o4} = 0.3943$	0.5531	odd
$b_9 = b_{e5} = 0.2398$	0.5681	even
$b_{10} = b_{o5} = 0.0762$	0.6207	odd

For our work we have assumed a normalized operating frequency of V=30. This allows for 10 distinct mode shapes (i.e., 10 distinct b solutions) in both the x and y directions, five of which are even solutions and five of which are odd solutions. That is, there are five solutions apiece for the two eigenvalue equations given in equations (25). These solutions, along with the appropriate normalization constants C_N are given in the above table. We have also indicated whether the eigenvalue solution corresponds to an even or odd mode variation. Note that in Table 1 we have labeled the first even mode propagation constant as b_{e1}, the fourth odd mode propagation constant as b_{o4}, etc.; and for simplicity we have relabeled all the propagation constants, in descending magnitude order, as b_1 through b_{10}.

Now consider the composite mode designation $\xi^{x}_{pq}(\bar{\rho})$. As previously mentioned, the square waveguide mode solutions are a superposition of the x and y spatial variations given in equations (23) and (24). The mode designation $\xi^{x}_{pq}(\bar{\rho})$ then is meant to indicate that b_p and b_q, where $\{p,q = 1,2,3,\ldots,10\}$, are utilized in the choice of the $\xi_x(x)$ and $\xi_y(y)$ modal variations, respectively. In addition, the superscript x is used to remind us that we are considering only modal solutions linearly polarized in the x-direction. For example, $\xi^{x}_{25}(\bar{\rho})$ represents a square waveguide mode for which the x-variation $\xi_x(x)$ corresponds to that given in equation (24) when we set $b_o = b_{o1} = b_2$, and for which the y-variation $\xi_y(y)$ corresponds to that given in equation (23) when we set $b_e = b_{e3} = b_5$. Furthermore, it can be shown that the subscripts p and q correspond physically to the number of mode intensity maxima in the x and y directions, respectively. It would seem then that for V=30, we would have 100 possible guided mode solutions for our square waveguide. In fact, though, only seventy-five mode solutions are allowed. This results because of an additional constraint which says that for square waveguide modes to exist, we must choose b_p and b_q such that $b_p + b_q \geq 1$. Thus $\xi^{x}_{27}(\bar{\rho})$ represents a valid guided mode solution for V=30, while $\xi^{x}_{97}(\bar{\rho})$ does not. It is thus the seventy-five allowed guided mode solutions which we have used in equation (15) for the determination of the matrix elements $B_{k\ell}$, and then the coupling and mixing efficiencies.

V. NUMERICAL SIMULATIONS AND RESULTS

After back substituting the appropriate terms into equations (10) and (11) and performing numerous simplifications, the coupling and mixing efficiencies were calculated for the square waveguide mixer, as a function of the truncation ratio R, based on the following system parameters:

V = 30	NA = 0.211
n_1 = 1.5	λ = 1.06 μm
Δ = 0.01	a = 10 cm
n_2 = 1.485	f = 23.62 cm
d = 23.918 μm	M = 75

(Recall that the numerical aperture of the optical waveguide NA_{wg} is given by $NA = n_1\sqrt{2\Delta}$, while

the numerical aperture of the receiving optics NA_{lens} is given by $NA_{lens} = a/2f$, and that for optimum coupling of received optical power into the waveguide mixer we have chosen $NA_{lens} = NA_{wg}$.) All simulations were performed on a Sun SPARCstation™ using MATLAB™ computational software, and our results are shown in Figure 3.

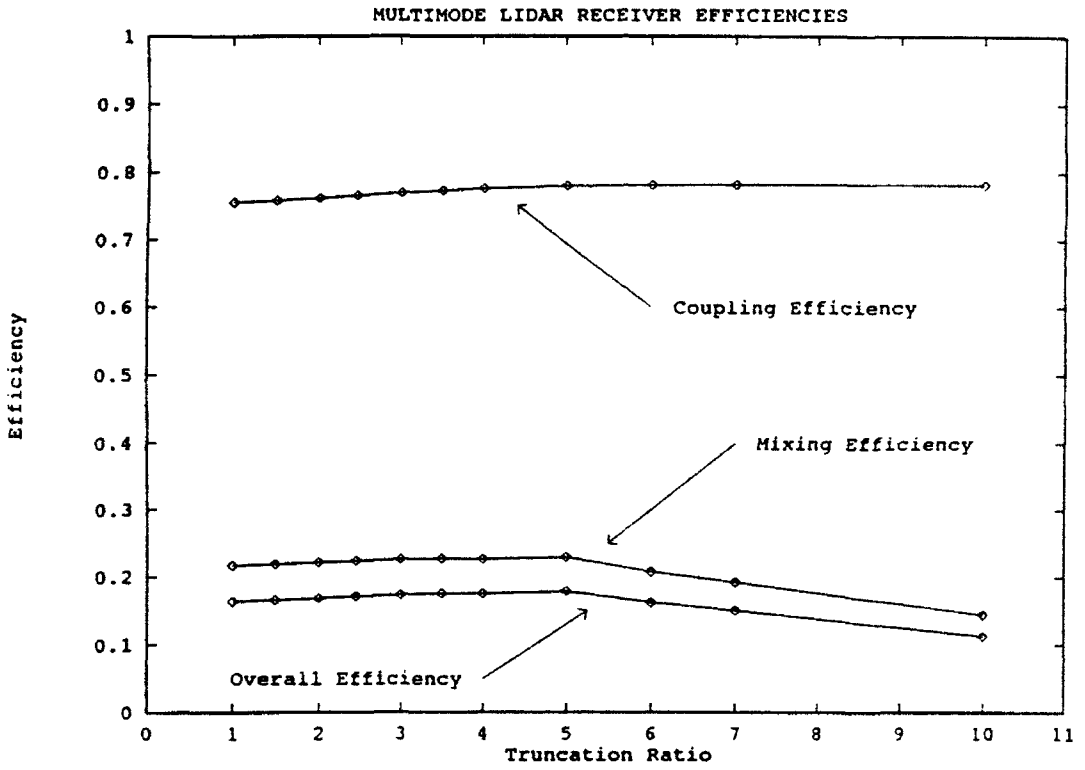


FIGURE 3
Mixing and Coupling Eficiencies

Note the following characteristics of Figure 3: first, the coupling efficiency is from 75-78%; second, the mixing efficiency is from 21-23%; and third, the overall efficiency (i.e., the product of the mixing and coupling efficiencies) is from 11-17%. Note also that all efficiency terms are, for the most part, independent of the truncation ratio. Now, at first glance it would seem that a maximum overall efficiency of 17% is quite low and that our multimode waveguide mixer has performed poorly. Primarily in order to restrict the required computation time, however, we have been considering a waveguide which supports only seventy-five guided modes. Though in the strictest sense this is a multimode waveguide, more realistic multimode waveguides commonly support many hundreds or thousands of guided modes. Since, loosely speaking, the coupling efficiency is related to the waveguide's ability to "sample" the many random spatial modes making up the returning signal field focused on the waveguide end face (see equation (12)), we should in practice expect coupling efficiencies approaching 100%. The mixing

efficiency is then the term which will ultimately limit the performance of the multimode optical mixer design. We see then, based on the mixing efficiency, that our multimode waveguide mixer is thus expected, in practice, to perform as well as many common free space and/or single mode fiber lidar mixers.

Much insight into the primary factors limiting the mixing efficiency can be gained from close inspection of equation (15), which describes the matrix elements $B_{k\ell}$. Note that as defined in equations (18) and (24), the Fourier transforms $\hat{U}_T$ and $\hat{W}$ of both the transmitted field U_T and the monostatic aperture function W, respectively, are even functions with on-axis maxima (as will likely be the case for most practical lidar system implementations). We would then expect that any given value of $B_{k\ell}$ would be maximum if the guided modes of interest $\xi_k(\overline{\rho})$ and $\xi_\ell(\overline{\rho})$ are also both even functions with on-axis maxima. $B_{k\ell}$ is then reduced if, say, one or both of the modes is even but does not have an on-axis maxima, or if one mode is even while the other is odd. In essence, then, we see that as the spatial correlation between the waveguide modes, the transmitted beam field and the monostatic aperture function increases, so does the mixing efficiency. This correlation will obviously be influenced by the optical waveguide geometry and thus by the form of the guided mode solutions. As for the square waveguide we have studied, it is then clear that the mixing efficiency is limited by the interaction of the many even and odd modal variations described in equations (23) and (24). Had we chosen a waveguide structure supporting many even modes with on-axis maxima, the the mixing efficiency would clearly have increased. Fortunately, the commonly available parabolically graded index multimode optical fiber preferentially supports such modes, and will thus likely be the candidate waveguide for implementation in practical multimode lidar optical mixer designs.

Finally, we mention that (at least in hind sight) it is clear that the efficiency terms plotted in Figure 3 should largely be independent of the transmit beam truncation ratio for the heterodyne lidar system we have investigated. This is because, historically speaking, beam truncation has been used in order to cause a stronger spatial matching bewteen LO and signal fields in free space mixing arrangements. Clearly, truncatating the transmit beam in our system does not strongly influence the spatial matching of the guided LO and signal fields in the waveguide mixer.

VI. CONCLUSION

The performance of a heterodyne lidar system incorporating a multimode optical waveguide mixer has been studied. We have seen that for a simple square optical waveguide supporting only seventy-five guided modes, mixing and coupling efficiencies on the order of 78% and 23%, respectively, can be achieved. For more realistic multimode waveguides, coupling efficiencies approaching 100% should be possible. In addition, for optimum waveguide geometries (for instance, a parabolically graded index multimode optical fiber) preferentially supporting many even modes with on-axis maxima, much higher mixing efficiencies should be expected as well. A multimode mixing arrangement is thus likely to afford a heterodyne lidar system many advantages. Not only are the mixing and coupling efficiencies expected to be high, but the robust nature of fiber optic systems in general will also yield the practical advantages of easy alignment, external noise immunity and ease of handling. In addition, as the mixing and coupling efficiencies for a multimode waveguide mixer are largely independant of the transmitted beam, only loose constraints need be applied in the lidar transmitter design.

Finally, future work in this area will likely center on the theoretical/numerical and experimental evaluation of multimode mixers employing graded index fibers. In addition, intermodal dispersion effects, especially as they would relate to pulsed lidar systems, should be more carefully studied and accounted for, as should be the effects of evanescent wave couplers which will likely be employed in practical multimode mixer designs. A less critical area of investigation may be to more fully study the excitation of the LO field modes.

VII. ACKNOWLEDGEMENTS

The author would like to gratefully acknowledge Major Martin B. Mark, Ph.D., of the Wright Laboratory Electro-Optic Techniques Group (WL/AARI-2), without whose patient guidance and assistance the successful completion of this research endeavor would have been impossible. The author would also like to thank the other WL/AARI-2 group members, especially Paul McManamon, Don Tomlinson, Mike Salisbury, Jay Overbeck and Larry Barnes, for their hospitality, encouragement and many enlightening discussions. Finally, special thanks goes to The University of Dayton Research Council and Drs. Mohammad Karim and Donald Moon, for their continuing support of the author's academic and scholarly pursuits.

VIII. REFERENCES

1. Gagliardi, Robert M. and Sherman Karp: Optical Communications, John Wiley and Sons, New York, 1976.

2. Rye, Barry J. and Rod G. Frehlich: "Optimal Truncation and Optical Efficiency of an Apertured Coherent Lidar Focused on an Incoherent Backscatter Target," Applied Optics, Vol. 31, No. 15, pp. 2891-2899, 20 May 1992.

3. Mark, Martin B.: "A Comparison of Free-Space and Fiber Mixer Performances," Wright Laboratories Technical Memo (to be published), Wright Patterson AFB, OH, 18 March 1992.

4. Mark, Martin B.: "Multimode Fiber Coupling Efficiency," Wright Laboratories Technical Memo (draft, to be published), Wright Patterson AFB, OH, April, 1992.

5. Stremler, Ferrel G.: Introduction to Communication Systems: Second Edition, Addison-Wesley, Reading, Massachusetts, 1982.

6. Snyder, Allan W.: "Excitation and Scattering of Modes on a Dielectric or Optical Fiber," IEEE Transactions on Microwave Theory and Techniques, MTT-17, No. 12, pp. 1138-1144, December 1969.

7. Shapiro, J.H., B.A. Capron and R.C. Harney: "Image and Target Detection with a Heterodyne-Reception Optical Radar," Applied Optics, Vol. 20, No. 19, pp. 3292-3313, 1 October 1981.

8. Marcuse, Dietrich: Theory of Dielectric Optical Waveguides: Second Edition, Academic Press, Boston, 1991.

TURBULENT HEAT TRANSFER IN COUNTER-ROTATING DISKS WITH THERMOGRAPHIC PHOSPHOR TEMPERATURE DETERMINATION

Jamie S. Ervin
Assistant Professor
Department of Mechanical & Aerospace Engineering

University of Dayton
121 Kettering Lab
Dayton, Ohio 45469-0226

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

August 31, 1992

TURBULENT HEAT TRANSFER IN COUNTER-ROTATING DISKS WITH THERMOGRAPHIC PHOSPHOR TEMPERATURE DETERMINATION

Dr. Jamie Ervin
Department of Mechanical & Aerospace Engineering
University of Dayton
Dayton, Ohio 45469-0226

Abstract

A state-of-the-art gas turbine design which implements counter-rotating blades has been recently submitted to the Air Force. As a first step in the study of the associated complex heat transfer phenomena, a counter-rotating disk system was designed. No experimental data concerning source flow between two parallel counter-rotating disks appears to have yet been published, and this planned study will be the first experimental examination of turbulent flow and local heat transfer between counter-rotating disks with radial outward flow.

The velocity field in rotating disk systems with source flow is complex. The fluid velocity vector has a component in the tangential direction from shear stresses induced by each rotating surface, and the mean radial velocity decreases with increasing radial distance from the source since the fluid moves through an increasing cross-sectional area. Centrifugal forces induce radial outward motion on the fluid near each disk surface in the presence of Coriolis forces. For non-isothermal systems, buoyancy forces present an additional complication. Furthermore, in the region of large pressure gradients, particularly as the radius tends to zero, flow separation and recirculation may occur.

One of the counter-rotating disks will be heated uniformly, and surface temperatures will be measured using a thermographic phosphor, lanthanum oxysulfide doped with Europium (La_2O_2S:Eu). When excited by ultraviolet radiation, thermographic phosphors emit fluorescent spectra characterized by unique emission lines, and the intensity of these lines varies with temperature. The ratio of the intensity of a high energy level emission line to that of a low energy level can be accurately correlated with temperature with geometric effects appropriately taken into account.

TURBULENT HEAT TRANSFER IN COUNTER-ROTATING DISKS WITH THERMOGRAPHIC PHOSPHOR TEMPERATURE DETERMINATION

Jamie S. Ervin

INTRODUCTION

A state-of-the-art gas turbine design which implements counter-rotating blades has been recently submitted to the Air Force and is shown in Figure 1.

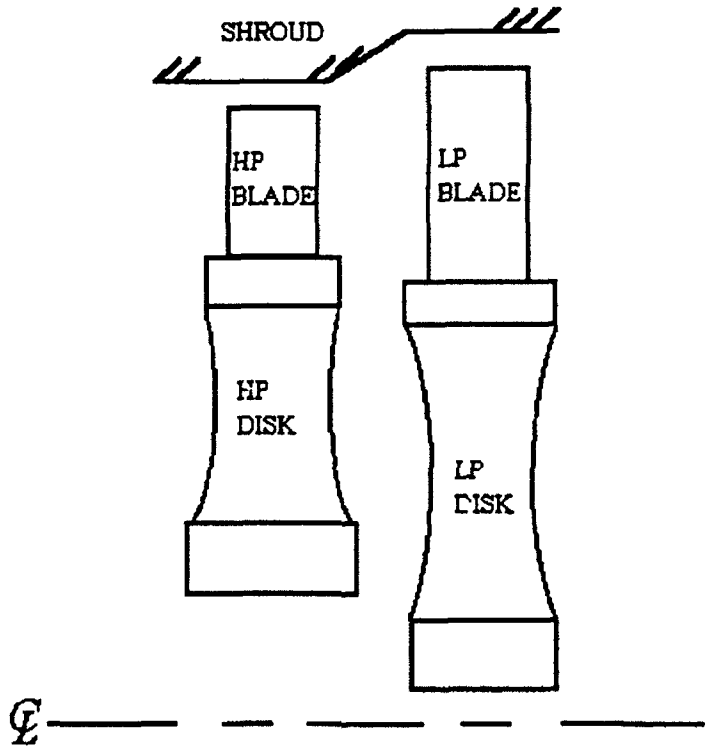


Figure 1 Counter-Rotating, Vaneless Design

As a first step in the study of the complex heat transfer phenomena associated with this turbine design, a counter-rotating disk system was designed. No experimental data concerning source flow between two parallel counter-rotating disks appears to have yet been published, and this planned study is the first experimental examination of turbulent flow and local heat transfer between counter-rotating disks with radial outward flow.

The velocity field in rotating disk systems with source flow is complex. The fluid velocity vector has a component in the tangential direction from shear stresses induced by each rotating surface, and the mean radial velocity decreases with increasing radial distance from the source since the fluid moves through an increasing cross-sectional area. Centrifugal forces induce radial outward motion on the fluid near each disk surface in the presence of Coriolis forces. For non-isothermal systems, buoyancy forces present an additional complication. Furthermore, in the region of large pressure gradients, particularly as the radius tends to zero, flow separation and recirculation may occur.

One of the counter-rotating disks will be heated uniformly, and surface temperatures will be measured using a thermographic phosphor, lanthanum oxysulfide doped with Europium (La_2O_2S:Eu). The use of thermographic phosphors for thermometry is more advantageous than conventional, direct contact devices such as thermocouples. Optical temperature sensing with thermographic phosphors eliminates the utilization of slip rings and thermally conductive probes which require the machining of holes in turbine blades for sensor placement. Thermographic phosphors can be applied to surfaces in very thin layers by plasma spray, vapor deposition, sputtering, or even by very fine air brushing. These phosphors are sensitive to relatively small changes in temperature and can be used at temperatures exceeding 1200 °C, as described in Dowell (1992).

Fluorescent emission of the thermographic phosphors with varying temperature can be used to determine the two dimensional temperature distribution at a heated surface. When excited by ultraviolet radiation, thermographic phosphors emit spectra characterized by unique emission lines, and the intensity of these lines varies with temperature. The ratio of the intensity of a high energy level emission line to that of a low energy level can be accurately correlated with temperature with geometric effects appropriately taken into account. This method is distinctly different from previous methods which correlate fluorescence lifetime with temperature as in Dowell (1992).

BACKGROUND

Previous studies involving the flow in rotating disk systems may be classified into three categories:

1. One fixed disk with one rotating disk

2. Co-rotating disks

3. Counter-rotating disks.

Although there are no experimental studies concerning turbulent convective heat transfer with source flow in counter-rotating disks, selected studies involving turbulent flow in rotor-stator disk configurations and in co-rotating disk systems are listed in order to give the reader some insight into the physics rotating disk systems and demonstrate that understanding of even these 'well studied' systems is far from complete: Kreith et al (1963), Bayley and Owen (1970), Metzger (1970), Bakke et al (1972), Owen et al (1974), Haynes and Owen (1975), Phadke and Owen (1983), Suryanarayana et al (1983), Uzkan and Lipstein (1986), Qureshi et al (1989), Iacovides and Theofanopoulos (1991), Bhavnani et al (1992), Bunker et al (1992a, 1992b), and Ko and Rhode (1992). A comprehensive review of the literature involving heat transfer in rotating disk systems, except for the counter-rotating case, is presented in Owen (1984).

Electron transitions in the dopant atoms cause fluorescence in thermographic phosphors. Dowell (1992) presents current theories about the change in fluorescence with temperature and summarizes research on rare-earth-doped ceramics over the last fifty years. He describes previous works involving fluorescent thermometry which rely on the correlation of fluorescence lifetime with temperature. He admits that fluorescence amplitude and ratios of amplitudes of different spectral lines also depend on temperature, but he purports that sensitive amplitude measurements traceable to metrological standards are difficult. (It may be parenthetically added that the ratios of amplitudes of spectral lines are presumed to eliminate geometric effects associated with the fluorescent surface.) However, temperature determination by correlation of fluorescence lifetime with temperature for non-isothermal surfaces can only provide point measurements rather than two dimensional field measurements. Nevertheless, for many engineering situations, the correlation of ratios of fluorescence amplitudes of spectral lines may be completely adequate for many engineering situations.

Taliaferro et al (1991) and Bizzak (1991) confirmed the feasibility of correlating ratios of fluorescence amplitudes with temperature for two dimensional temperature measurement. Taliaferro et al

(1991), using an ultraviolet lamp correlated the natural logarithm of the ratio of fluorescence intensity amplitude at the same wavelength with temperature. For calibration, the spectral intensity was measured initially at 25 °C and then at greater temperatures. The ratio of these intensities at greater temperatures to the intensity at 25 °C was taken, and then the natural logarithm of these ratios were correlated with temperature. Bizzak (1991), in generating a calibration curve, used intensities from two different wavelengths at a given temperature to form an intensity ratio. La_2O_2S:Eu provided fluorescence at the 510 nm and the 620 nm lines, but only the intensity of the 510 nm line varied with temperature. Also, the ultraviolet excitation source, a mercury arc lamp, produced undesirable infrared radiation resulting in surface heating. Clearly, more work needs to be done. The investigations of Taliaferro et al (1991) and Bizzak (1991) did not explore the effect of non-uniform irradiance of the ultraviolet radiation and the influence of a non-uniform phosphor coating, and the sources of uncertainty in temperature measurement in these studies was not fully explored.

OBJECTIVES

There were three objectives for this project:

1.) To design a counter-rotating disk system, arranged for utilization of the two spectral line thermographic phosphor technique, for the study of the associated turbulent convective heat transfer. Several planned tests will be scaled to actual turbine design by Ro, here between 0.1 and 0.3, with two gap spacings of 1.27 cm (0.5 inches) and 3.81 cm (1.5 inches). The resulting local and mean heat transfer rates will be determined with both disks rotating simultaneously at either 200, 400, or 800 rpm in opposite directions.

TABLE 1 INDEPENDENT VARIABLES WITH CORRESPONDING ROSSBY NUMBER

rpm	g (cm)	$\dot{m}$ (kg/s)	Ro
800	3.81	0.670	0.3
800	1.27	0.670	0.9
800	3.81	0.223	0.1
800	1.27	0.223	0.3
800	3.81	0.074	0.03
800	1.27	0.074	0.1
400	3.81	0.670	0.6
400	1.27	0.670	1.8
400	3.81	0.223	0.2
400	1.27	0.223	0.6
400	3.81	0.074	0.1
400	1.27	0.074	0.2
200	3.81	0.670	1.2
200	1.27	0.670	3.6
200	3.81	0.223	0.4
200	1.27	0.223	1.2
200	3.81	0.074	0.1
200	1.27	0.074	0.4

Two levels of uniform heat flux will be considered, and flow visualization will be used for selected tests along with turbulence intensity measurements. The tests to be considered are summarized above in Table 1. Following the evaluation of the experimental results determined at the conditions in Table 1, the effects of adding a shroud will be investigated. At this point, methods of heat transfer augmentation with multiple sites of air injection or swirl at the inlet radius will be explored.

2.) To improve the current thermographic phosphor technique which relies on the detection of two specific spectral intensities from the fluorescing phosphor by using an Nd:YAG laser as an ultraviolet source together with a gated camera. Additionally, to demonstrate the practical implications of the thermographic phosphor method, by measuring the radial temperature distribution of a cylinder with a uniformly imposed heat flux in cross flow.

3.) To establish the capabilities to perform in-house phosphor coating.

APPARATUS AND PROCEDURE

The description of the apparatus and procedures consists of three parts. The first part discusses the design of the counter-rotating disk arrangement, the second section concerns an exploration of different ultraviolet sources, and the third part involves radial temperature measurement on a heated cylinder in cross flow.

Counter-Rotating Disks

A counter-rotating disk system was designed and most items were fabricated and purchased for later use due to time limitations. An overall view of the proposed experimental arrangement is presented in Figure 2 and consists of two counter-rotating 61 cm diameter parallel disks. In order to reduce possible flexing of the polycarbonate disk, the disks are in an orientation parallel with the gravity vector.

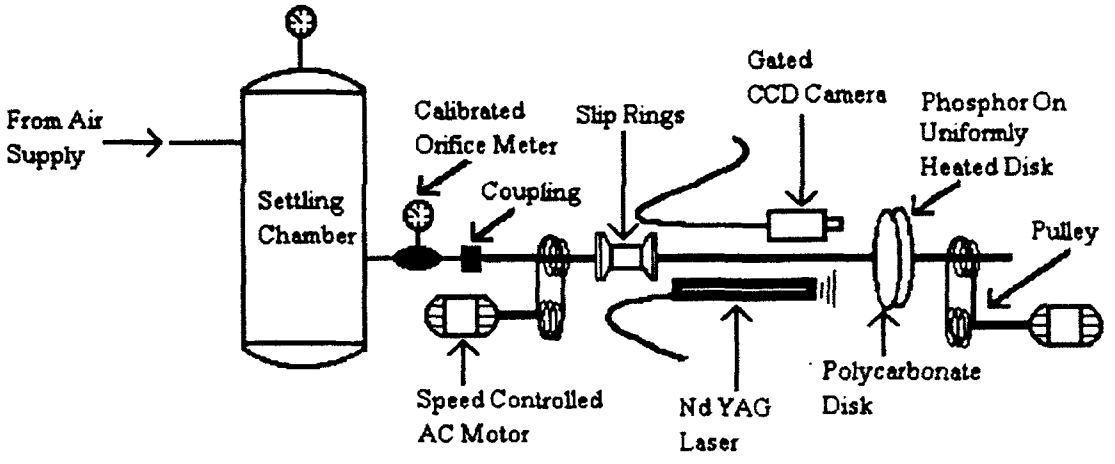


Figure 2 Counter-Rotating Disk Arrangement For Two Dimensional Temperature Measurement Of Uniformly Heated Disk

Air flows from the settling chamber through the calibrated orifice meter and enters the center of the rotating, unheated polycarbonate disk. The air then impinges on the counter-rotating heated disk, flowing radially outward. The CCD camera of Figure 2 is focused through the rotating polycarbonate disk on the heated disk and records the fluorescent surface emissions which are caused by ultraviolet excitation from the YAG laser. These fluorescent emissions from the heated disk correspond to the surface temperature distribution and are

used in the determination of local Nu. When not being used for the thermographic phosphor technique, the YAG laser can produce a light sheet for paraffin mist flow visualization. The transparent polycarbonate disk permits high speed photography of the mist from orthogonal viewing directions, either parallel or perpendicular to the polycarbonate disk orientation. When velocity measurements are desired, the probing rays from the TSI laser velocimetry system are selected to be within the disk spacing. More specific details of the experimental components are given in the following paragraphs, beginning with the counter-rotating disks.

The heated disk, shown in Figure 3, is composed of a retaining aluminum cylindrical shell which contains polyurethane insulating material below a laminated heater. This laminated heater consists of three layers: a heating element etched from an electrically resistive material on a printed circuit board, an electrically insulating, but thermally conducting epoxy, and an inconel outer surface (0.05 mm thick).

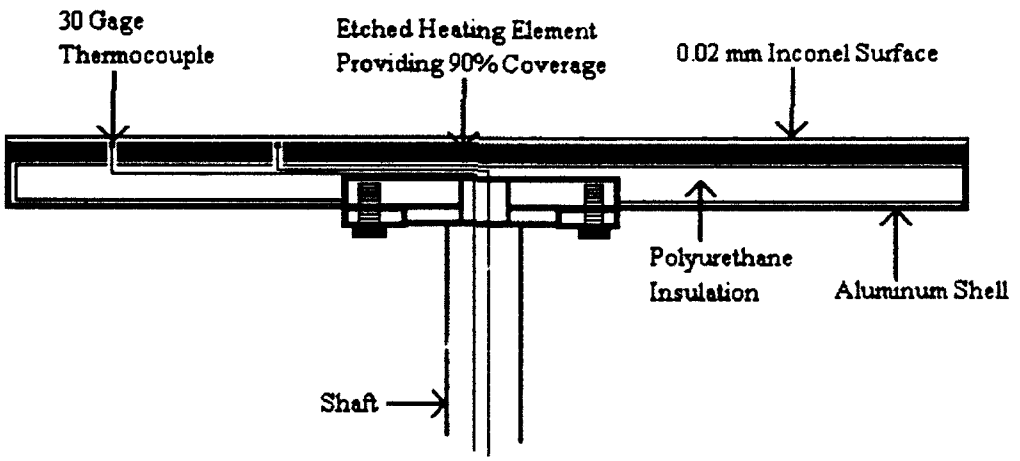


Figure 3 Sectional View of Disk For Uniform Heating

The heating element is etched to provide a uniform coverage of 90% of the available surface area, insuring a uniform wall heat flux with the application of a constant voltage. A single piece of inconel is used for the outer surface of the laminated heater because it does not readily form surface oxides which results in a nearly constant emissivity over the temperature range of -134 °C to 560 °C. As a result, reasonable estimates of radiation losses can be made. The outer surface of the heated disk is to be airbrushed with an extra-fine

airbrush with La_2O_2S:Eu phosphor and then oven cured. Type J 30 gage thermocouples are also attached to the underside of the inconel surface for verification of the phosphor surface temperature determinations and for the initial calibration required for the thermographic phosphor technique. A thermocouple slip ring assembly transmits the thermocouple signals to PC based data acquisition.

The polycarbonate disk has a transmissivity of 92% for ultraviolet spectra so as not to absorb ultraviolet radiation during surface temperature measurement using the thermographic phosphors. Also, the polycarbonate disk is transparent to visible light, permitting use of a laser light sheet for flow visualization.

The disks are each mounted to a pulley driven shaft, which are in turn, connected to two speed-controlled 20 hp ac motors by belts and pulleys. The angular speed of each disk can then be independently controlled up to 1750 rpm, and the spacing between the disks is set by adjusting the shaft bearings. Each rotating shaft has an optical encoder used to monitor the rotational speed of the disks and for synchronization with the YAG laser.

The heated disk is powered by means of a German-silver power slip ring assembly which can transmit up to 80 Amps at 1000 rpm. The ac power source is capable of providing a maximum of 15 kW of power to the heated disk surface. Voltage measurement across a calibrated shunt electrically in series with the heated surface permits determination of the current, and voltage measured across the heated surface allows the calculation of the power supplied.

Temperature measurements at the heated disk surface for determination of local Nu are planned to be made using a thermographic phosphor technique, and the components necessary for this method are described below. In this study, the ratio of the intensity of the fluorescent emissions at 510 nm and 620 nm will be calibrated with the surface temperature of the heated disk in the range of 24 °C to 66 °C. The La_2O_2S:Eu will be excited by ultraviolet radiation from the pulsed YAG laser, described above. Image acquisition of the radiation emitted by the fluorescing phosphors will be accomplished using a CCD video camera with nanosecond gating capabilities, optical filters, and a PC based image acquisition card. The video camera, laser, and ac motor controllers are to be synchronized with TTL circuitry such that for a given radial location, the laser beam strikes the same location on the heated disk in each revolution.

The procedure for obtaining heat transfer data from the lower heated disk is now outlined. Prior to conducting the thermophosphor experiments, regions of the phosphor coated surface above the thermocouples would be used to calibrate the optical system for temperatures between 24 °C and 66 °C. By means of a valve, the mass flow rate of air leaving the settling tank would be adjusted to the desired level, and the voltage impressed on the heating element of the heated disk would be set to give the required uniform heat flux. Following attainment of steady-state conditions, the thermographic phosphor technique would be used to determine the two dimensional temperature distribution of the motionless heated disk, permitting calculation of local Nu and heat transfer coefficients. This set of Nu obtained from the stationary disk forms the baseline to which other local Nu resulting from counter rotation and modifications for heat transfer augmentation, including adding swirl to the flow and multiple jet injection, can be compared.

It should be mentioned in passing that heat loss is energy generated by the heating element of the heated disk that is not transferred to the air flow. Calculated estimates of radiation and conductive heat transfer losses from the heated disk surface are to be performed.

Ultraviolet Sources

Previously in a study by Bizzak (1991), a 100 Watt mercury arc lamp was used at WL/POTC as an ultraviolet source for excitation of La_2O_2S:Eu. However, it was found that the mercury arc lamp produced significant optical noise in the visible range, and infrared radiation produced surface heating of the illuminated target. In addition, the ultraviolet irradiation was highly non-uniform. As a result, it was decided to explore alternate sources of ultraviolet excitation. To reduce optical noise, the overhead lights were switched off for all of the following described tests. The video camera used in the current tests, Xybion model ISG205, has a resolution of 776 pixels by 488 pixels and has a spectral response in the 380 nm to 920 nm range, but the spectral response falls off sharply near the spectral lines of interest, 510 nm and 620 nm.

The first ultraviolet source considered here was a high intensity Mineral Light model UVG-54 manufactured by UVP Inc. which is a continuous source of emission at 254 nm. This lamp, which appeared to have a nearly uniform irradiance, was directed at a 279 mm x 152 mm x 0.05 mm sheet of stainless steel foil that was air brushed with La_2O_2S:Eu and located at a distance of 20 cm from the Mineral Light.

Although the ultraviolet light functioned satisfactorily, it would be more desirable to control the power output of the illumination source. A pulsed Nd:YAG laser with a high power density (100 mJ), a short pulse (8 ns), and a Gaussian beam profile address this concern. The suitability of a Quanta-Ray GCR-11 pulsed Nd:YAG laser as an ultraviolet source was examined next. The manufacturer of this model claims that the optical resonator produces a 'near Gaussian' spatial profile of beam intensity. However, illumination with the YAG beam expanded with a spectrally coated negative lens on the foil target coated with phosphor resulted in a spatial distribution of fluorescent emission that consisted of a series of concentric rings, resembling the cross-section of a tree trunk. The size and shape of these rings appeared to vary with time. As a result, attempts were made using diffuse transmission and reflection to filter the beam profile to eliminate the spatial variation in beam intensity.

Diffuse reflection did not provide adequate target irradiation and also presented a safety concern. Diffuse transmission through a quartz filter permitted the YAG laser to be operated as an adjustable ultraviolet source. A diffusing filter was constructed from a 0.32 cm thick rectangular quartz piece. It was found that sandblasting only one side of the quartz produced the fine, randomly sized and spaced surface pits required for diffusing the beam, without degrading the beam intensity beyond a useful level. The gated Xybion camera was synchronized with the Q switch TTL output of the YAG laser. The Q switch output, however, had to be inverted in order to provide a low external trigger signal for the camera, and a TTL inverter was constructed for this purpose.

Heated Cylinder in Cross Flow

The third portion of the description of the apparatus and procedures concerns measurement of the surface temperature of a heated cylinder in crossflow and begins with a description of the cylinder. The test cylinder was constructed from a 279 mm x 152 mm x 0.05 mm sheet of stainless steel foil wrapped around a 8.9 cm diameter insulating material. This material, polyurethane, came in sheets from which disks were cut and glued together to form a cylinder. The stainless foil was firmly clamped by two-piece thick aluminum clamps which had a nominal inner diameter of 8.9 cm to two 8.9 cm diameter brass disks located at each end of the cylinder of insulation. Electrical grease was used between the clamps and the foil to ensure radially uniform electrical contact. The outer surface of the foil was coated with an air brushed coating of La_2O_2S:Eu, while the mid plane of the inner surface is instrumented with thermocouples placed at 22.5 ° intervals, beginning at the front stagnation point. Since the resistance of the stainless steel changes very little in the temperature range of interest, application of a constant voltage between the two disks imposes a uniform heat flux at the surface of the cylinder. Voltage measurement across a calibrated shunt electrically in series with the heated surface permits determination of the current, and voltage measured across the heated surface permits the calculation of the power supplied. This combination of the insulating material and the thin sheet of stainless steel results in a relatively low thermal inertia, which is expected to reduce the time period required to reach steady-state conditions from ambient conditions for a given heat flux level.

The test cylinder of this study is schematically represented in Bizzak (1991), along with the optical surface temperature measurement equipment. However in the present investigation, the ultraviolet excitation is provided by an Nd:YAG laser, while image acquisition is accomplished with the gated Xybion camera described previously. The arrangement and use of these components is more fully described below, and details of the wind tunnel construction appear in Bizzak (1991).

The wavelength and bandwidth of the fluorescent emission captured by the video equipment is selected by positioning either a 510 nm or a 620 nm bandpass filter (each with a bandwidth of 10 nm) in front of the camera. The Xybion camera with a macro lens provided an image of the illuminated portion of the cylinder. The analog video signal is digitized, and each pixel is assigned an 8 bit intensity value, which corresponds to some level out of a possible 256 gray scale levels. The corresponding images at the two

wavelengths are acquired sequentially in real time and stored in binary files for processing on newly written and acquired image enhancement and analysis software.

Selected regions of the phosphor surface were intended to be calibrated. With the cylinder in still air, a quasi-steady surface temperature for a given voltage is eventually reached and measured with calibrated thermocouples attached to the interior stainless steel surface. The uncertainty in the temperature determination with these thermocouples and instrumentation is $\pm$ 0.3 °F over the temperature range of interest. The acquired and digitized images are processed to obtain a calibration curve for each rectangular region, which relates emission intensity ratio to temperature. After sections of the cylinder are calibrated for no flow conditions in the wind tunnel, tests at a Reynolds number of 2×10^5 were to be performed. Due to time constraints, actual testing of a phosphor coated cylinder was not performed. However, initial tests were performed to test the uniformity of the imposed heat flux and the recently improved method of attaching 30 gage thermocouples to the inner side of the cylinder and appear in following paragraphs.

RESULTS

The results of preliminary tests are now presented, beginning with the tests of the ultraviolet light and phosphor coated foil. Since the foil was assumed to be isothermal at the ambient temperature, it was expected that the spatial distribution of the fluorescent emission at 510 nm from the stainless steel target would ideally be a constant. Analysis of images digitized with the gated camera and frame grabber indicate that this emission is not a constant and exhibits some variance about a mean value. Later, a soldering iron was held against the back surface of the illuminated foil and switched on. The graphic changes in the fluorescence near the point of soldering iron contact were particularly striking. In fact, a pattern resembling the buoyant plume rising from the soldering iron and foil was visible.

In another test with the ultraviolet light, a set of four 30 gage thermocouples were attached to the vertical midplane at the back of the foil. No rise in temperature was detected by the thermocouples after one hour of irradiation. It appears that this ultraviolet lamp is much more satisfactory for temperature measurement than the mercury arc lamp. The power output of the ultraviolet lamp, however, is not

adjustable and is limited to the factory setting. Since the Mineral Light is not an ideal continuous source and requires 115 VAC, the camera would have to be synchronized with the line voltage to avoid apparrent temporal variations in irradiation.

Another isothermal test, similar to that using the Mineral Light, was performed with the YAG laser. Again the intensity of the fluorescent emission at 510 nm from the steel target was found to have a mean value and variance. Figure 4 shows the normalized probability density function for the emitted 510 nm line and the quartz filter for the YAG laser excitation, and a probability density function of this shape would be expected: a relatively 'small' standard deviation. Here, the intensity is defined in terms of a pixel gray scale value between 0 and 255. For purposes of illustration, Figure 5 shows the normalized probability density function near the same section of foil only now with a soldering iron held behind the foil. As can be observed in Figure 5, the standard deviation has now doubled. Figures 6 and 7 show 64 level gray scale images (64 gray scale levels is the resolution of the available laser printer). Figure 6 corresponds to Figure 4, with no soldering iron present; Figure 7 corresponds to Figure 5 and the application of a soldering iron.

Verification of the uniformity in heating of the cylinder was performed by measuring the temperature distribution at the mid plane of the cylinder periphery, and these temperatures are presented in Table 2. Measured temperatures at corresponding negative and positive angular locations are symmetric as expected, with 0° being the front stagnation point on the cylinder. Similar results were also obtained for conditions of no flow in the wind tunnel.

TABLE 2 TEMPERATURE DISTRIBUTION AT CYLINDER RADIUS FOR $Re_D = 2.3 \times 10^4$

Angular Location (Deg.)	Temperature (°C)
0	34.9
45	35.1
-45	35.1
90	41.1
-90	41.1

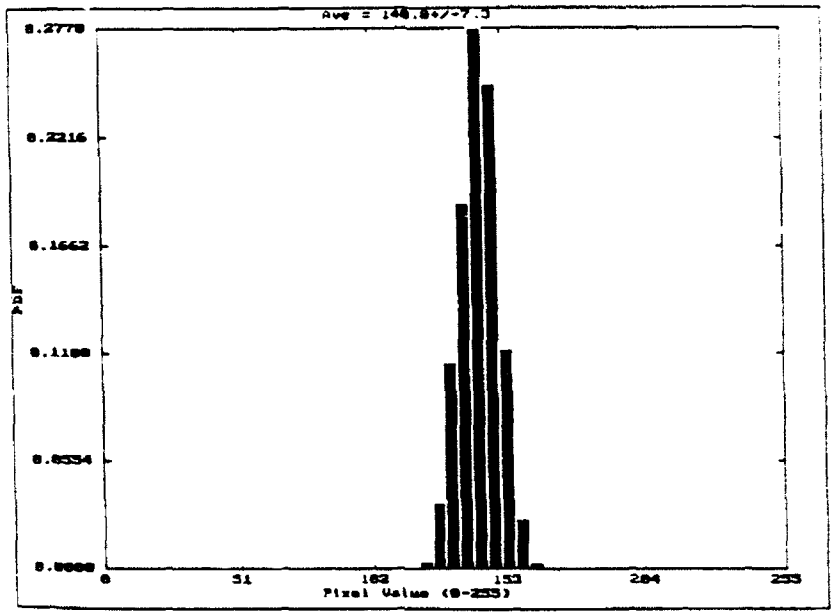


Figure 4 Normalized probability density function for YAG laser excitation and isothermal conditions.

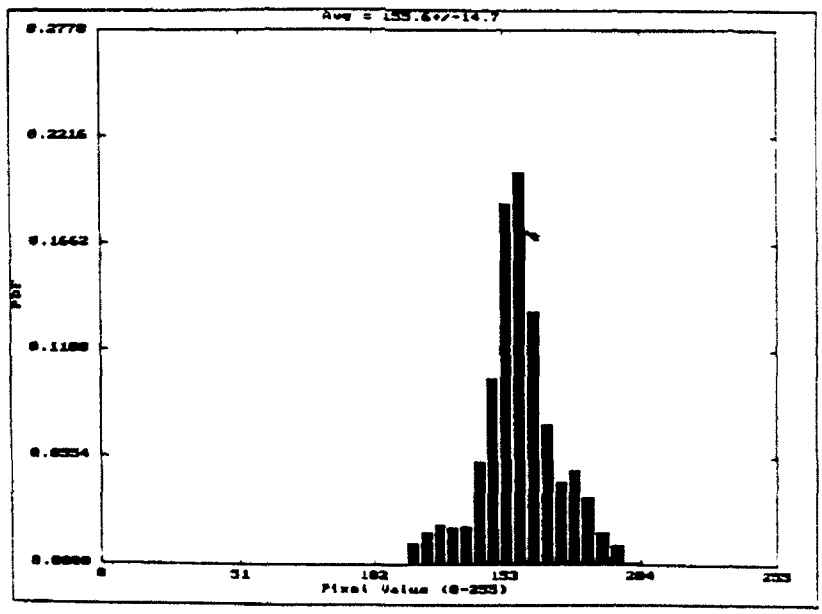


Figure 5 Normalized probability density function for YAG laser excitation with heating.

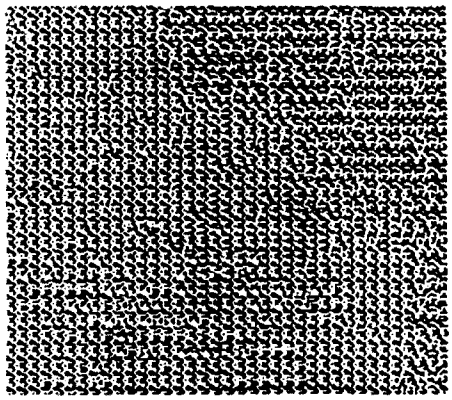


Figure 6 64 gray scale representation of isothermal surface.

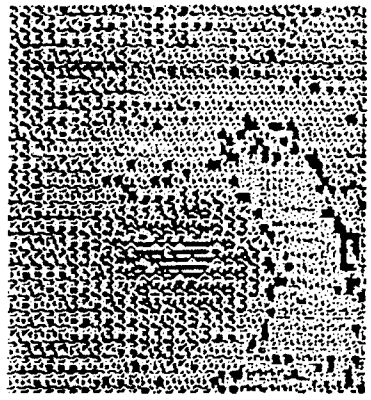


Figure 7 64 gray scale representation of heated surface.

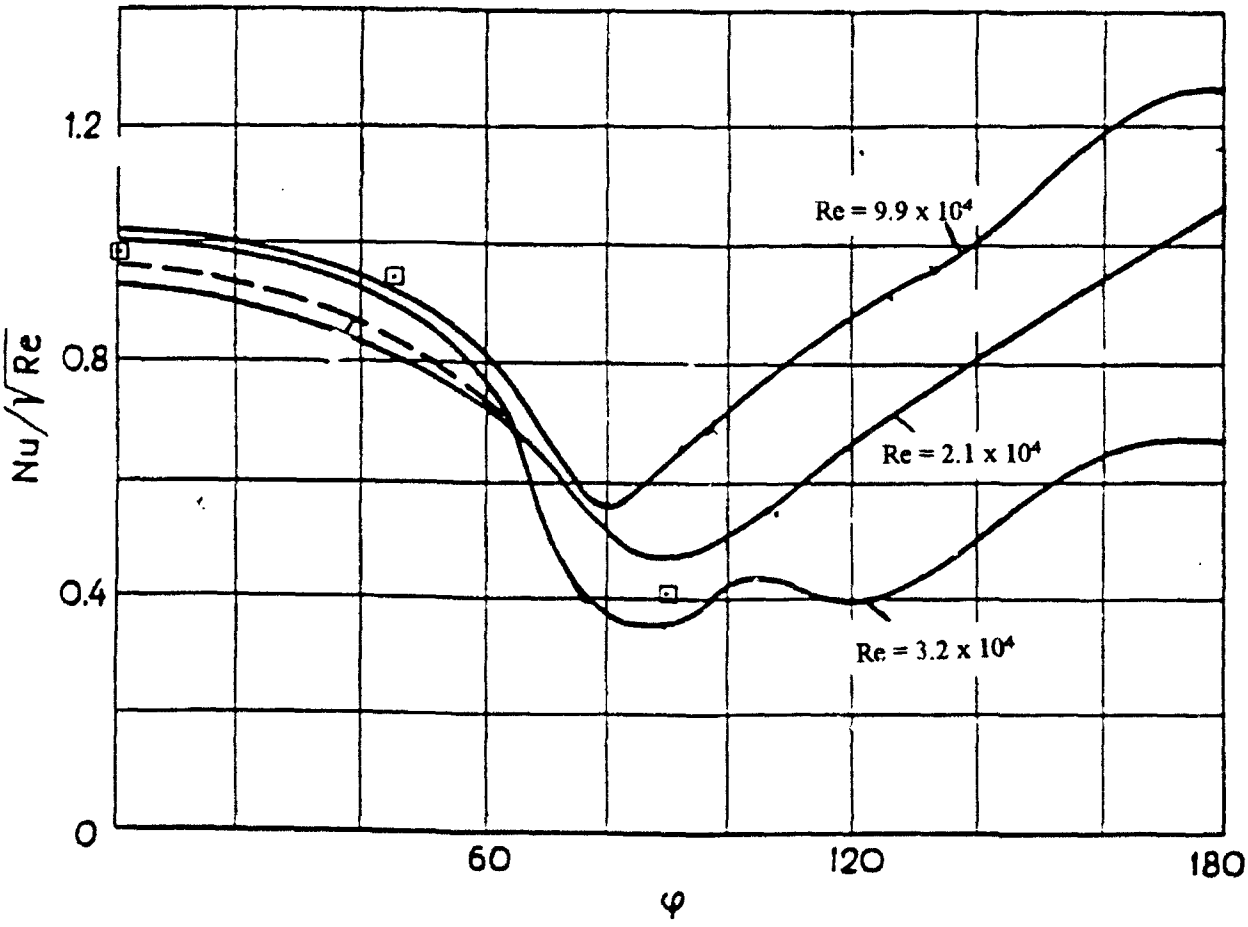


Figure 8 Heat transfer parameters and angular location on cylinder.

Figure 8 presents dimensionless heat transfer parameters measured at $Re_D = 3.2 \times 10^4$ for the current set up with data curves from prior works in Zaukaskas (1972). The uncertainty in the present measurements of Figure 8 are estimated to be 12 % of the indicated value. From this Figure, it is observed that the present cylinder and wind tunnel combination provide heat transfer data consistent with prior works.

RECOMMENDATIONS

As a result of this study, the following recommendations are made:

1. Previous work has used either a ratio of intensities involving a reference temperature and a single wavelength or a ratio of intensities with two spectral lines. The technique which uses a reference temperature with a single wavelength offers some advantage in that it requires only one bandpass filter. In contrast, the two spectral line method requires two bandpass filters with either mechanically stable and movable filter holders or additional optical components to combine images at both spectral lines on the same video frame. The error involved with each of the two methods should be analyzed and compared on a statistical basis.

2. The effect of the uniformity of the phosphor coating and the method of application of the coating on the uncertainty of the resulting temperature measurement should be examined. The effects from vacuum deposition, sputtering, or air brushing on temperature accuracy is unknown.

3. An alternate video system with gating capabilities should be procured. This new system should be sensitive to illumination in the 500 to700 nm range as opposed to the existing system which is more sensitive to radiation in the ultraviolet range. The accuracy of the temperature measurement could be enhanced if a 10 or 12 bit digitization were employed.

4. Further work using the YAG should be done to evaluate the appropriateness of the present quartz filter. Additionally, the suitability for fiber optics of the peak power associated with ultraviolet radiation from the YAG source should be discerned by experimentation.

5. The effect of a software bandpass filter on acquired optical data should be examined.

BIBLIOGRAPHY

Bakke, E., Kreider, J., and Kreith, F., Turbulent Source Flow between Parallel Stationary and Co-rotating Disks, *J. Fluid Mech.*, 1973, **58**, pp. 209-231.

Bayley, F. and Owen, J., The Fluid Dynamics of a Shrouded Disk System with a Radial Outflow of Coolant, *J. of Eng. for Power*, July 1970, pp. 335-341.

Bhavnani, S. et al, An Experimental Study of Fluid Flow in Disk Cavitites, *J. of Turbomachinery*, April 1992, **114**, pp. 454-461.

Bizzak, D., The Use of Thermographic Phosphors for Surface Temperature Measurement, AFOSR Summer Faculty Report, 1991.

Bunker, R., Metzger, D., and Wittig, S., Local Heat Transfer in Turbine Disk Cavities: Part I-Rotor and Stator Cooling with Hub Injection of Coolant, *J. of Turbomachinery*, January 1992a, **114**, pp. 211-220.

Bunker, R., Metzger, D., and Wittig, S., Local Heat Transfer in Turbine Disk Cavities: Part II-Rotor Cooling With Radial Location Injection of Coolant, *J. of Turbomachinery*, January 1992b, **114**, pp. 221-228.

Dowell, L., Fluorescence Thermometry, *Appl. Mech. Rev.*, July 1992, **45**, pp. 253-260.

Haynes, C. and Owen, J., Heat Transfer from a Shrouded Disk System with a Radial Outflow of Coolant, *J. of Eng. for Power*, January 1975, pp. 28-36.

Iacovides, H. and Theofanopoulos, Turbulence Modeling of Axisymmetric Flow Inside Rotating Cavities, *I.Int. J. Heat and Fluid Flow*, March 1991, **12**, pp. 2-11.

Ko, S. and Rhode, D., Thermal Details in a Rotor-Stator Cavity at Engine Conditions with a Mainstream, *J. of Turbomachinery*, April 1992, **114**, pp. 446-453.

Kreith, F., Doughman, E., and Kozlowski, H., Mass and Heat Transfer from an Enclosed Rotating Disk with and without Source Flow, *J. of Heat Transfer*, May 1963, pp. 153-163.

Metzger, D., Heat Transfer and Pumping on a Rotating Disk with Freely Induced and Forced Cooling, *J. of Eng. for Power*, July 1970, pp. 342-348.

Owen, J., Haynes, C., and Bayley, F., Heat Transfer from an Air-Cooled Rotating Disk, *Proc. R. Soc. Lond.*, 1974, A **336**, pp. 453-473.

Owen, J., Fluid Flow and Heat Transfer in Rotating Disc Systems, *Heat and Mass Transfer in Rotating Machinery*, D. Metzger and N. Afgan, eds. Hemisphere, Washington DC, 1984, pp. 81-103.

Phadke, U. and Owen, J., An investigation of Ingress for an "Air-Cooled" Shrouded Rotating Disk System with Radial-Clearance Seals, *J. of Eng. for Power*, January 1983, **105**, pp. 178-183.

Qureshi, G. et al, Heat Transfer Measurements for Rotating Turbine Discs, Presented at the Gas Turbine and Aeroengine Congress and Exposition, Toronto, 1989. ASME Paper No. 89-GT-236, pp. 1-8.

Suryanarayana, N., Scofield, T., and Kleiss, R., Heat Transfer to a Fluid in Radial, Outward Flow Between Two Co-axial Stationary or Corotating Disks, *J. Heat Transfer*, August 1983, **105**, pp. 519-526.

Uzkan, T. and Lipstein, N., Effects of Honeycomb-Shaped Walls on the Flow Regime between a Rotating Disk and a Stationary Wall, *J. of Eng. for Gas Turbines and Power*, July 1986, **108**, pp. 553-561.

Zukauskas, A., Heat Transfer from Tubes in Crossflow, Advances in Heat Transfer, 1972, **8**, pp. 93-159.

Molecular Modeling of Materials for Non-Linear Optical Applications

B. L. Farmer
Professor
Department of Materials Science and Engineering

University of Virginia
Thornton Hall
Charlottesville, VA 22903-2442

Final Report for:
Summer Research Program
Materials Laboratory
Wright-Patterson Air Force Base

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D.C.

October 1992

Molecular Modeling of Materials for Non-Linear Optical Applications

B. L. Farmer
Professor
Department of Materials Science and Engineering
University of Virginia

Abstract

Molecular modeling calculations were undertaken on a series of materials having potentially useful non-linear optical properties for laser-hardening applications. Molecular mechanics and molecular dynamics calculations examined the conformational properties of several-siloxane based liquid crystalline materials. Siloxanes having cyclic, star, and linear geometries were examined. Biphenyl and cholesteryl mesogens attached by leader groups having vinyl or allyl components were considered. Irrespective of the geometry of the siloxane core, the molecules all had sufficient flexibility to allow close associations between mesogens to develop. Associations between cholesteric mesogens were particularly stable.

The properties of several graphite-based materials were also examined. Semiempirical molecular orbital calculations were used to determine the deformation behavior of benzene, naphthalene and pyrene. All showed remarkable tendency to remain planar, even though imposed constraints introduced severe geometric distortions. Calculations on cyclic and linear acene oligomers were also carried out. The energies of linear acenes were consistently lower than those of the corresponding cyclic molecule. For small cycles containing 6-10 benzene units, the energy dependence on ring size is less sharp than it is for larger cycles.

Molecular Modeling of Materials for Non-Linear Optical Applications

B. L. Farmer

INTRODUCTION

With the proliferation of laser weapons, materials having appropriate optical properties must be developed which can serve, by any of a variety of mechanisms, as protection devices for guidance and surveillance equipment and the operators of those systems. The most appropriate criteria by which materials should be evaluated are not yet well defined, and will, no doubt, depend strongly on the specific application. Nonetheless, studies of the optical properties of materials of potential use in laser hardening applications are proceeding on several fronts. Computational materials science promises to be a very useful technique for evaluating the optical - and other - properties of materials, and offering the prospect of tailoring, or even designing, a material to have a specific set of properties. Of course, suitable optical properties must be coupled with necessary mechanical and chemical properties before the result is a useable material. Computational modeling is also well suited to investigating these kinds of properties as well.

The work carried out during the summer program falls into three areas. First, a computational modeling study of a series of cyclic siloxanes was carried out. Second, a study of graphitic materials was also undertaken using computational modeling. Third, a two-day mini-symposium on Computational Materials Science for Laser Hardened Materials Research, was co-organized with Dr. Ruth Pachter of WI/MLPJ.

COMPUTATIONAL MODELING METHODS

Three methods of computational chemistry were used. Semiempirical molecular orbital calculations were used to study deformation and structure of graphitic materials. Molecular mechanics was used to generate and evaluate starting conformations of star and cyclic siloxanes. Molecular dynamics was used to evaluate and compare the associations between mesogens attached to siloxane stars and cycles.

Molecular mechanics (MM) and molecular dynamics (MD) calculations are based on a classical description of the bonding in a molecular system. They use an empirical force field to describe the interactions between atoms forming a molecule and the interactions between molecules. Bond lengths, bond angles, torsion angles, out-of-plane bending, and non-bonded interactions (van der Waals and coulombic) are described in terms of force constants derived from a large body of vibrational and crystallographic data. MM calculations use this description to survey and optimize possible geometries and conformations of individual molecules, and to evaluate and optimize possible packing modes between molecules.

Molecular dynamics uses the same force-field description and extends the method to consider effects of temperature and time. A minimum energy structure determined by MM is a static representation of the system as it might exist near 0 K. MD includes time and temperature effects by calculating the forces acting on each atom, determining each atom's velocity by integrating the equations of motion over a short time interval, and scaling the total kinetic energy to the desired temperature. The MD simulation thus

yields a picture of the system as a function of time at each temperature examined.

A primary limitation of MM and MD methods is that the bonding is fixed throughout the computation. Chemical reactions cannot occur, nor can changes occur in the electron distribution or bonding as a function of time or molecular geometry. The advantages of the methods are that they can deal with systems containing very large numbers of atoms (hundreds), and the dynamics of the system can be explored. Semiempirical molecular orbital (MO) calculations, while limited to many fewer atoms (tens), can be used to explore chemical (or photochemical) processes, electron distributions as a function of substituent and conformation, and electronic interactions. MO methods are thus invaluable in investigations of optical properties and of large scale molecular deformation.

COMPUTATIONAL MODELING OF SILOXANE-MESOGEN SYSTEMS

Siloxane-based liquid crystalline materials have been studied as potentially providing a chemically- and mechanically-suitable conveyance system for low molecular weight or substituent NLO moieties. A molecular dynamics study has previously been carried out for a cyclic pentasiloxane (i.e. five -Si-O- units in the ring) having biphenyl and cholesteryl mesogens attached to the siloxane ring by an allyloxybenzoate spacer group. That study especially focused on three different conformations (suggested by x-ray studies) of the siloxane ring, which gave overall molecular shapes characterized as cone, disk and cylinder models.

Star siloxanes having two biphenyl and two cholesteryl mesogens have also been studied using MD. The tetrahedral geometry of the siloxane star molecule keeps the mesogens tethered while making the B-B, C-C and B-C interactions equally accessible. The tendencies of the star toward specific pairwise interactions, the influence of a proximate third (or perhaps fourth) mesogen, and the role of the spacer (its length and structure) in promoting these interactions (collectively or specifically) were investigated. The frequency and longevity of associations and dissociations and the ability to achieve optimum interactions characterized.

This study extended the earlier investigations of the star siloxane to consider the effects of changing the length of the spacer group used to connect the mesogen to the siloxane. Figure 1a shows the siloxane star, substituted with methyl branches, and figure 1b shows the star substituted with four (B) mesogens, with vinyl groups in the leader instead of the allyl groups previously studied. Stars having four (C) mesogens or two (B) and two (C) mesogens were also studied. Linear siloxane dimers and trimers, examples of which are shown in figure 2, having (B) and (C) mesogens were also examined. Finally, a cyclic siloxane having four -Si-O- units in the ring and having (B) and (C) mesogens attached was also studied. The methyl-substituted cycle is shown in figure 3a and the biphenyl substituted cycle is shown in figure 3b. In general, the MD simulations showed that irrespective of the geometry of the siloxane moiety, there was sufficient flexibility in the leaders for the mesogens to associate with each other very rapidly during the simulation. Associations between cholesteryl mesogens, once formed, tended to be quite stable, while associations between biphenyl mesogens or between

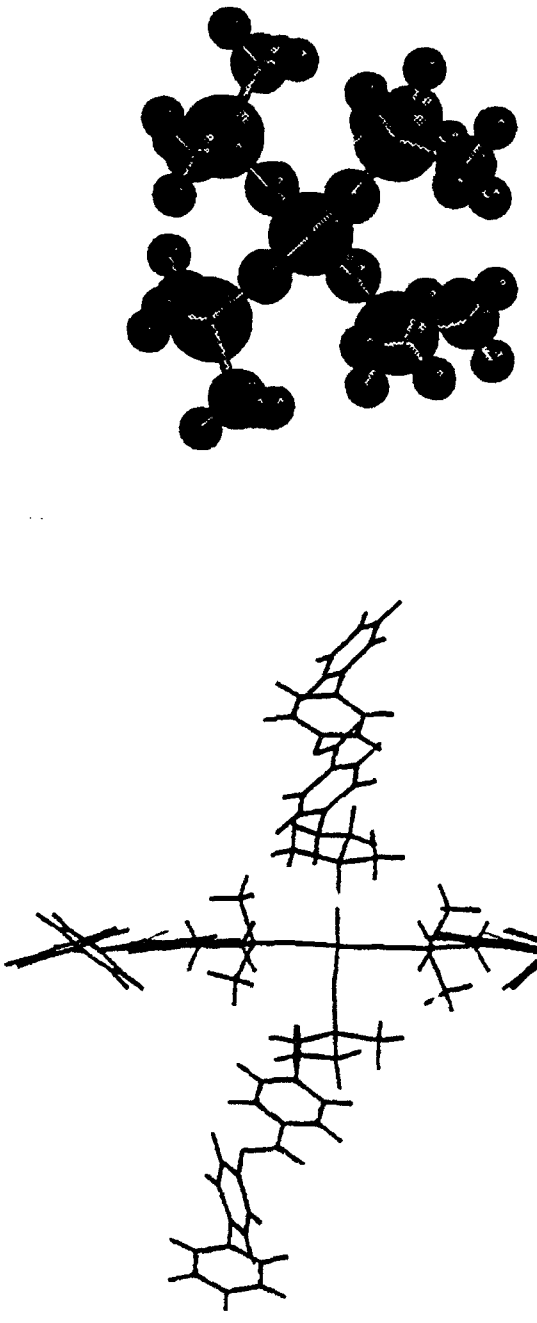


Figure 1a. Methyl-substituted star siloxane molecule.
b. Star siloxane substituted with four biphenyl mesogens having vinyl leader groups.

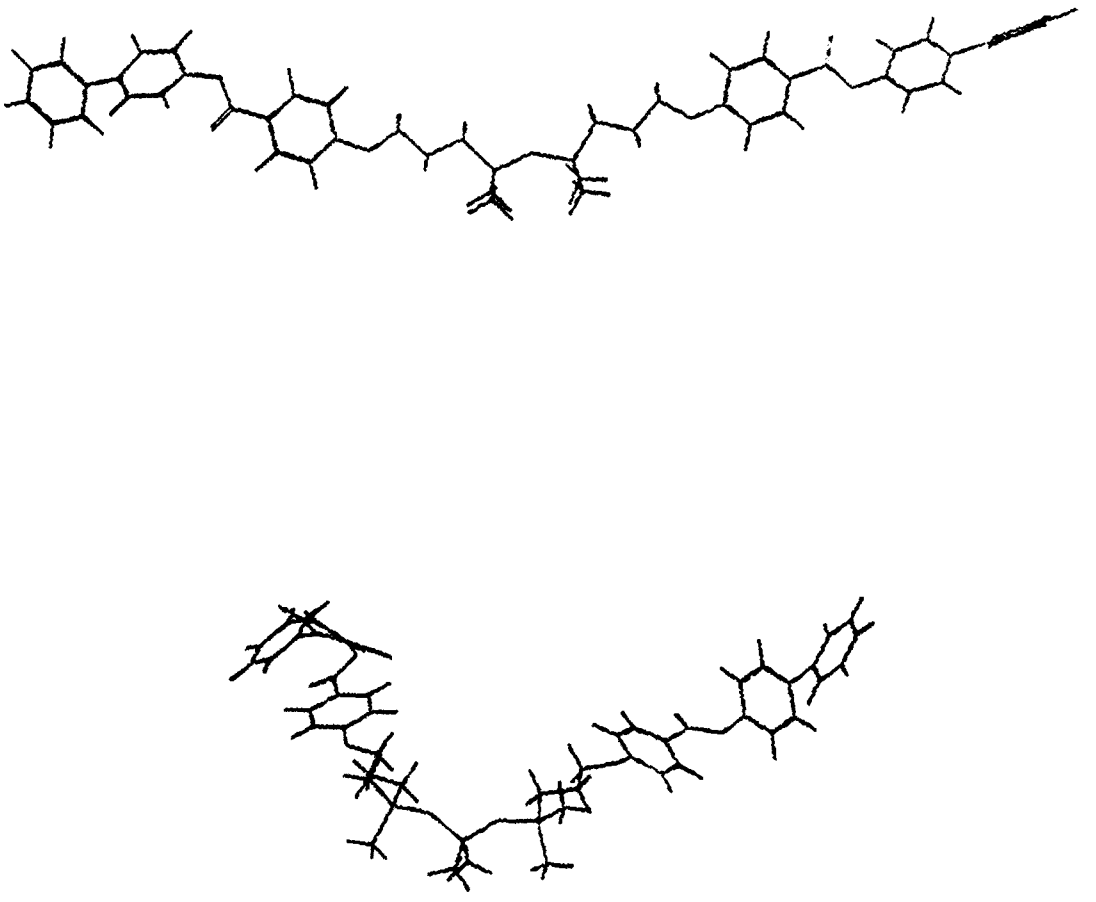
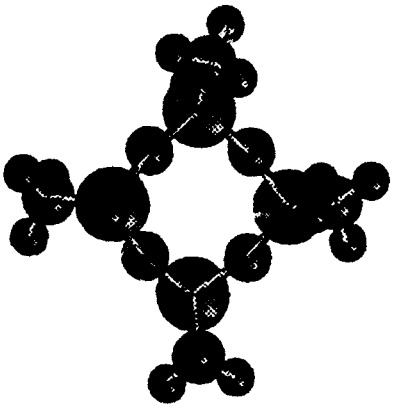


Figure 2. Linear siloxane dimer (top) and trimer (bottom) substituted with two biphenyl mesogens.



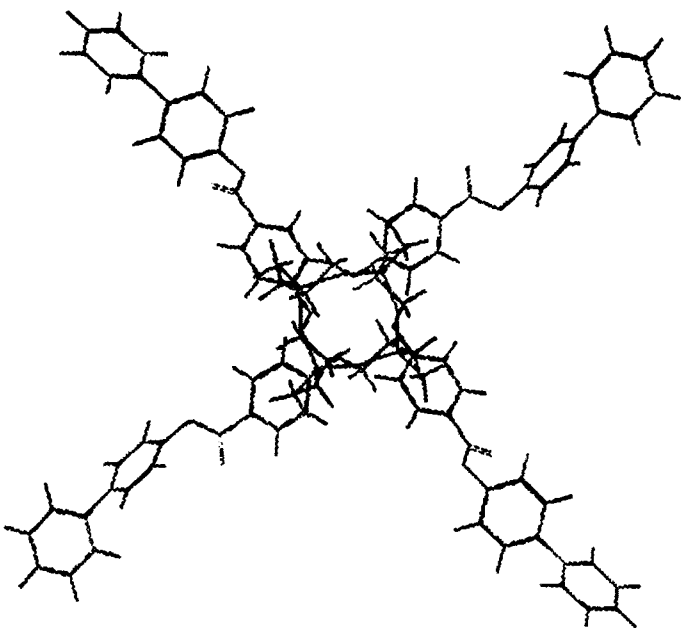


Figure 3a. Methyl-substituted cyclic siloxane having four -Si-O- groups in the ring.
b. Cyclic siloxane substituted with four biphenyl mesogens.

dissimilar mesogens fluctuated more widely with regard to inter-mesogen distances. Detailed analysis of the MD results on the effects of geometry and the various associations is currently in progress.

COMPUTATIONAL MODELING OF GRAPHITIC MATERIALS

Previous research on the mechanical properties of rigid-rod polymers, including poly(para-phenylene), leads naturally to studying other geometries available to hydrocarbon structures. The recent flurry of research activity focused on the structures and properties of Fullerenes, including the observation of their optical-limiting capabilities by scientists at MLPJ, provides added impetus for applying computational materials science techniques to such materials. MO calculations using MOPAC were undertaken to investigate the large scale deformation behavior of small aromatic-ring molecules such as benzene, naphthalene and pyrene. MO calculations were also used to compare the energies of linear and cyclic oligomers of polyacene. MD calculations were used to study the flexibility of graphitic cylinders and to compare their behavior with that of graphitic sheets.

Figure 4 shows orthogonal views of the minimum energy geometries of benzene, naphthalene and pyrene when the molecules were constrained to be bent 90 deg. For benzene, the bend was through the mid-points of bonds on opposite sides of the ring. For naphthalene and pyrene, bending was about the middle bond, causing the molecule to fold in butterfly fashion. The resistance to deformation implicit in the resulting geometries of benzene and pyrene is quite striking, and the tendency to retain planarity is obvious.

Figure 5 shows the energy versus bend angle for these three molecules. The bend energies are very high, and such deformations may exceed the validity of the MOPAC AM1 parameterization. Bending the pyrene molecule involves deformation of four rings and consequently is most energetically prohibitive. Further analysis will determine if the various types of single ring deformations contribute additiviely to the total deformation energy.

The plotted curves are least squares best fits to the calculated data of a second degree polynomial. It is interesting to note that for both naphthalene and pyrene, the curvature of the quadratic line is too large at small torsion angles. This indicates that the energy actually rises less rapidly than the quadratic potentials often used to represent such deviations from planarity.

The energies of linear and cyllic acenes were also calculated. Figure 6 shows the cyclic acene composed of 16 benzene rings and the linear acene having 18 rings. For cyclic and linear molecules having equal numbers of rings, the cycle has four less hydrogen atoms, two less carbon atoms and two more carbon-carbon aromatic bonds. The cycles also have consistently higher energies (heats of formation), as shown in figure 7. For linear acenes, the energy dependence, not surprisingly, is linearly dependent on the number of rings in the molecule. Such is clearly not the case for the cycles. For small cycles (6-12 rings), the energy is nearly independent of the number of rings, while for larger cycles (12-18 rings) the energy increases in fashion similar to the linear molecules. Extrapolation of just the small-ring data would suggest (incorrectly) that a cycle of 16 rings might have an energy comparable to that of the linear chain. The reason for the change in behavior

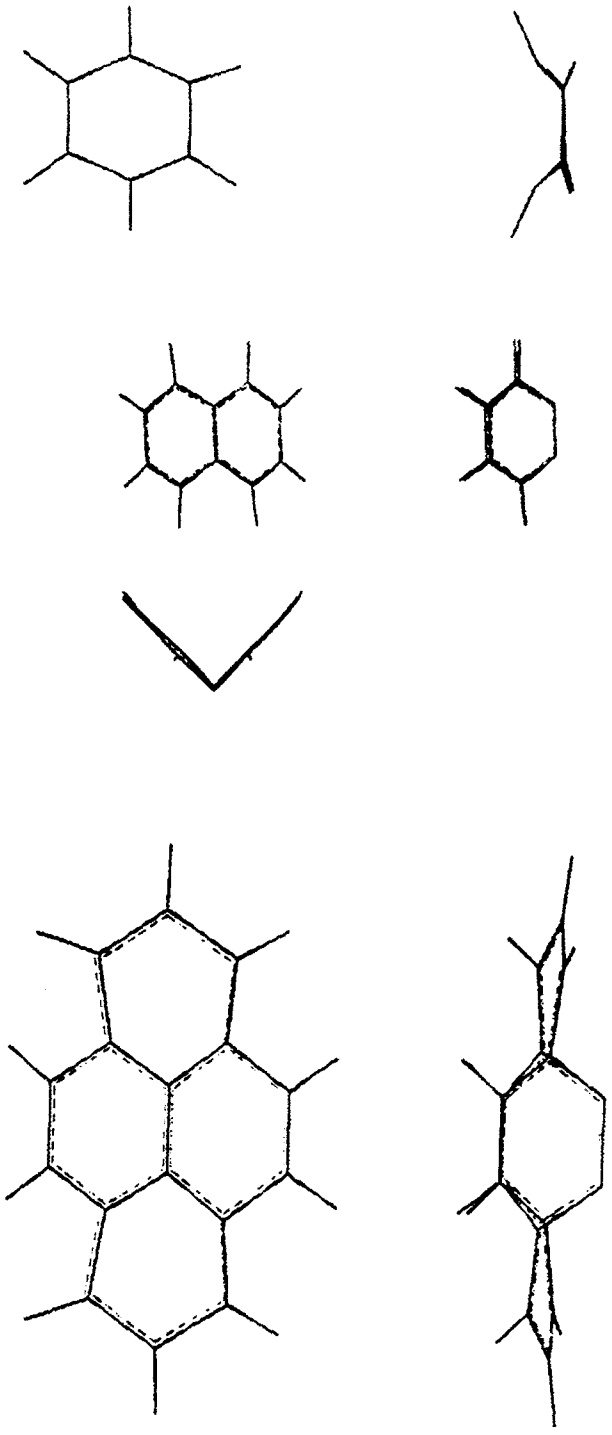


Figure 4. Orthogonal views of the minimum energy geometries for benzene, naphthalene and pyrene, constrained at 90 deg bends.

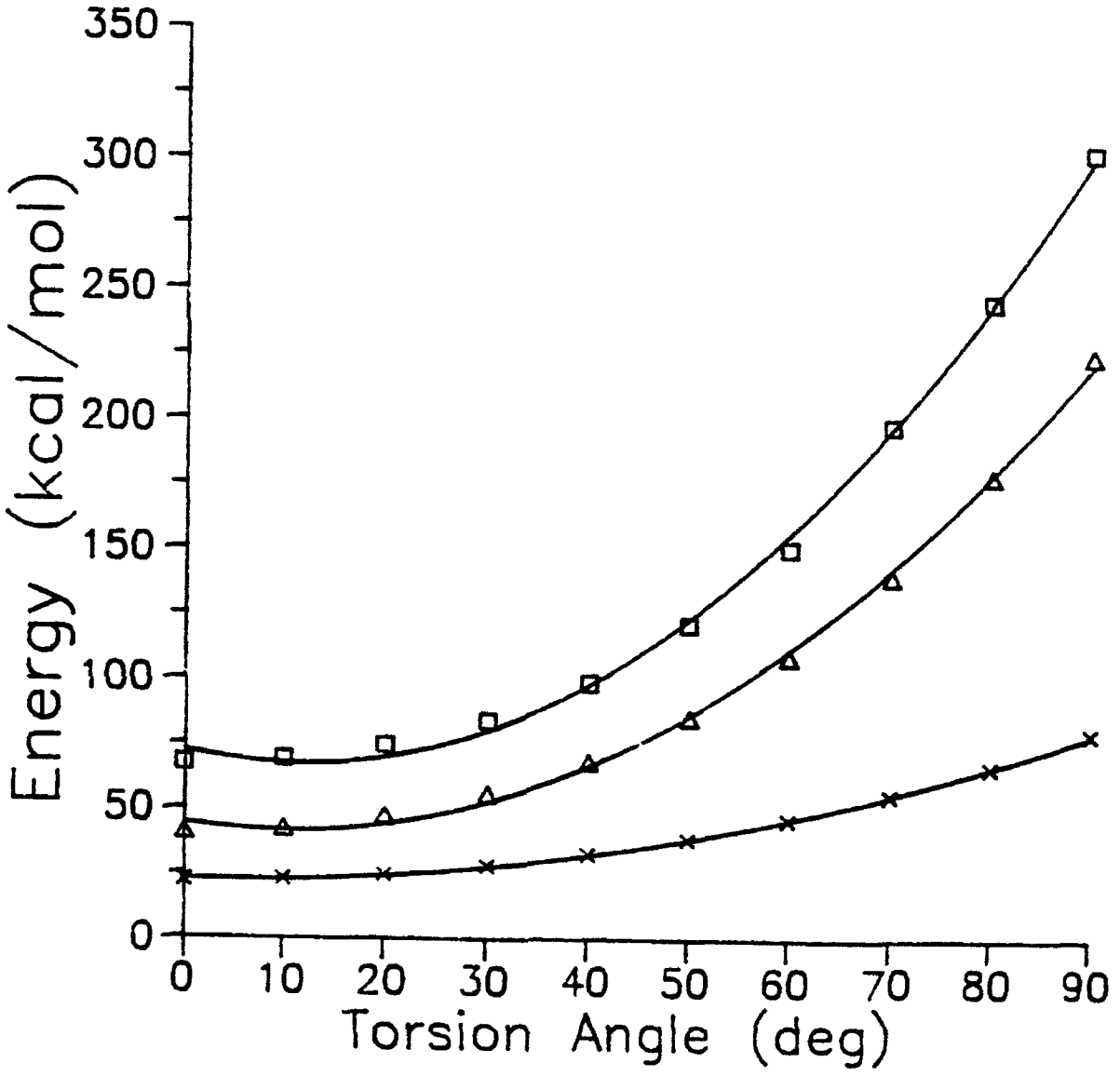


Figure 5. Energy versus bend angle for benzene (x), naphthalene (Δ), and pyrene (□).

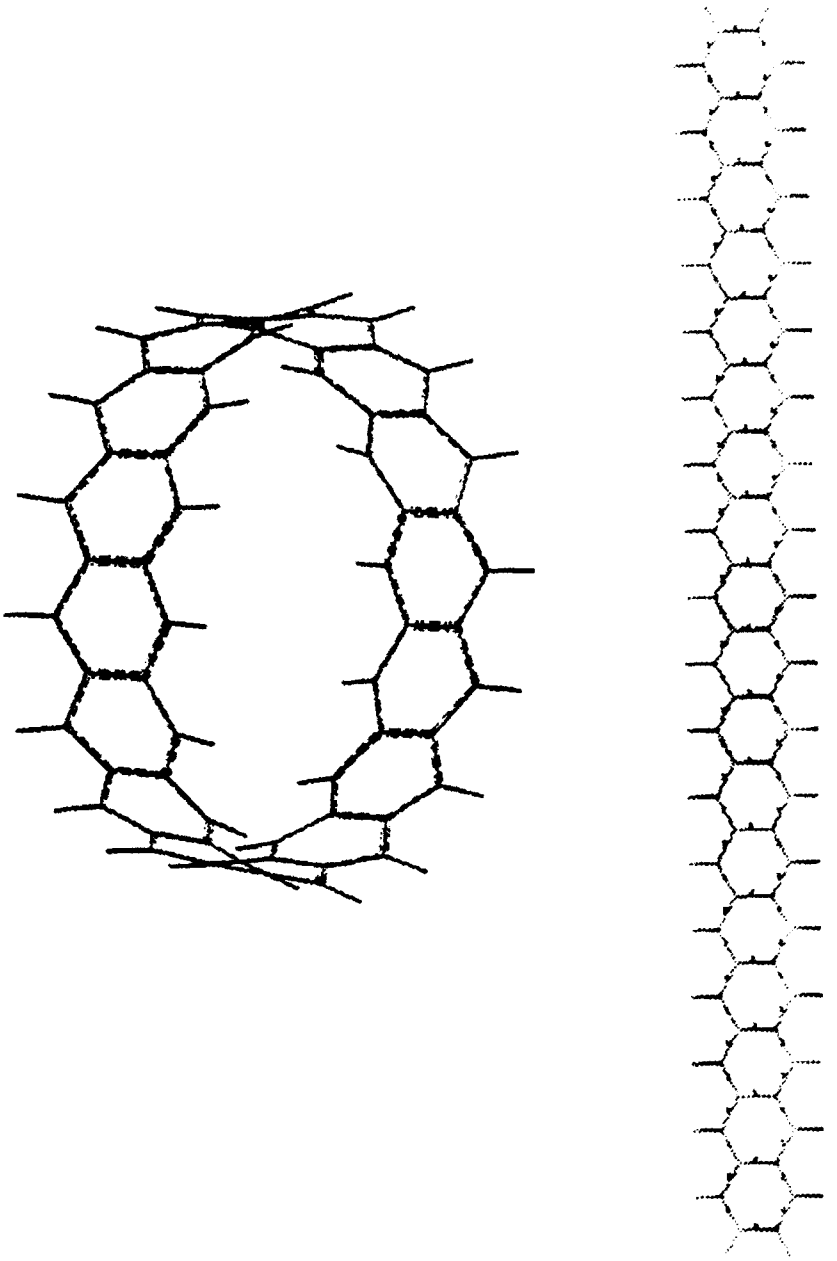


Figure 6 Cyclic acene consisting of 16 benzene rings and a linear acene having 18 benzene rings.

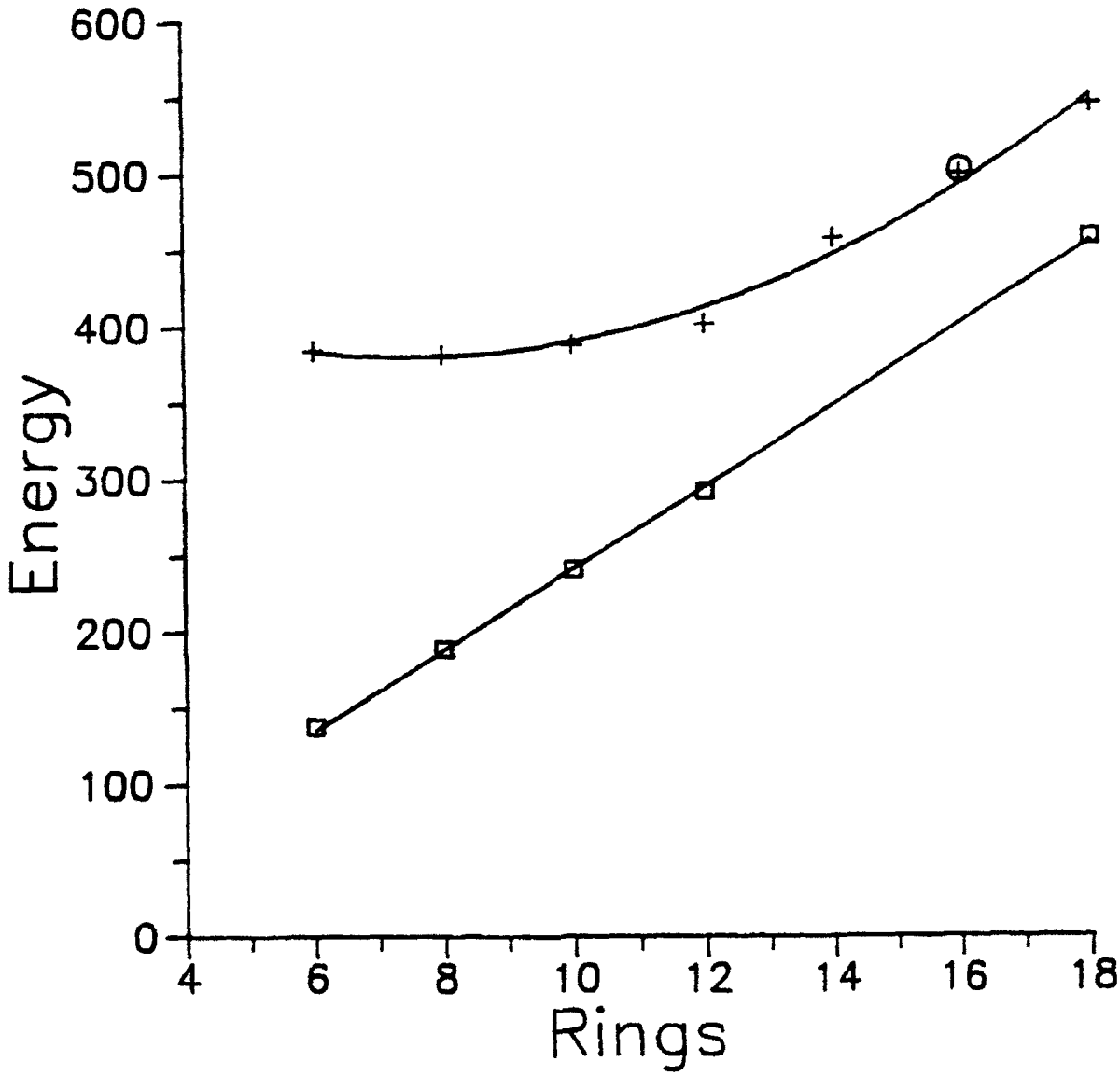


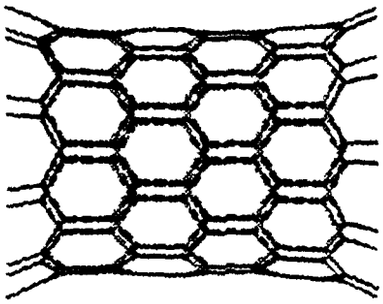
Figure 7 Energies of cyclic (+) and linear (□) acenes. The circle represents the result for a cycle whose starting geometry was an oval.

is not apparent. One possibility might be that interactions across the diameter of the cycles are important. These interactions are mandatory for small cycles, but for large cycles having a circular cross-section these would be lost. Energy minimization techniques often identify minima close to the starting geometries. To see if the cross-cycle interactions and this starting-point effect might be responsible for the change in the nature of the energy curve, an energy minimization was carried out (for a cycle of 16 rings) from an initial geometry having an oval shape. Virtually the same energy and geometry as that from a circular starting geometry was obtained, as indicated by the datum point represented by the circle.

Molecular dynamics simulations of cyclic and tubular molecules were performed. Figure 8a shows the tube having a length of four rings and a circumference of 10 rings, and figure 8b shows orthogonal views of a tube 16 rings in length and 10 rings circumference. Not surprisingly, the data show that the longer tube is very stiff indeed, with changes in its length less than 2% over the 100 ps dynamics trajectory. Further analysis and comparison with graphitic sheets are in progress.

MINI-SYMPOSIUM

The program for the two-day mini-symposium is attached as the Appendix Scientists from the University of Akron, University of Michigan, University of Virginia, Air Force Institute of Technology, Armstrong Lab and the MLPJ, MLPO, MLBP and MLIM Branches of the Materials Lab attended and presented papers. The symposium was structured to allow adequate discussion and coverage of Air Force interests and to bring the participants up-to-date on current activities



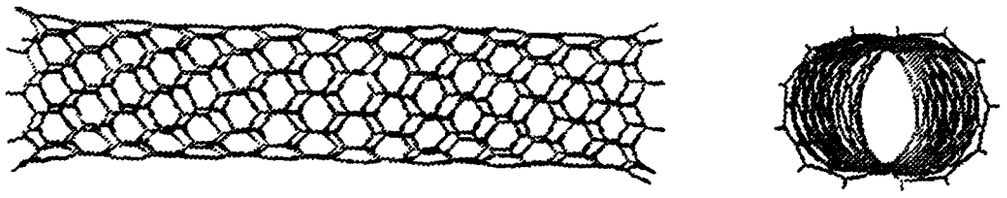


Figure 8 Graphitic tubes having a circumference of 10 benzene rings and a lengths of 4 and 16 rings.

in computer modeling relevant to laser hardening applications. The symposium was very well received, and planning is underway for a similar symposium in the summer of 1993.

From the perspective of intent of the RDL Summer Program, I believe the Mini-Symposium was especially valuable. It proved to be an excellent, condensed way to learn about the myriad research interests of the Air Force in areas relevant to my own interests. I believe it was equally valuable to the other participants.

MLPJ Mini-Symposium on Computational Materials Science for Laser Hardened Materials Research

15-16 July 1992

Bldg. 653 Cafetorium

PURPOSE:

Inform collaborators of scope of work and ensure complementary nature, reduce direct overlap of work while looking for uncovered subjects, produce cross fertilization of ideas, brainstorm coverage of relevant AF interests, critique our work, and provide students a chance to present current work to a "friendly" audience.

July 14

6:00 p.m.		Mixer: Trolley Stop

July 15

8:30-9:40 — **Introduction and Overview**

Chairman:	Dr. Wade Adams MLPJ	
8:30 - 8:50	Registration	
8:50 - 9:00	Dr. Wade Adams MLPJ	Welcome
9:00 - 9:20	Dr. Robert Crane MLPJ	Overview of Laser Hardened Materials
9:20 - 9:40	Dr. Ruth Pachter MLPJ	Overview of Computational Initiatives in DoD/DARPA/NSF Overview of Computational Projects at MLPJ
9:40 - 10:00		Coffee Break

10:00-12:00 — **Properties Calculations**

Chairman:	Dr. Barry Farmer U. Virginia	
10:00 - 10:20	Dr. Wade Adams	Prediction of Mechanical Properties
10:20 - 10:40	Dr. Peter Haaland AFIT	Properties of C_{60} Molecules
10:40 - 11:00	Dr. Tom Horn AFIT	Strain-Dependence of Normal Mode Frequencies: A "Biased-Hessian" Approach
11:00 - 11:20	Dr. Ruth Pachter	Prediction of Strain-Dependent Properties of Biopolymers
11:20 - 11:40	Mr. Philip Klunzinger U. Akron	Computations of Moduli and Coefficients of Thermal Expansion in Extended Chain Polymers
11:40 - 12:00	Dr. Doug Dudis MLBP	Cooperative Effects in Rigid-Rod Polymer Dynamics
12:00 - 13:30		Box Lunch at AFIT Park

13:30-15:30 — **Structure-to-Property Relationships**

Chairman:	Dr. Ruth Pachter	
13:30 - 13:50	Dr. Barry Farmer	Simulation of Star and Oligomeric Siloxanes
13:50 - 14:10	Dr. Soumya Patnaik MLPJ	Molecular Modeling of Siloxanes
14:10 - 14:30	Mr. Ed Socci U. Virginia	Calculations of Mesogen Interactions in a Cyclic Siloxane-Based Liquid Crystals
14:30 - 14:50	Dr. Harvey Paige MLBT	*Computational Chemistry in Liquid Lubricant Research*
14:50 - 15:10	Dr. Steven Pollack U. Cincinnati	Ab Initio Studies of Torsional Potentials in Mesogens Exhibiting Conformational Flexibility
15:10 - 15:30	Dr. Don Dorsey MLPO	Application of Computational Chemistry Techniques to Modeling of Semiconductor Thin-Film Epitaxy

15:30 - 17:30 — Coffee Break: **Poster Session**

18:00 — *Conference Dinner: China Cottage*

July 16

8:30-10:30 — **Biopolymers Structure Determination and Prediction**

Chairman:	Dr. David Martin U. Michigan	
8:30 - 8:50	Dr. Ruth Pachter	Aspects of Biomolecular Structure Determination
8:50 - 9:10	Mr. Ernie Fuentes MLPJ	Molecular Simulation of Beta-Sheet Polypeptides
9:10 - 9:30	Mr. Marc Martin MLPJ	Modeling of Polypeptide Bound Chromophores
9:30 - 9:50	Mr. David Mahoney U. Akron	Approaches to Spider Silk Modeling
9:50 - 10:10	Mr. Philip Anderson U. Michigan	Morphology and Modeling of Genetically Engineered Silk-Like Polypeptides
10:10 - 10:30	Mr. Steven Fairchild MLPJ	Protein Structure Prediction using Artificial Intelligence Neural Networks, and Modeling Techniques
10:30 - 10:50		Coffee Break

10:50-12:30		**Macromolecular Simulation**
Chairman:	Dr. Ronald Eby	
10.50 - 11:10	**Dr. David Martin**	Intermolecular Twist Defects in Extended-Chain Polymers
11:10 - 11:30	Ms. Patricia Wilson U. Michigan	Defects and Defect Modeling in Poly(diacetylenes)
11:30 - 11:50	Mr. Ken McTurk U. Akron	High Temperature Properties of PBZT Fibers
11:50 - 12:10	Mr. Rafael Ramirez U. Akron	Deformation of Fibers and Birefringence
12:10 - 12:30	Mr. Jaime Ojeda U. Michigan	Atomic Force Microscopy of Poly(imides) Near Surfaces
12:30 - 13:30		Lunch: Flying Pizza
13:30 - 15:10		**Algorithms and Other Modeling Applications**
Chairman:	Dr. Harvey Paige	
13.30 - 13.50	**Mr. Walt Johnson**	Modeling of Rugate Filters
13:50 - 14:10	Dr. Jim Lupo MLPJ	Nano-Crystalline Silicon: Simulation and Visualization
14.10 - 14.30	Dr. Garry Lamont AFIT	Parallel Computation
14.30 - 14.50	Mr. Chris Hasser Armstrong Lab Dr. Ruth Pachter	Modeling of Intermolecular Interactions using a Force Feedback Arm
14.50 - 15.10	Dr. Steven Leclair MLIM	The Application of Genetic Algorithms to Understanding Protein Structures
15.15		Coffee Break
15:30 - 17:00		**Round Table Discussion**
17:30		Farewell Dinner: Young's Dairy

NONLINEAR DYNAMICS AND CONTROL ISSUES FOR AEROELASTIC ENHANCEMENT USING PIEZOELECTRIC ACTUATORS

George T. Flowers
Assistant Professor
Department of Mechanical Engineering

Auburn University
201 Ross Hall
Auburn University, AL 36849–5341

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

Nonlinear Dynamics and Control Issues for Aeroelastic Enhancement Using Piezoelectric Actuators

George T. Flowers
Assistant Professor
Department of Mechanical Engineering
Auburn University

Abstract

A topic of current interest in the aeronautical community is the use of adaptive and 'smart' materials for health monitoring and performance enhancement for aircraft. One such material is piezoelectrics, which produce an electric current when subjected to strain and conversely is strained when an electric current is passed through it. As a result, such devices are being considered as actuators/sensors for a variety of applications. The goal of this investigation is to develop an understanding of issues relevant to the use of piezoelectric actuators/sensors in order to enhance the aeroelastic behavior of advanced aircraft.

Nonlinear Dynamics and Control Issues for Aeroelastic Enhancement Using Piezoelectric Actuators

George T. Flowers

Introduction

An important issue with regard to aeroelastic performance is wing/store flutter. The addition of wing stores serves to reduce aircraft flutter speeds and can lead to destructive vibrations. Military aircraft have a wide variety of possible wing stores. It has been recognized for some time that it is desirable to develop control schemes for the active suppression of flutter that will allow for many different store configurations in a convenient manner.

A flutter suppression strategy has been proposed that consists of using piezoelectric actuator/sensors as active elements in the store pylon. Each store could have a separate and independently controlled system (hence decentralized) that could respond to changing conditions (hence adaptive). The current work is aimed at understanding dynamics and control issues related to the implementation of such a strategy and evaluation of the performance of the resulting system. In addition, a better understanding of the impact of nonlinearities associated with piezoelectrics applied to aeroelastic systems is sought. The following sections describe relevant background material and discuss a research project aimed at investigating these issues.

Background

Most wings are designed to be free of flutter problems in the standard design configuration. Problems typically arise when non-design conditions are imposed upon the craft. This is particularly a problem during wartime when it is vital to prepare an aircraft for a mission as quickly as is possible and to include as much armament as possible. The addition of stores to a wing serves to alter the frequencies of the wing sufficiently so that a coalescense of the bending and torsional frequencies may occur due to aerodynamic effects, resulting in the unstable vibration mode known as bending-torsion flutter. Figure 1 illustrates the frequency coalescense behavior that marks the onset of this phenomenon. This type of flutter will be the focus of the current research effort.

A number of strategies have been tried with regard to the use of active controllers for gust alleviation and flutter suppression. They fall into three basic categories. The first is the use of aerodynamic control surfaces to provide the actuation forces. The most recent work is related to the AFW programs (AF/NASA/Rockwell) and (NASA/Rockwell). This work is well documented in the literature. Some sources for other information in this area are Newsom (1979), Mahash, et al, (1981), and Noll, et al, 1989.

The second control strategy consists of actuation supplied directly to the wing, particularly at the root. This strategy has recently been investigated with promising results by Heeg (1992). She investigated the application of piezoelectrics to flutter suppression. This work was a combined analytical/experimental study. The piezoelectrics were attached to the base of the wing to provide direct actuation.

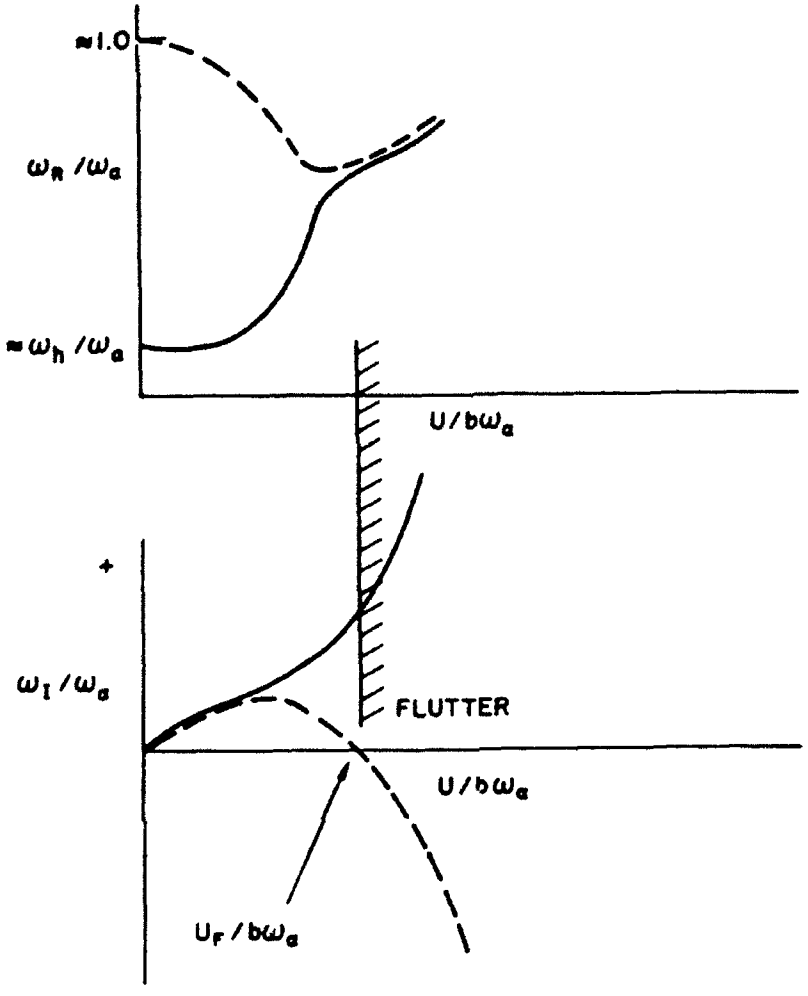


Figure 1: Real and imaginary parts of eigenvalues for bending–torsion flutter.

The current work will consider a much different approach for the suppression of flutter. The basic strategy was first studied by Triplett, et al. (1973). It consists of using active control techniques to decouple the dynamics of the store from those of the wing using an active pylon between the store and the wing. This approach has been implemented using hydraulic actuators with mixed results (Reed, et. al, 1980). The dynamic behavior of this scheme worked quite well. However, the hydraulic flow rates required were too high to be

practical.

An alternative to an active approach is a semi-active approach studied by Reed, et al. (1980). This approach used a passive spring/damper combination to isolate the store in pitch at the flutter frequency. The physical configuration is illustrated in Figure 2. Store alignment is maintained by an active control system. Good results were reported for this strategy. The expected qualitative performance of this system is shown in Figure 3. However, there were drawbacks. In particular, stiction of the passive damper was a significant problem.

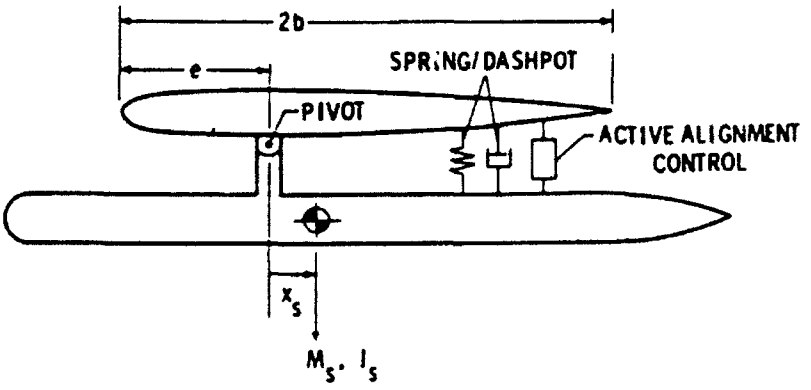


Figure 2: Decoupler pylon system [Reed, et al, (1980)]

'Smart' structure concepts have received considerable attention in recent years. An important area of that work is concerned with the use of piezoelectrics as actuators and sensors. Background material in this regard can be found in Crawley and de Luis (1987)

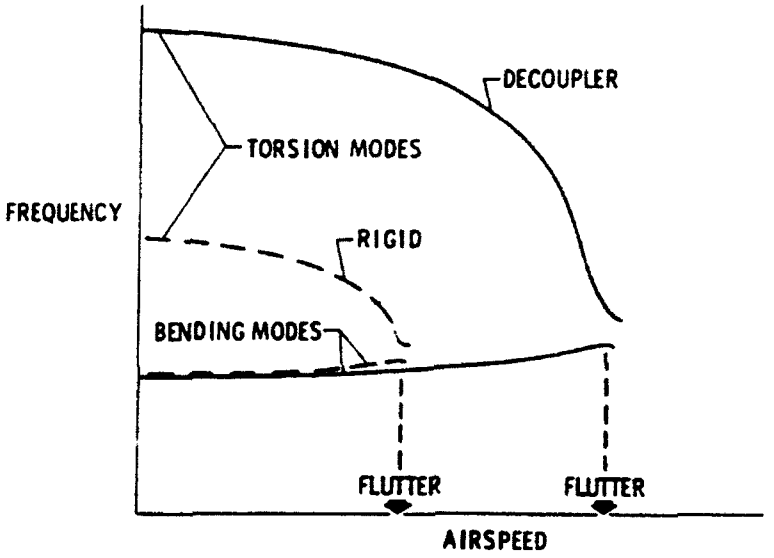


Figure 3: Decoupler pylon increases frequency separation of flutter critical modes [Reed, et al, (1980)]

which discussed the use of piezoelectric actuators in structural control and in Wada, et al., (1990) which presented a discussion of adaptive structures with piezoelectric actuators as an important component. Hagood, et al., (1990) presented a modelling approach for the dynamics of a piezoelectric actuator coupled with a structure. Anderson and Hagood (1992) discussed the use of piezoelectrics as self–sensing actuators for use in structural control. They presented a design for a piezoelectric truss member that would appear to be appropriate for use in the current work as an active pylon. The device is illustrated in

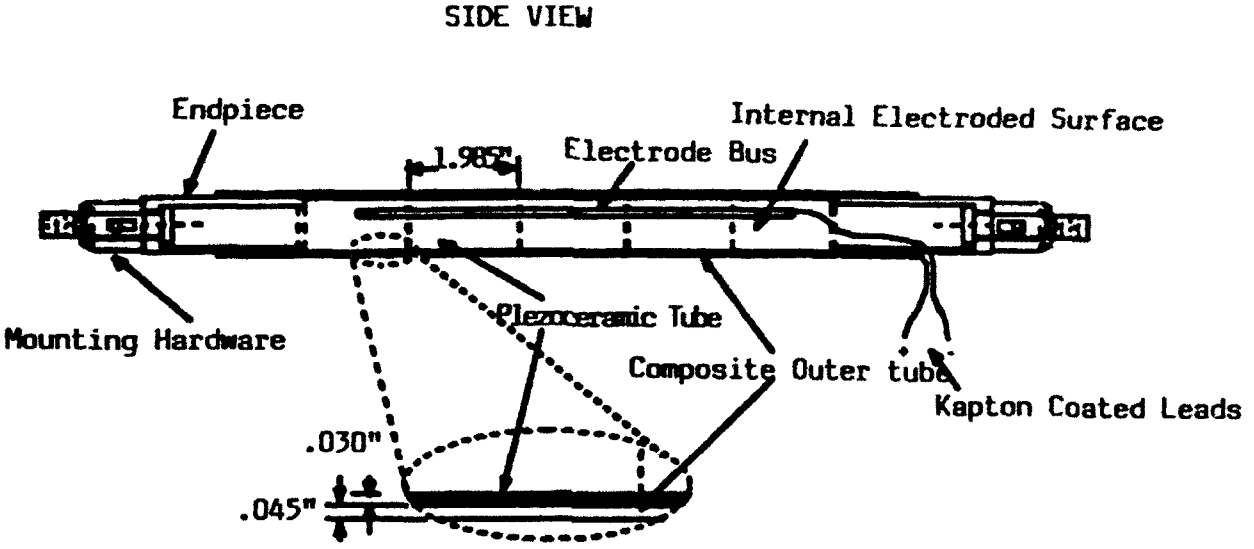


Figure 4: Piezoelectric truss member [Anderson and Hagood (1992)]

Figure 4.

Most piezoceramics studied for actuator use have been PZT. This material has primarily a linear force/strain relation within a limited total strain region. If strained beyond this region, the material ceases to be piezoelectric. Alternative materials are available with higher strains possible. However, the force/strain relations are nonlinear. PMN is a particular example of such material (Cross, 1992). The applicability of such materials as actuator elements in an active pylon and the implications of the associated nonlinerities need to be considered.

Available control strategies consist of two basic categories. They are fixed-gain and adaptive techniques. Adaptive control schemes offer several valuable advantages to the control system designer. Such approaches are typically more tolerant of modelling errors

and uncertainties. Under appropriate conditions, certain stability characteristics can be assured even though there may be a great deal of modelling uncertainty. In addition, adaptive strategies usually produce lower closed loop feedback gains and, thus, significantly improved noise rejection characteristics as compared to fixed–gain methods. The application of adaptive techniques to decentralized control is a topic of current interest and one that promises to produce important advances.

A number of investigators have made important contributions to the field of adaptive decentralized control. Among them are Hmamed and Radouane (1983), Ioannou and Kokotović (1983), Ioannou and Reed (1988), and Siljak (1991). Much of this work has been directed toward using model reference adaptive controllers (MRAC) for the decentralized control of unknown subsystems as if they were disconnected from each other. An important category of this work in decentralized control consists of direct control strategies for which the interconnection effects are directly reachable by the control actuators. One such approach has been developed by Gavel and Siljak (1989). They describe an adaptive method for designing a decentralized control strategy for a system with known subsystems that will be stable regardless of the magnitude of the interconnection forces.

The current work draws together flutter suppression, piezoelectric sensors/acutators, and adaptive decentralized control. The proposed study will build on the previous work in these areas. A simulation study of an active pylon flutter suppression strategy will be conducted. This study will incorporate several innovative concepts. First, it will use piezoelectrics as the actuation source for the active pylon. Second, a control strategy

will be employed that is aimed at using the vibration energy to provide the actuation for suppressing the vibration. However, this strategy is much more innovative than a simple attempt to passively damp out the vibration. Instead, an adaptive control scheme will be used to actively control the stiffness of the store pylon and raise the flutter frequency above the current speed range. Each store pylon will be controlled separately, thus the need for decentralized control. The basic issues to be addressed and the specific work to be done for such a project are detailed below.

Basic issues to be addressed

The proposed work will address the following basic issues with regard to aeroelastic enhancement using piezoelectrics.

1. Feasibility and effectiveness of frequency separation strategy and comparison with other control approaches. An innovative strategy for the control of bending–torsion flutter will be applied to the aeroelastic models. This strategy will consist of using the energy of the flutter mode to increase the bending stiffness of the wing and move the structure out of the flutter regime. Such a strategy is particularly exciting because of the electromechanical capabilities of the piezoelectric materials. Many implementation approaches are possible. A possible preliminary design concept is as follows. A piezoelectric film will be attached to the wing so that twisting and bending of the wing will cause straining of the material and a resulting electrical potential across two electrodes with a capacitor connected between them. The capacitor will serve as a storage cell for

the vibrational energy. This energy can then be called upon to power an actuator that will adjust the bending stiffness of the wing and move it away from the flutter regime.

2. Impact of piezoelectric nonlinearities on aeroelastic performance. The chief types of piezoelectric materials that are currently available consist of piezoelectric films and piezoceramics. The films are plastics and have the advantage that they can be subjected to large strains without losing their piezoelectric properties. The disadvantage is that the strength of such materials is relatively low. The piezoceramic materials are much more brittle and subject to fracture. However, they can produce relatively large actuation strengths, albeit for very low strains. The primary piezoeceramic that is currently used is lead zirconium titanate (PZT), which has been studied extensively in investigations of piezoelectric materials. An alternative material, lead magnesium niobate (*PMN*), is currently being investigated by researchers at Pennsylvania State University. This material offers the advantage over PZT in that it has considerably more potential maximum strain. The disadvantage is that the material has significant nonlinear characteristics that must be considered. An important goal of the suggested work is to incorporate the nonlinear effect of such materials into the aeroelastic model and evaluate the impact of such nonlinearities on dynamical behavior and control characteristics.

3. Effectiveness of decentralized, adaptive control strategies. The proposed frequency separation strategy lends itself naturally to the use of decentralized control techniques. Such an approach would not require communication between the various controllers

(presumably one for each store) and eliminate the need for hardwired connections between them. The controller algorithms that will be developed for this study will be based around an adaptive, decentralized control formulated by Gavel and Siljak (1989). This strategy is a model reference adaptive controller designed for decentralized systems with uncertain and nonlinear interconnections.

Specific Suggested Research Task

The specific items that should be accomplished in the proposed research are listed below.

I. Modelling

Models should be developed that simulate the dynamics of a blade subjected to bending–torsion flutter and the dynamics of the proposed piezoelectric sensor/actuator configuration. The wing models should consist of two types. The first is a simple typical section model similar to those used in many qualitative studies [(Dowell (1989), Heeg (1991)]. The second should be a bit more sophisticated. It should consist of a beam subjected to aerodynamics obtained from the Wright Laboratory simulation program called ADAM [Noll (1987)].

The data for the model of the piezoelectric actuator and sensor can be obtained from models developed by Hagood (1990). Strain as a function of electrical properties for PZT and PMN are available from a number of sources. Cross (1992) has data for PMN.

A suggested controller strategy consists of an adaptive decentralized strategy based upon the approach developed by Gavel and Siljak (1989). A number of stores are typically

attached to each wing. The controllers for each store could be independent of the others.

II. Simulation Studies.

The complete models can be used as testbeds for simulation studies aimed at investigating issues related to the proposed control strategy. The specific studies are detailed below.

1. Develop implementation strategies and study their effectiveness for flutter suppression. An effective technique for the suppression of bending–torsion flutter is to raise the bending frequency and eliminate the frequency coalescence in the operating speed range. The proposed strategy is aimed at supplying the energy for this approach using the energy of the 'unstable' vibration being controlled. The result could be truly decentralized and low maintenance controller designs. Important concerns need to be addressed with regard to the feasibility of this strategy and the method(s) of implementation. This study will examine various approaches in this regard and evaluate their effectiveness.

2. Study control structure interaction. Concerns are the interaction between the dynamics of the structure and the dynamics of the actuators. Of particular concern is the impact of nonlinearities associated with the piezoelectric materials (Heeg, 1992). Reliability is also of special concern when dealing with the catastrophic behavior that may result from the onset of flutter. The proposed strategy consists of multiple actuators and controllers. An important problem is concerned with understanding the possible responses if one or more of the actuators or controllers fails.

Significance and benefits of the proposed work

The proposed research effort is meant to develop knowledge in the area of aeroelastic enhancement using piezoelectric and other 'smart' materials. The work will investigate the potential of the proposed strategies and point out possible benefits and areas where problems may occur. The use of piezoelectric actuators in aircraft has many obvious potential benefits, some of which are:

(1.) Adequate control power with lighter weight;

(2.) Truly collocated sensing and actuation;

(3.) Wider bandwidth than hydraulics; and

(4.) Smart pylons with independent controllers; self contained and possibly powered by the aeroelastic vibration phenomena that is being controlled.

There is excellent potential for significant contributions in the work described above and for future work in this area. Some suggestions for further areas of investigation that are natural extensions of the proposed research effort are discussed below.

1. **Perform simulation studies with more sophisticated models.** The models that have been suggested for the proposed study are relatively simple. More complex structural and aerodynamic models should be simulated using the proposed flutter suppression strategy in order to more accurately assess its performance.

2. **Study suppression strategies for other types of flutter and for gust alleviation.** Bending–torsion flutter is the simplest type of flutter to model and is the behavior that has be focused upon by the current study. However, other types of flutter exist and can have serious consequences for aircraft performance. In addition,

gust loadings can be a serious problem for civilian as well as military aircraft. Important benefits in terms of aircraft performance and passenger safety may result from extending the proposed control techniques to these areas.

3. **Perform experimental tests to study the nonlinear dynamical characteristics of piezoceramics and other appropriate piezoelectric materials.** PMN may have significant nonlinear characteristics at high electric field levels and may exhibit hysteresis effects. In addition, aging and temperature may alter physical properties. Figures 5, 6 and 7 illustrate some of these effects. Such properties can have a significant impact on dynamical behavior and need to be better understood if acceptable implementations for aeroelastic control are to be developed.

4. **Experimentally validate the simulation work.** Simulation models can provide very useful tools for the evaluation of dynamical behavior and the effectiveness of control strategies. However, it is vitally important that the results be tested with reasonable experimental models. Otherwise, important dynamical behavior may remain unmodelled and the conclusions of the simulation studies will be suspect. An appropriate experimental program would include the construction of a scaled wing model with a piezoelectric pylon and stores attached. Wind–tunnel tests using this model should be conducted to evaluate the performance of the proposed control strategy and of the strut type actuator design.

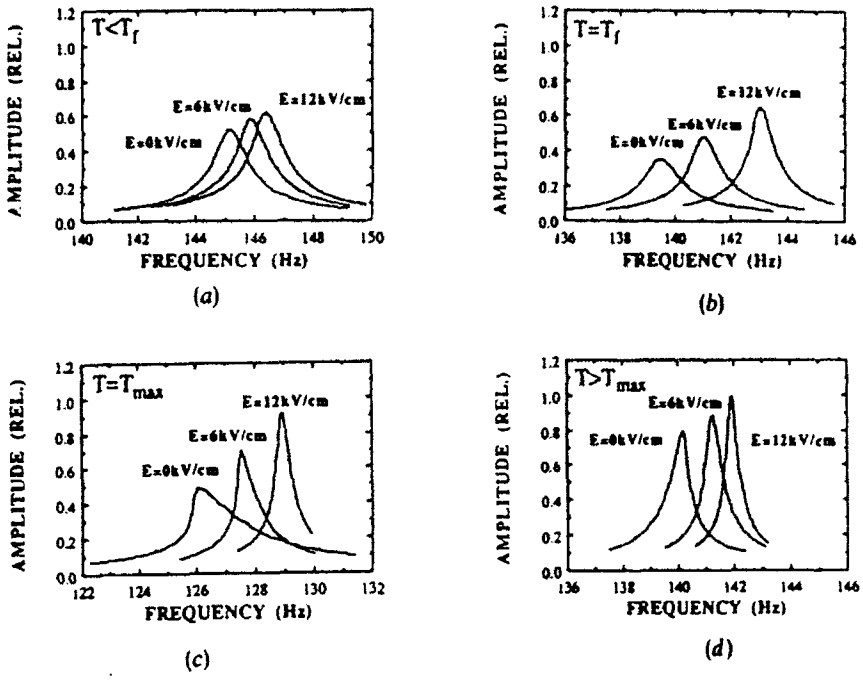


Figure 5: Sample Resonance Curves for PMN–10PT (a) T=-15 deg C, (b) T=15 deg C, (c) T=40 deg C, (d) T=65 deg C [Viehland, et al., (1991)]

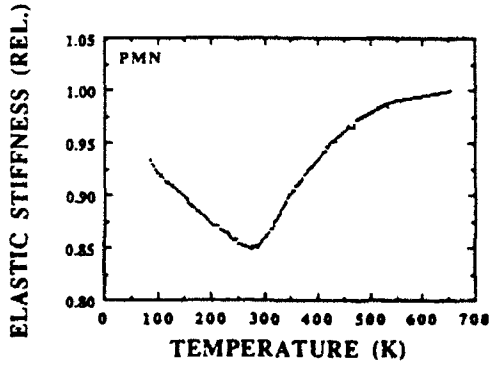


Figure 6: Elastic Stiffness as a Function of Temperature for PMN [Viehland, et al., (1991)]

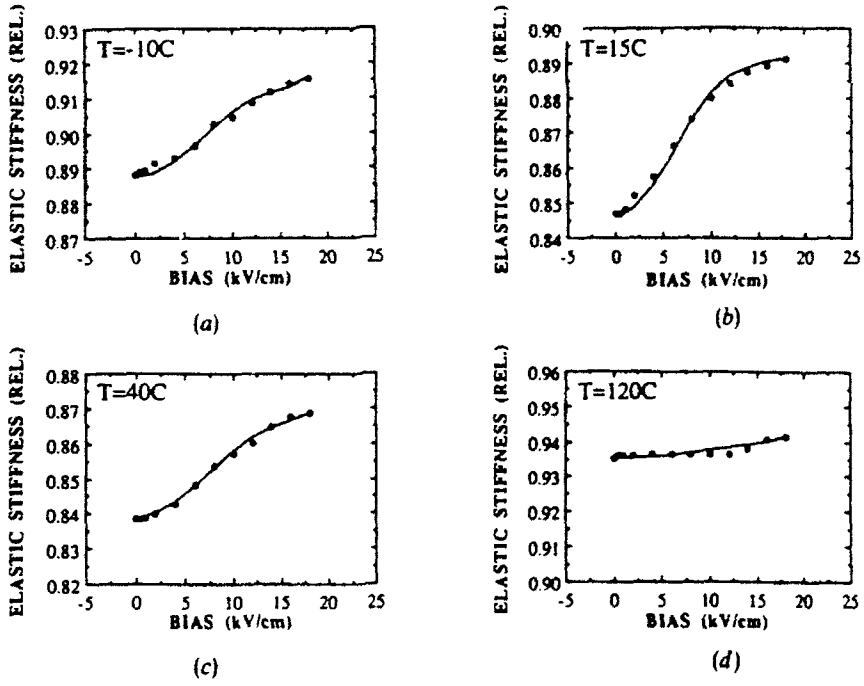


Figure 7: Elastic Stiffness as a Function of D.C. Electrical Bias for PMN–10PT (a) T=-10 deg C, (b) T=15 deg C, (c) T=40 deg C, (d) T=120 deg C [Viehland, et al., (1991)]

References

Anderson, E.H., Hagood, N.W., amd Goodloffe, J.M., *Self-Sensing Piezoelectric Actuation: Analysis and Application to Controlled Structures*, AIAA paper no. 92–2465.

Cagle, A., and Ozguner, U., *Optimal Decentralized Feedback Control for a Struss Structure*, Proceedings of the AIAA Guidance, Navigation, and Control Conference, Boston, MA, August, 1989, pp. 1199–1206.

Craig, J.J., *Adaptive Control of Mechanical Manipulators*, Addison–Wesley Publishing Company, Reading, MA, 1988.

Crawley, E.F., amd de Luis, J., *Use of Piezoelectric Actuators as Elements of Intelligent Structures*, AIAA Journal, Vol. 25, No. 110, 1987, pp. 1373–1385.

Derecho, A.T., Freskakis, G.N., and Fintel M., *A Study of the Effects of the Frequency Characteristics of Ground Motion on Nonlinear Structural Response*, Proceedings of the International Symposium on Earthquake Structural Engineering, St. Louis,

Missouri, USA, August 1976, pp. 21–36.

Dix, P., Ozguner, U., and Gordon, R., *Decentralized Experiments on a Truss Structure*, Conference in Decision and Control, Honolulu, Hawaii, Dec., 1990.

Gavel, D.T., and Siljak, D.D., *Decentralized Adaptive Control: Structural Conditions for Stability*, IEEE Transactions on Automatic Control, Vol. 34, No. 4, April 1989, pp 413–426.

Hagood, N. W., and Crawley, E. F., *Experimental Investigation into Passive Damping Enhancement for Space Structures*, AIAA paper no. 89–3436.

Heeg, J., "An Analytical and Experimental Investigation of Flutter Suppression Via Piezoelectric Actuation," *Proceedings of the 33rd AIAA/ ASME/ ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, pp. 237—243, AIAA paper no. 92–2106–CP.

Hmamed, A., and L. Radouane, *Decentralized Nonlinear Adaptive Feedback Stabilization of Large-scale Interconnected Systems*, IEE Proceedings, Vol. 130-D, 1983 PP. 57–62.

Ioannou, P.A., and P.V. Kokotović, Adaptive Systems with Reduced Models, Springer, New York, 1983.

Ioannou, P.A., and P.V. Kokotović, *Robust Design of Adaptive Control*, IEEE Transactions on Automatic Control, Vol. AC–29, 1984, pp. 202–211.

Ioannou, P.S., *Decentralized Adaptive Control of Interconnected Systems*, IEEE Transactions on Automatic Control, Vol. AC–31, No. 4, April, 1986, pp. 291–298.

Ioannou, P.A., and J.S. Reed, *Discrete-time Decentralized Control*, Automatica, Vol. 24, 1988, pp. 419–421.

Kailath, T. Linear Systems, Prentice–Hall, Inc., 1980.

Lewis, F.L., Optimal Estimation, John Wiley and Sons, New York, 1986.

Mao, Cheng–Jyl, and Lin, Wei–Song, *On–Line Decentralized Control of Interconnected Systems With Unmodelled Nonlinearity and Interaction*, Proceedings of the 1989 Automatic Controls Conference, Vol. 1, pp. 248–252.

Mahesh, J. K., Stone, C. R., Garrard, W. L., and Dunn, H. J., *Control Law Synthesis for Flutter Suppression Using Linear Quadratic Guassian Theory*, Journal of Guidance

and Control, Vol. 4, No. 4, 1981.

Newsom, J.R., *Control Law Synthesis for Active Flutter Suppression Using Optimal Control Theory*, Journal of Guidance and Control, Vol. 5, No. 4, 1982.

Noll, T., Perry, B. III, Tiffany, S., and others "Aeroservoelastic Wind–Tunne Investigations Using the Active Flexible Wing Model – Staus and Recent Accomplishments," *Proceedings of the 30th AIAA/ ASME/ ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, pp. 68–79, AIAA paper no. 89–1168–CP.

Ozguner, U., Yurkovich, S., and Gordon, R., *A Case Study in Control Design and Implementation for Flexible Structure Systems*, Proceedings of the AIAA Dynamic Specialists Conference, April, 1990, pp. 322–333.

Ozguner, U., Khorrami, F., and Iftar, A., *Two controller design approaches for decentralized systems*, Proceedings of the AIAA Guidance, Navigation, and Control Conference, Minneapolis, MN., Aug. 1988, pp. 237–244.

Rao, S.S., Pan, T.S., and Venkayya, V.B., *Modeling, control, and design of flexible structures: A survey*, Applied Mechanics Review, vol. 43, no. 5, part 1, may 1990, pp. 99–117.

Reed, W. H. III, Foughner, J.T. Jr., and Runyan, H.L. Jr., *Decoupler Pylon: A Simple, Effective Wing/Store Flutter Suppressor*, Journal of Aircraft, Vol. 17, No. 3, 1980, pp. 206–211.

Sandell, N.R., Varaiya, P., Athans, and Safonov, M.G., *Survey of Decentralized Control Methods for Large Scale ystems*, IEEE Transactions of Automatical Control, Vol. AC–23, No. 2, April 1978, pp. 108–128.

Siljak, D.D., *Parameter Space Methods for Robust Control Design: A Guided Tour*, IEEE Transactions on Automatic Control, Vol. 34, No. 7, July 1988, pp. 674–688.

Siljak, D.D., Decentralized Control of Complex Systems, Academic Press, Inc., 1991.

Singh, M.D., Decentralised Control, North–Holland Publishing Company, Amsterdam, 1981.

Triplett,, W.E., Kappus, H.P.F., and Landy, R.J., *Active Flutter Suppression Systems for Military Aircraft: A Feasibility Study*, AFFDL–TR–72–126, February, 1973.

Yang, C.Y., Random Vibration of Structures, John Wiley and Sons, New York, 1986.

Young, K.D., *A Distributed Finite Element Modeling and Control Approach for Large Flexible Structures*, AIAA Guidance, Navigation, and Control Conference, Minneapolis, MN, Aug. 1988, AIAA, pp. 253–263.

Viehland, D., Jang, S.J., Cross, E., and Wuttig, M., *Anelastic Relaxation and Internal Strain in Lead Magnesium Niobate Relaxors*, Philosophical Magazine A, Vol. 64, No. 4, 1991, pp. 835–849.

Wada, B.K., Fanson, J.L., and Crawley, E.F., *Adaptive Structures*, Proceedings of the ASME Winter Annual Meeting, December, 1990.

Witenhausen, H.S., *A Counter Example in Stochastic Optimal Control*, SIAM Journal of Control, No. 6, 1968, pp. 131–147.

PASSIVE RANGING AND ROLL-ANGLE APPROXIMATION FOR FUZE APPLICATION

Simon Y. Foo
Assistant Professor
Department of Electrical Engineering
FAMU/FSU College of Engineering

Florida State University
Tallahassee, FL 32306

Final Report for:
AFOSR Summer Faculty Research Program
Wright Laboratory/MNMF (Fuze Branch)

Sponsored by:
Eglin Air Force Base, FL

August 1992

PASSIVE RANGING AND ROLL-ANGLE APPROXIMATION FOR FUZE APPLICATION

Simon Y. Foo
Assistant Professor
Department of Electrical Engineering
Florida State University

Abstract

This research presents a stereo vision approach to passive ranging and roll-angle approximation using an array of image sensors. Specifically, a divergent multisensor imaging system for fuze application is presented. Although there may be a loss in accuracy due to small baselines of sensors and potential mismatches in pixel registration, this method is computationally quick and simple. Therefore this stereo vision technique is most suitable for fuze application where speed is critical and accuracy is not so important. A method for determining the roll-angle from divergent imaging sensors is also presented.

PASSIVE RANGING AND ROLL-ANGLE APPROXIMATION FOR FUZE APPLICATION

Simon Y. Foo

INTRODUCTION

There are numerous situations, particularly in military systems, where a fast moving observer (e.g., a projectile) with passive imaging sensors wishes to determine the location of a stationary or moving target. For example, in passive airborne automatic acquisition systems, it may be necessary to determine range to a candidate target for use in the target classification algorithms. Currently there are three basic approaches to passive ranging: stereo vision or triangulation methods, optical flow methods, and methods based on measuring the field strength of energy transmission.

Optical flow refers to the apparent motion of objects in the image plane due to either the target or the observer camera motion. Assuming all optical flow is produced by the motion of the camera (observer), the image of a single point on a stationary or near stationary target appears to move to the motion of camera. The amount a specific point on the image plane appears to move depends on (i) speed of camera, (ii) distance of point, and (iii) how far the point lies off the velocity vector of observer. It is obvious that a point normal to the velocity vector of observer does not appear to move. (This is the basis of a collision avoidance rule: watch out for objects which appears stationary in the line of sight!). For a "flat" scene (as seen from a forward-looking infrared sensor) the optical flow method utilizing mission parameters by Dvornychenko [1] would be appropriate for depth calculation. However, in our fuze application the target is assumed to be very close to a fast moving camera. Hence there is not much time for the camera to take successive pictures and determine the focus of expansion which is crucial in calculating the position along the observer's path at which the target is closest.

It is also possible to relate distance with the measurements of strength of energy transmission, as shown by Appling[2]. It is well known that the field strength of energy transmitted from a location, as measured at a second location, is related to the distance between the two locations by the equation $P = k/R^2$, where P is the power density, R is the relative distance between the two locations, and k is the coefficient of proportionality including effects of antenna gains, transmitted power level, receiver sensitivity, atmospheric effects, etc. The problem with this approach is that the constant k may not be precisely determined.

This study investigates a stereo vision approach to extracting the range of target and roll-angle information from imaging sensors. Specifically, a divergent multisensor imaging system for tracking a target is presented. The lower accuracy due to small baselines of sensors and potential offsets in pixel registration can be tolerated in our fuze application. The advantage of this stereo vision method is that it only requires few simple calculations, and therefore most suitable for real-time applications where speed is critical. In addition to range and roll-angle calculations, a neural network classifier is used to classify the segment of target in the field of view (FOV).

ARCHITECTURE OF DIVERGENT MULTISENSOR IMAGING SYSTEM

In this project we assumed the target is *within* range of observer but not necessary located in the observer's direction of motion. Furthermore the target is assumed to be stationary (at least relative to the fast approaching observer). Assuming the velocity of observer is 1000 meter per second and the scanning rate of camera is 1 frame per millisecond (which means that a 1-meter displacement occurs in 1 millisecond), then it is necessary to process each frame in $\leq$ 1 millisecond. Therefore an important consideration when choosing an algorithm or method for range and roll-angle calculation is that it must perform *very quickly.*

Figure 1 shows two possible placements of imaging sensors on a convex surface. Configuration (a) with a 2nd row of sensors will only provide redundant information, since at an oblique angle, the column pair sensors appear as a single sensor. In other words, both sensors in each column shares approximately the same point of view. In contrast, the adjacent sensors in each row of configuration (b) are focussed at different points of view to cover a wider area of a close-by target.

There are basically two scenarios in the encounter between the observer (imaging sensor array) and the target. The first case is when the velocity vector of observer is normal to the centroid of target (i.e., the observer is heading right into the center of target). In this case, the target may be visible on all 8 sensors (depending on size of target) and the roll-angle is insignificant. The other scenario is when the observer flies over, or under, or by the side of target. In this case, the target is only visible on maybe 2 or 3 adjacent sensors (again, depending on size of target). Assuming each wide-angle lens can provide a 100° FOV, a 100x100-element CCD or infrared focal plane array to implement each sensor is suggested, such that each image pixel represents a 1° angle of view. Figure 2 shows the corresponding set of 2-dimensional maps for passive ranging and roll-angle calculation based on our choice of configuration (b). The roll-angle is the angular position where the centroid of target is located.

LOW-LEVEL IMAGE PROCESSING

The low-level image processing steps prior to feature classification, roll-angle calculation, and passive ranging, are as follows:

1(a). Noise Cleaning/Suppression

An image may be subject to noise from several sources such as electrical sensor noise. A simple noise-cleaning method is to examine each pixel sequentially, and if the magnitude of a pixel is greater than the average intensity of its immediate neighbors by some threshold, it is replaced by the average value.

Median filtering (developed by Tukey, 1971) is a useful technique for noise suppression in images. In one-dimensional form, the median filter consists of a sliding window encompassing an odd number of pixels. The center pixel in the window is replaced by the median of the pixels within the window. The median of a discrete sequence $a_1, a_2, \ldots, a_N$ for N odd is the member of the sequence for which $(N-1)/2$ elements are smaller or equal in value, and $(N-1)/2$ elements are larger or equal in value. It has been shown that the median filter is much more effective in reducing the effect of discrete impulse noise than smoothly generated noise. Furthermore, median filtering should be used with care so that there is no significant signal loss. Median filtering is particularly useful in heavily-cluttered images such as an infrared image.

1(b). Geometric Distortion Correction

Perspective distortion encountered in most imaging systems can be corrected by fitting a polynomial curve to each distorted grid line on the fish-eye view, Figure 3(a), and then computing the inverse correction functions for each cell. This process is called spatial warping. The extended object viewed from an oblique angle can be warped so that the object appears as if it is viewed at a normal angle. Corrective warping may be considered a two-stage process. First, the coordinate of the distorted image point (p, q) that maps into the corresponding ideal image point (j, k) of Figure 3(b) is determined. Next, the intensity value of pixel at point (p, q) is estimated by interpolation of the intensity levels of surrounding pixels.

1(c). Edge Enhancement

Since there is blurring of image due to a fast moving camera, edges of objects need to be enhanced. Edge enhancement or crispening is performed by convolving the image with high-pass masks such as:

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} \qquad \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix} \qquad \begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

These edge enhancement masks have the property that the sum of their elements is unity. A Weiner filter can also be applied to remove the blurring effects due to a fast moving object, fast moving camera (as in our case), or atmospheric turbulence.

2. Target Acquisition

A traditional approach to the problem of finding the location of an object or part of an object in a given FOV has relied on the use of an intensity histogram to determine a threshold level, and using this threshold to locate and separate the object from the background. One problem with this simple approach is that the background intensities may vary significantly throughout the FOV and portions may exceed the chosen threshold, giving rise to false candidate targets. Another problem is that in some situations the background and target intensities are not well resolved, as in the case of optical blur due to motion of camera. Thus it may be difficult to determine a reasonable threshold level on the basis of the intensity histogram alone. One solution is the size-contrast filter approach by Dudani et al [3] which uses two windows on an approximate target location. Other popular target detection algorithms include the "Superslice" (developed at the University of Maryland), "Spoke Filter" (U. S. Army Missile Command, Hunstville), "contrast box" (Texas Instruments), and double-gated contrast filter (Ford Aerospace). The performance of these algorithms can be found in a report by Burton and Benning [4].

3. Target Extraction (Segmentation)

Once an object is acquired in the FOV, the next step is to extract (separate) it from the background so that the shape of the object becomes apparent. Feature extraction or segmentation is a process that finds boundaries of meaningful objects in an image. One of the simplest approaches to target extraction is to use an intensity threshold for assigning pixels to either the target or the background. Another approach is the intensity-gradient based approach by Dudani et al [3] which uses both the intensity histogram and gradient information to select a threshold to separate the feature of interest from the background.

4. Edge Detection

Edge detection is performed to generate an edge map of the object for classification. There are two popular gradient operators for edge detection, namely the Prewitt operator and the Sobel operator, defined as follows:

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$ Prewitt edge masks

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$ Sobel edge masks

To locate diagonal edges, the following edge masks are used:

$$\begin{bmatrix} -1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}$$

Notice the gradient masks have zero weighting (the sum of array elements is zero) so that there is zero output response over constant luminance regions of image.

Edge-thinning is then used to remove undesired edge broadness in gradient images while retaining edge continuity of the image. This process is required so that an edge map of the target can be generated for target classification purposes. Keshavan and Lo [5] performed a comparative study of some popular edge-thinning algorithms, namely the non-maximum suppression technique, modified USC thinning algorithm, and Reeves' algorithm.

NEURAL NETWORK CLASSIFIER

Current approaches to target identification include statistical (Bayesian) classifiers and neural network classifiers. One of the most popular neural network architecture for implementing a classifier is the multilayer feedforward network (or multilayer Perceptrons) which consists of sets of nodes (or neurons) arranged in layers. The network is "trained" to recognize objects through a learning algorithm called the backpropagation algorithm. The backpropagation algorithm, formulated and popularized by Rumelhart et al [6] is basically a gradient descent method that minimizes the error between an actual output and the desired

output. For example, a 3-layer network, shown in Figure 4, consists of an input layer, a hidden layer, and an output layer. The outputs of nodes in one layer are transmitted to nodes in the following layer through weighted links that amplify or inhibit such outputs. Except for the input layer nodes, the net input to each node is the sum of the weighted outputs of the nodes in the previous layer. Each node is activated according to a nonlinear activation function (typically a sigmoid function) which takes into account the sum of the weighted inputs and the bias of the node. In the training phase of the network, a vector or pattern is presented as input and the network uses a learning rule to adjust the set of weights in all connecting links and the biases of all nodes such that the desired output(s) is obtained at the output node(s). In the testing phase, the network calculates the output(s) of the network using information from the input vector, biases at each node, and the stored weights between each layer. The resulting activated output node denotes the class the object belongs.

ROLL-ANGLE DETERMINATION

For an arbitrary shape represented by a region $\Re$ containing N pixels, the center of mass $(\bar{m},\bar{n})$ is computed as

$$\bar{m} = \frac{1}{N}\sum_{(m,n)\in\Re}\sum m \quad , \quad \bar{n} = \frac{1}{N}\sum_{(m,n)\in\Re}\sum n$$

where m and n are horizontal and vertical axes of the image plane, respectively. As shown in Figure 2, the roll-angle is defined as the angle where the centroid of image (partially or whole) is located with respect to the sensor orientation. In practice, the true centroid of target may not be known since the whole target may not be in the FOV. Depending on the orientation of image sensors and the FOV of each lens, there may be two or three images which have partially overlapping views, as shown in Figure 5. Therefore the procedure for calculating the centroid of mass must take these overlapping pixels into account.

PASSIVE RANGING IN DIVERGENT BINOCULAR IMAGING

It is possible to obtain range information by using a divergent binocular imaging (*stereo vision*) technique. If two images from different viewpoints can be placed in correspondence, then the intersection of the lines of sight from two matching image points determines a point in three-dimensional space. This technique is conceptually simple:

(1) Consider two images separated by a baseline.

(2) Register points between the two images.

(3) Derive the equations of the two lines from the image planes to the target.

(4) Intersect the lines.

The result is the target's position in three-dimensional world coordinates. The most difficult part of this method is that of identifying corresponding points in the two images. One solution is to use correlation, or pattern matching. The idea is to take a patch of one image and match it against the other image, finding the location of the best match.

This above concepts form the basis of this passive ranging technique. Figure 6 illustrates a divergent binocular imaging system as part of the sensor array for our fuze application. The line from the target through lens i to image plane i can be described by

$$\bar{x}_i = \bar{x}_{Li} + \alpha_i(\bar{x}_{Li} - \bar{x}_{Ri}) \tag{1}$$

where

$\bar{x}_i$	: any point on the line,
$\bar{x}_{Li}$	: one known point on the line
$(\bar{x}_{Li} - \bar{x}_{Ri})$	: a known segment giving the direction of the line
α_i	: a scalar (to be determined)

In 3-dimensional space, the position of target, i.e., $\bar{x}$, is given by the point of closest approach between the two lines $\bar{x}_1$ and $\bar{x}_2$. The distance between any two points on the lines, squared, is given by

$$d^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 \tag{2}$$

where

$$x_1 = x_{L1} + \alpha_1(x_{L1} - x_{I1})$$

$$x_2 = x_{L2} + \alpha_2(x_{L2} - x_{I2})$$

and

$$\begin{aligned}(x_1 - x_2)^2 = (x_{L1} - x_{L2})^2 &+ \alpha_1^2(x_{L1} - x_{I1})^2 + \alpha_2^2(x_{L2} - x_{I2})^2 \\ &+2\alpha_1(x_{L1} - x_{L2})(x_{L1} - x_{I1}) \\ &-2\alpha_2(x_{L1} - x_{L2})(x_{L2} - x_{I2}) \\ &-2\alpha_1\alpha_2(x_{L1} - x_{I1})(x_{L2} - x_{I2})\end{aligned}$$

The same applies for the y's and the z's. The positions of the centers of lens *1* and *2* relative to the global origin are

$$\bar{x}_{L1} = \begin{pmatrix} R\sin\Theta \\ R\cos\Theta \\ 0 \end{pmatrix}, \qquad \bar{x}_{L2} = \begin{pmatrix} -R\sin\Theta \\ R\cos\Theta \\ 0 \end{pmatrix}$$

where R is the radius of curved surface, and Θ is the angle of lens from y-axis. (In our case, $\Theta = 22.5°$). Similarly, the positions of image points on image planes *1* and *2* relative to the global origin are

$$\bar{x}_{I1} = \begin{pmatrix} (R-f)\sin\Theta + u\cos\Theta \\ (R-f)\cos\Theta - u\sin\Theta \\ z \end{pmatrix}, \qquad \bar{x}_{I2} = \begin{pmatrix} -(R-f)\sin\Theta + u\cos\Theta \\ (R-f)\cos\Theta + u\sin\Theta \\ z \end{pmatrix}$$

where image position on image plane is (u, z) and f is the focal length.
After summing all terms of equation (2), the result is

$$d^2 = C_0 + C_1\alpha_1^2 + C_2\alpha_2^2 + C_3\alpha_1 + C_4\alpha_2 + C_5\alpha_1\alpha_2 \qquad (3)$$

To find the intersection between the two lines, we minimize equation (3) with respect to α_i, i.e.,

$$\frac{\partial(d^2)}{\partial\alpha_1} = 2C_1\alpha_1 + C_3 + C_5\alpha_2 = 0$$

$$\frac{\partial(d^2)}{\partial\alpha_2} = 2C_2\alpha_2 + C_4 + C_5\alpha_1 = 0$$

or

$$\begin{pmatrix} 2C_1 & C_5 \\ C_5 & 2C_2 \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} -C_3 \\ -C_5 \end{pmatrix} \qquad (4)$$

From equation (4), solve for $\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ and calculate $\bar{x}_1$ and $\bar{x}_2$. Then estimate

$$\bar{x} \approx \frac{1}{2}(\bar{x}_1 + \bar{x}_2)$$

This gives the target's position relative to the global origin, and hence the relative distance to the target is determined.

As mentioned earlier, the major concern of this stereo imaging technique is that of point or pixel registration. The basic problem with correlation in stereo imaging is that objects can look significantly different from different points of view. In some cases of

much obscuration, an important feature of a scene may only be present in one particular view. This problem can be alleviated by decreasing the baseline, but at the expense of lower accuracy in depth measurements. However, in some applications where accuracy is not so crucial, this passive ranging technique may be appropriate.

CONCLUSION

A divergent multisensor imaging system for a fuze application is presented. A roll-angle determination technique and passive ranging method based on stereo vision have been described. The major advantage of this passive ranging method is speed. Furthermore, the accuracy could be improved with better correlation or pattern matching techniques for pixel registration.

EQUIPMENT USED

The imaging equipment used is the Dark Invader (Model #3000), an ultra-low light night vision system. The system's spectral window is approximately 400 nanometer to 980 nanometer (i.e., from visible to near IR). A 850 nanometer filter is mounted to a 25 millimeter *f*1/1.4 wide-angle lens. Frame grabbing is performed through KBVision running on a Sun SparcStation. The neural network classifier is implemented using MIGRAINES neural network software, a public domain neural network development tool, and runs on a DecStation 3100.

ACKNOWLEDGEMENT

I would like to express my appreciation to Ron Boulet (Chief, Fuze Branch) and Dr. Duane Finello (my lab focal point) for the opportunity to work at the fuze laboratory and for their warmest friendships. Special thanks to my friends: John Mittleman of Coastal Systems Station (Navy) for his invaluable contribution to the phototriangulation method; Dr. Thad Roppel (Auburn University) for his help with frame-grabbing; Darryl Huddleston also for help in frame-grabbing and IR imaging; Joseph Kagel of McDonnell Douglas for discussions on multispectral sensor fusion, and Dr. Harold Szu of Naval Surface Weapons Center for discussions on IR imaging. I also wish to thank Robert Orgusaar, Bryce Sundstrom, and Dr. Min for productive discussions on passive ranging. Appreciation also goes to the Mitre Corporation for the use of MIGRAINES neural network software.

REFERENCES

[1] V. N. Dvornychenko, "Mission Parameters Derived from Optical Flow," (in preparation).

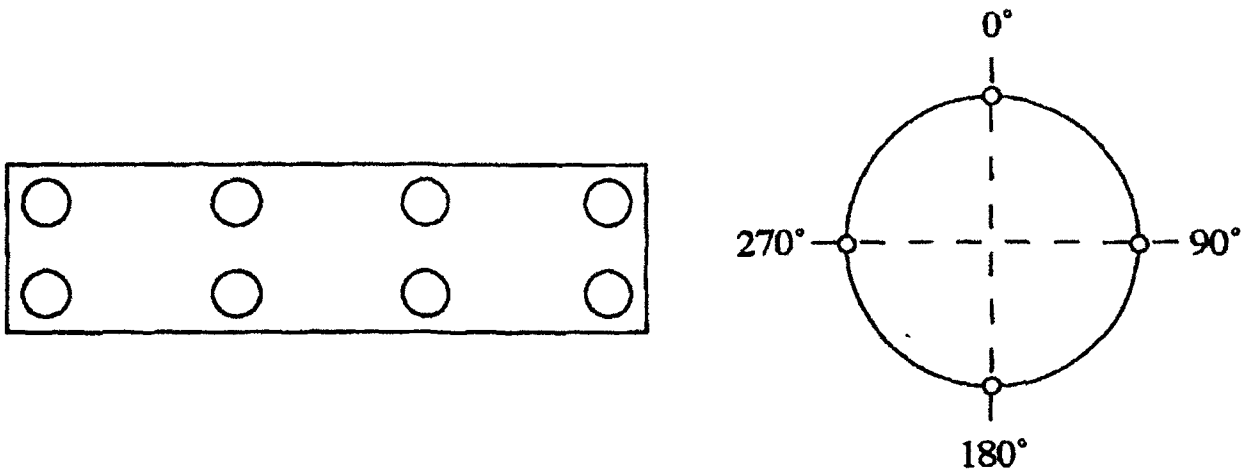
[2] B. C. Appling. Unpublished memo.

[3] S. A. Dudani, B. Smithgall, and P. Robins, "Target Acquisition and Extraction from Cluttered Backgrounds, " Proc. of SPIE, Vol. 302, 1981, pp. 43-50.

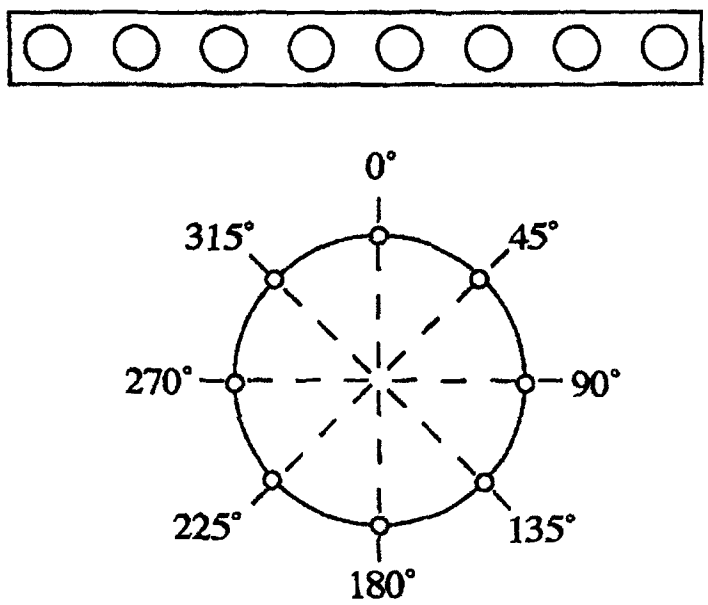
[6] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," Parallel Distributed Processing: Explorations in the Microstructures of Cognition, Vol. I, pp. 318-362, MIT Press, Cambridge, MA, 1986.

[4] M. Burton and C. Benning, "Comparison of Imaging Infrared Detection Algorithms," Proc. of SPIE, Vol. 302, 1981, pp. 26-31.

[5] H. R. Keshavan and C. M. Lo, "Comparative Study of Edge Thinning Algorithms for Target Identification," Proc. of SPIE, Vol. 302, 1981, pp. 26-31.



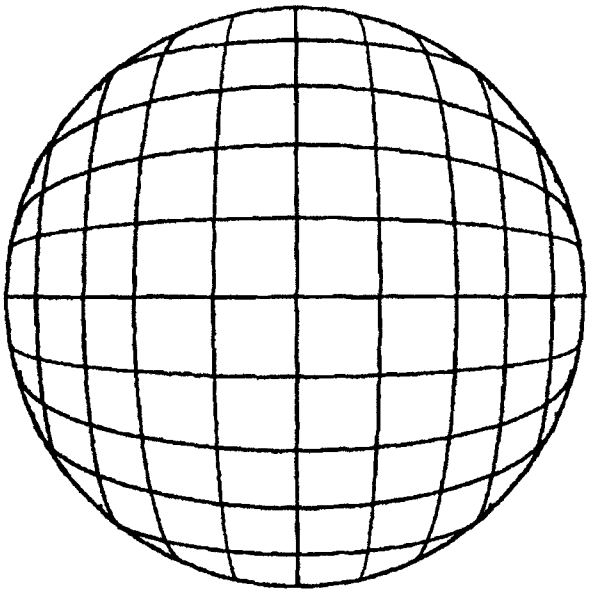
(a)



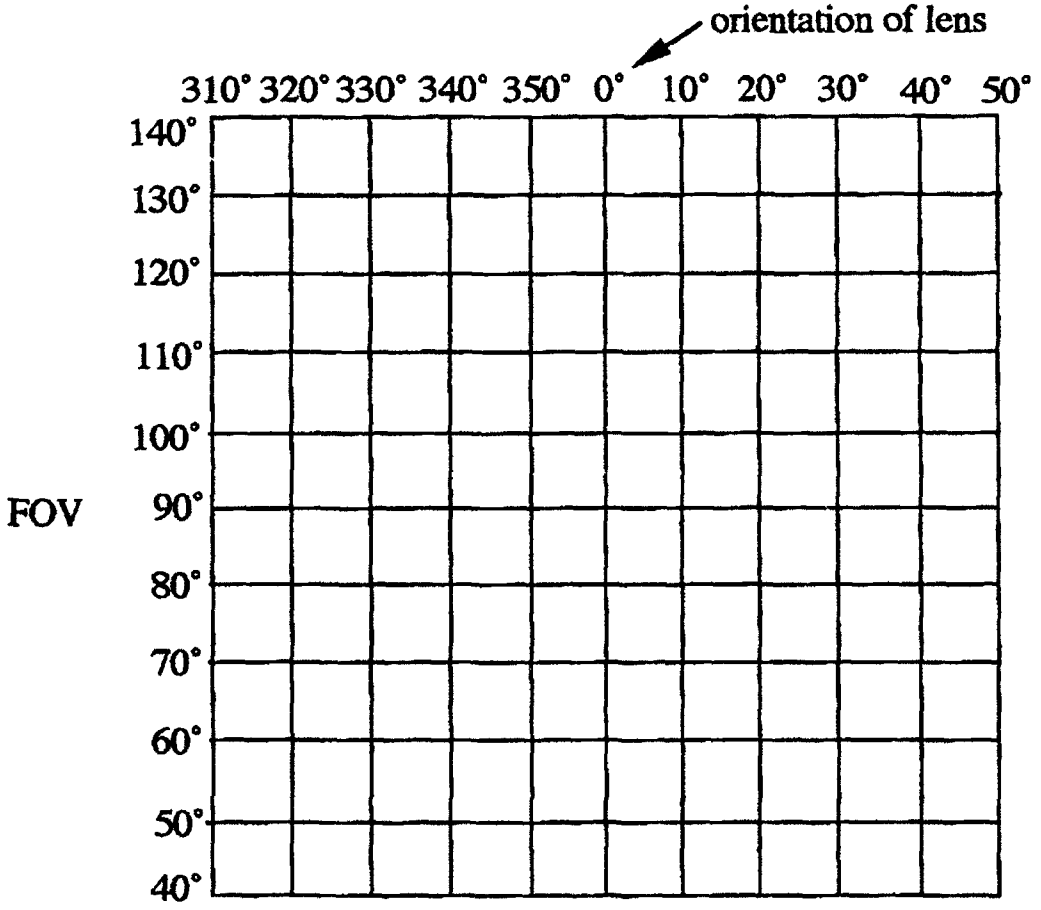
(b)

Figure 1. Possible sensor arrangements and their orientations in a divergent multisensor imaging system.

Figure 2. Two-dimensional maps for roll-angle determination.



(a) Distorted image



(b) Ideal image

Figure 3. Fish-eye view with geometric distortion, and an ideal image.

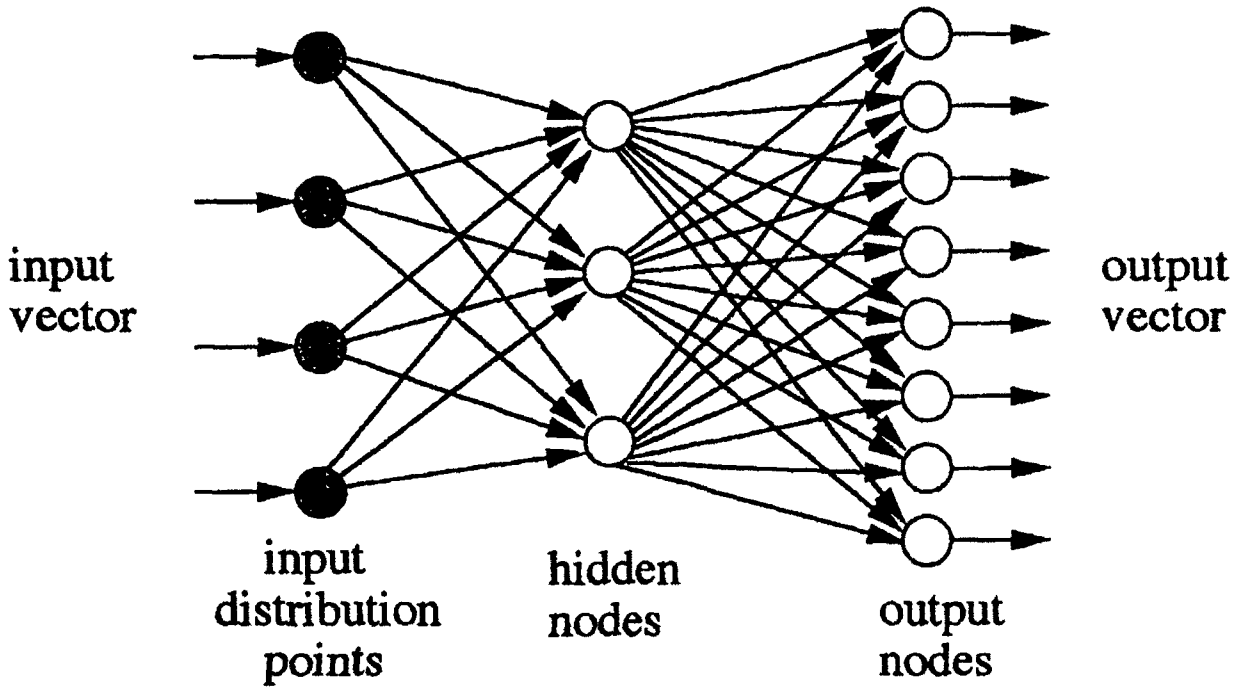
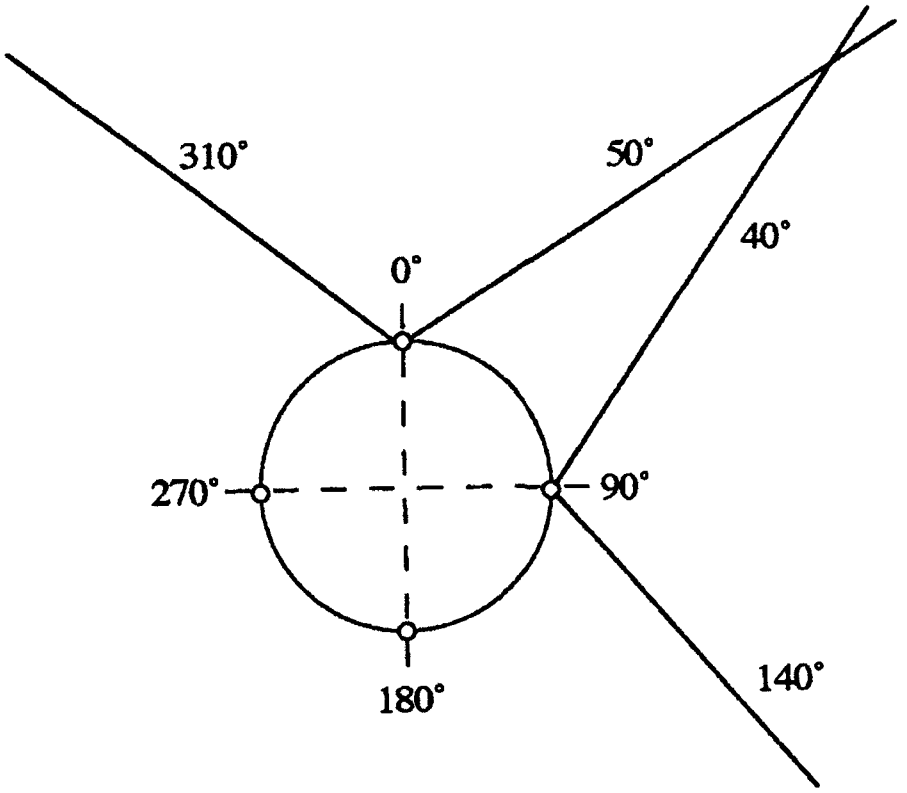
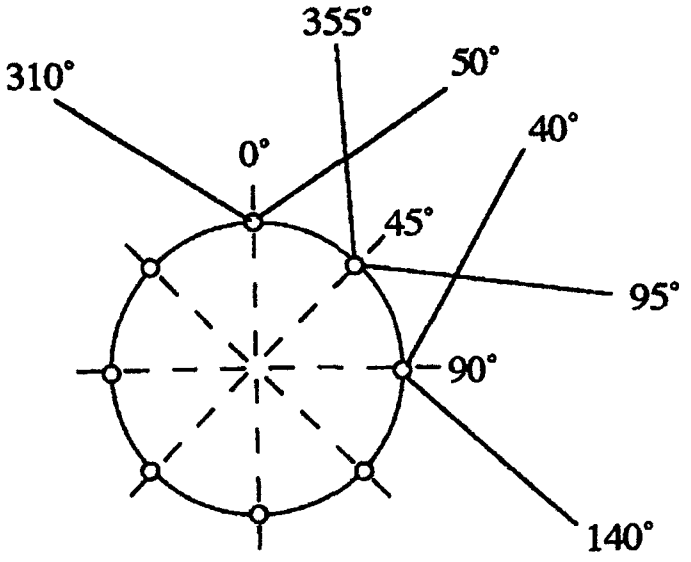


Figure 4. A feedforward neural network classifier.



(a)



(b)

Figure 5. Overlapping views with FOV = 100°.

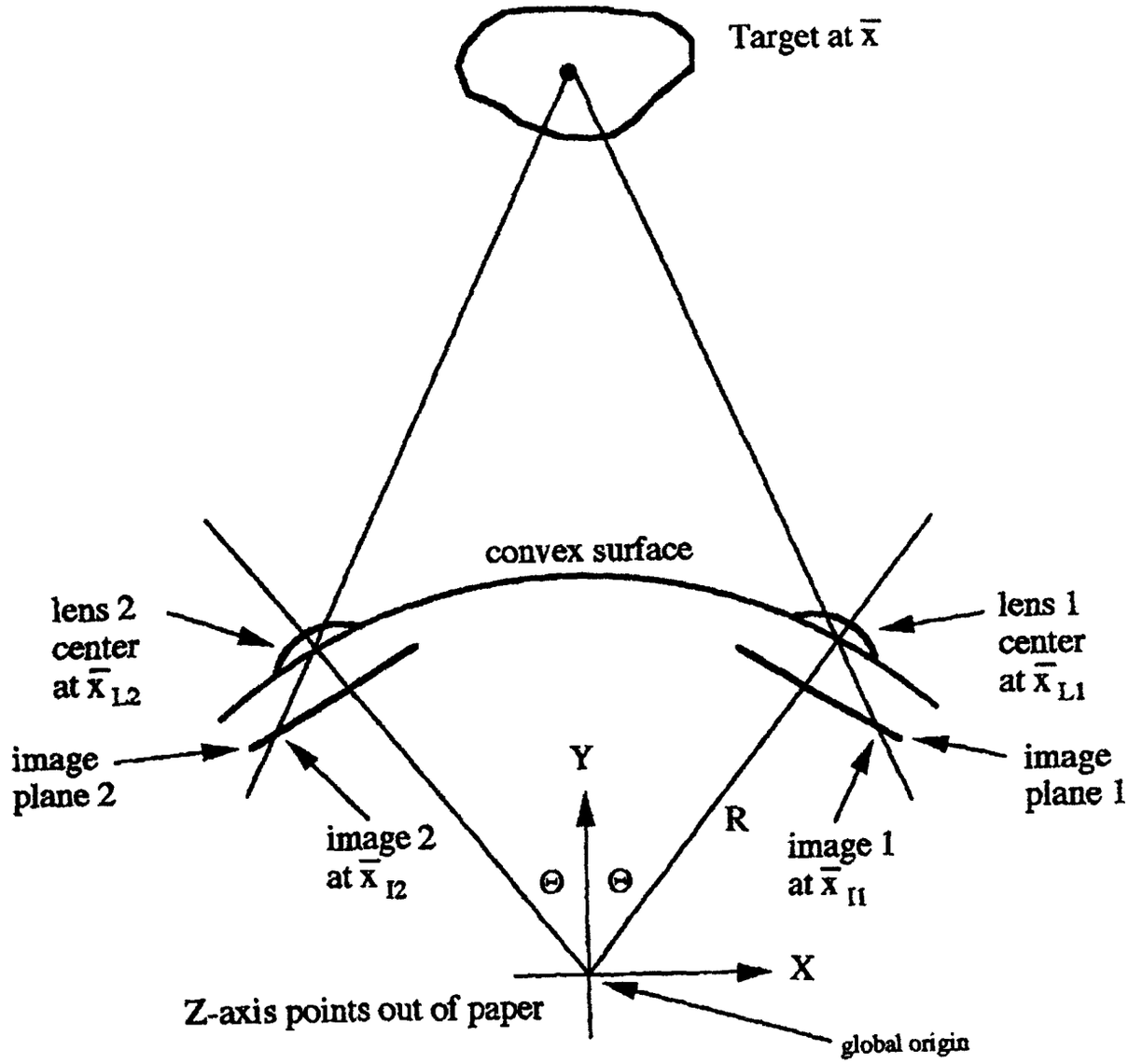


Figure 6. Passive Ranging Using Divergent Binocular Imaging

VELOCITY AND TEMPERATURE MEASUREMENTS IN A HIGH SWIRL DUMP COMBUSTOR

Dr. Richard D. Gould and Mr. Lance H. Benedict

Mechanical and Aerospace Engineering
North Carolina State University
Raleigh, NC 27695

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

August 1992

VELOCITY AND TEMPERATURE MEASUREMENTS IN A HIGH SWIRL DUMP COMBUSTOR

Dr. Richard D. Gould and Mr. Lance H. Benedict
Mechanical and Aerospace Engineering
North Carolina State University
Raleigh, NC 27695

ABSTRACT

Successful two component laser Doppler velocimetry (LDV) and single point temperature measurements were made in the highly swirling flow field of a model dump combustor. A type S Platinum/10 % Platinum-Rhodium thermocouple probe was used to make temperature measurements. A lean propane-air diffusion flame with an overall equivalence ratio of $\Phi = 0.45$ was stabilized in the combustion chamber by the flow pattern of the Wright Laboratory/Rolls Royce (WL/RR) swirler. The reacting flow case was found to have higher axial and tangential mean velocities than the isothermal flow case throughout most of the flow field due to heat release. A shorter central recirculation bubble was found in the reacting flow case due to the pressure gradients produced by these higher mean velocities. Turbulent normal stresses were found to be larger while the maximum value of the $\overline{u'w'}$ turbulent shear stress was found to be less in the reacting flow case when compared to the isothermal flow case. Maximum mean temperatures occurred were maximum turbulent stresses occurred thus suggesting that gradient transport modeling may be successful in predicting this flow field.

VELOCITY AND TEMPERATURE MEASUREMENTS IN A HIGH SWIRL DUMP COMBUSTOR

Dr. Richard D. Gould and Mr. Lance H. Benedict

INTRODUCTION

The interaction between combustion and turbulence has been the subject of much research in the past few years. The effect of combustion on turbulence structure and the effect of turbulence on combustion are of great interest to the combustion engineer. Knowledge of these effects would permit the design of more compact energy efficient combustion systems with low pollutant emission while minimizing costly test programs.

The objective of most recent experimental work in turbulent flows has been to obtain data for comparison with numerical prediction codes that model turbulence, combustion, and heat transfer. Velocity, temperature, and species concentration measurements in reacting flows have been made in various geometries including: 1) co-axial jets, 2) diffusion flame jets, 3) bluff body flame holders in ducts, 4) industrial furnaces, 5) two-dimensional and axisymmetric rearward facing steps, 6) confined swirling flows. All of these geometries have a common feature in that they induce a separated flow region which acts as an anchor for the flame front. High turbulence intensities and well defined recirculation zones are characteristic of these flows. Comparisons of mean velocities, turbulence intensities, temperature distributions and recirculation zone sizes have been made in flows with and without combustion in an effort to understand the turbulence-combustion interaction process. Turbulence structure parameters (*i.e.* skewness and flatness) are also often compared to give additional insight into interaction mechanisms. Although substantial literature exists, there is still an incomplete understanding of the turbulence-combustion interaction. This is because turbulent combustion is highly inhomogeneous, involving large fluctuations in temperature, composition, density and velocity. There also can be a strong interaction between the aerodynamic and heat release mechanisms. Modeling of turbulent aerodynamics requires the use of statistical properties including the fluctuation intensity, covariance, energy spectrum and probability density functions. The development of accurate models is dependent upon reliable experimental information concerning these quantities. This study was undertaken to provide data necessary to improve the turbulent combustion modeling process in a highly swirling diffusion flame dump combustor.

EXPERIMENTAL APPARATUS

A swirling axisymmetric sudden expansion flow geometry was produced by joining a 3.5 m long entry pipe having a 101.6 mm (4 in.) inside diameter to a water cooled 152.4 mm(6 in.) inside diameter test section pipe as shown in Figure 1. A specially built swirler, through a collaborative effort between Wright Laboratory and Rolls Royce, having a 30 mm inside diameter and 61 mm outside diameter was used to produce the swirl in the combustion chamber. The swirler had 20 curved vanes and was designed to give a mean airflow turning angle of 45°. Well regulated gaseous propane was fed through a central tube to the fuel injector. The fuel injector consisted of a reservoir, a mixing chamber and a movable pintle

for control of the fuel exit velocity. The fuel flow annulus area for this study was 0.000307 m^2 giving a fuel exit velocity of approximately 4.4 m/s fot the flow conditions used here. The entry pipe-swirler-fuel injector assembly and sudden expansion face were mounted on a movable table and thus could be positioned at various axial locations in the rigidly fixed test section. This arrangement allowed for measurements at various downstream positions within the sudden expansion flow field without having to move the LDV probe volume location in the axial direction. It should be noted that measurements were made only on the diameter of the test section. The face of the sudden expansion is moved to place the probe volume at a new axial position. High quality optical access can be limited to a small window with this apparatus. A metal ring was used to seal the small gap between the sudden expansion face and the inside diameter of the test section. Air was provided by well regulated air compressors and was monitored using a calibrated orifice plate located upstream of a large settling chamber which precedes the entrance pipe. Flat quartz windows 50 mm in diameter and 3.2 mm thick(2 in × .125 in) where mounted in optical access ports on both sides of the 152.4 mm diameter test section such that the inner flat surfaces were nearly flush (0.060 in. inset) with the inside diameter of the test section.

A TSI two-component dual-beam LDV system, operating in forward scatter mode, was used in this study. This system was oriented to make direct measurement of the axial and tangential velocity components on the diameter of the test section. Measurements were made at 16 radial locations and at 14 axial planes as indicated in Figure 2. The 514.5μm argon laser line was used to measure the axial velocity component while the 488μm argon laser line was used to measure the tangential velocity component with this system. Bragg cells shifted the frequency of one beam in each component by 40 MHz causing the fringes to move in the downstream direction for the axial component and vertically upward for the tangential component. Fringe spacing were measured and found to be 1.825 μm $\pm$.006 for the axial component and 1.731 μm $\pm$.006 for the tangential component. The LDV transmitting(fiber optic head) and receiving optics were mounted on a precision xyz positioning table with resolution of $\pm$2.5μm in each axis. The LDV system employed × 3.75 beam expansion optics and gave probe volumes approximately 60μm in diameter and 450μm in length. Beam blocks were placed in front of the collection optics when measurements were being made so that the Doppler signal could be detected. Large refractive index gradients in the combustion chamber, due to the combustion process, made it necessary to adjust the photodetector pinhole locations at every third or fourth measurement point along the radius. This procedure involved turning off the photodetectors, removing the beam blocks, adjusting the relay lens just upstream of the photodetector so that the image of the beam crossing is located on the pinhole plane, and lastly, adjusting the pinhole position so that the beams pass through. Narrow bandpass filters were placed in front of each photomultiplier tube to eliminate cross-talk between the two channels.

Two TSI Model 1990 counter processors interfaced to a custom built coincidence timing unit were used in the data collection and processing system. High and low pass filters were set to 10 MHz and 100 MHz for both channels, respectively. Both processors were set to make a single measurement per burst, count 16 fringes and use a 1 % comparator. A hardware coincident window was set at 20μs for all of the tests. Data(two velocities and the

running time for each realization) were transferred through two DMA ports to a MicroVax minicomputer and later uploaded to a VAX 8650 for analysis.

The flow field was seeded using titanium dioxide (TiO_2) particles generated by reacting dry titanium tetrachloride ($TiCl_4$) with the moist shop air. Craig et al. (1984) measured the particle sizes generated by this device and found that they were fairly uniform and in the $0.2 - 1\mu m$ diameter range. Data validation rates varied between 10000 and 200 per second on each counter processor and depended mainly on how clean the quartz windows were and on the location in the flow field. Low data validation rates occurred in the central recirculation zone and near the windows. Coincident data validation rates ranged from 3000 to 50 measurements per second. Velocity bias was not considered in this study.

Single point temperature measurements were made by manually traversing a type S Platinum/10 % Platinum-Rhodium thermocouple probe across the radius of the test section at each axial plane. The thermocouple bead was unshielded and was approximately 0.015 in. in diameter. A NEFF analog signal data acquisition system with a built in electronic ice point junction and ANSI standard thermocouple tables was used to acquire the mean temperature at each point in the flow field. The mean temperature at each point was calculated by averaging 225 temperatures sampled over a 10 second period. No attempt has been made to correct these mean temperatures for radiation losses.

EXPERIMENTAL PROCEDURE

All flow conditions were maintained at near constant values throughout the testing procedure. The inlet reference velocity, U_{ref}, was maintained at 30.5 m/s $\pm$ 0.9 m/s(100 ft/s $\pm$ 3.0 ft/s) giving an air mass flow rate of 0.091 kg/s $\pm$ 0.0027 kg/s (12 lbm/min $\pm$ 0.36 lbm/min). The propane fuel flow rate was maintained at 0.0026 kg/s $\pm$ 0.000078 kg/s (0.35 lbm/min $\pm$ 0.011 lbm/min) giving an overall equivalence ratio of $\Phi = 0.45 \pm 0.014$. Turbulent velocity statistics and histograms were formed by using 5000 individual realizations for each velocity channel at each measurement point. Mean temperature measurements were formed by using 225 individual measurements taken over a 10 second period. In computing statistical parameters from the LDV data a two step process was used to eliminate noise from the data. In the first step, a 5% threshold level was applied to the raw velocity data in an effort to estimate the standard deviation of the valid data. This estimate was made by putting the raw velocity data in 100 equally spaced bins bounded by the actual maximum and minimum velocity found in the raw data sample. Next, the bin with the maximum number of samples in it was found. Finally, all bins having at least 5% of the number of samples found in the "maximum" bin were located. The width of the data which met this threshold criteria was then used to estimate the standard deviation of the "good" data. Upper and lower cutoff limits were set by adding and subtracting, respectively, 2.5 times the half-width of the data which met this 5% threshold. Applying this technique to a Gaussian distribution is equivalent to setting cutoff limits which correspond to ± 4.1 standard deviations and thus this first step was used only to remove spurious data. This method is a variation of the method suggested by Meyers (1988) and was used to eliminate spurious data which if not removed would give an abnormally large value for the standard deviation and thus wider cutoff limits. In the second step, the mean and standard deviation of the remaining data(spurious data removed)

were calculated. This data was then further filtered to remove data which deviates more than 3 standard deviations from this new mean. Finally, revised statistics were calculated once these additional outliers were discarded. For a properly operating LDV system very few points are removed during the first step(typically less than 10 out of 5000) and less than 1% of the data should be discarded after both steps.

EXPERIMENTAL RESULTS

Two-component velocity measurements were made at 16 radial positions and at 14 axial planes in the swirling, axisymmetric reacting flow field of a model dump combustor. An overall lean diffusion flame was stabilized by the swirling shear layer. Figures 3 through 14 show the measured profiles of the mean and rms axial and tangential velocities, the $\overline{u'w'}$ turbulent shear stress, and the mean temperature in this combustor at 12 of these axial locations. The first axial position, $x = 25.4$ mm (1 in.), was located as close to the sudden expansion face as possible with the beam angles and beam orientation (*i.e.* 0^o and 90^o) used in this study. Measurements were made at axial stations every 12.8 mm (0.5 in.) from this first position up to $x = 152.4$ mm (6 in.) in order to capture the detailed velocity and temperature field in the highly swirling region of the flow field. Measurements were made at three additional downstream axial locations, x =203.2 mm (8 in.), 254 mm (10 in.) and 304.8 mm (12 in.), to complete the description of the flow field. Also included in these figures, for comparative purposes, are the velocity measurements made by Ahmed et al. (1992) for the isothermal swirling flow case using the same geometry, swirler, and inlet reference flow velocity as used in this reacting flow study.

The mean velocity profiles show that higher velocities are present in the reacting flow case when compared to the isothermal flow case. This is to be expected due to volumetric expansion caused by heat release associated with the combustion process. The higher mean velocities near the combustor wall produce a well defined central recirculation bubble. This central recirculation bubble was found to be approximately 140 mm (5.5 in.) long in the reacting flow case which is much shorter than in the isothermal flow case. This strong, short central recirculation zone is probably why this combustor is stable at low overall equivalence ratios. In fact, flow visualization studies indicated that lean blowout occurred at $\Phi = 0.21$. It would be interesting to calculate the swirl number for the reacting flow case to determine the combined effects of reduced density and increased mean velocity on the axial momentum flux. The maximum values of turbulent normal stresses were found to be larger in the reacting flow case than in the isothermal flow case. These larger fluctuations in velocity may be due to fluctuating heat release which is known to be highly non-linear. Further data analysis is required to determine if this increase in fluctuations is due to acoustic phenomena (*i.e.* unsteady flow). The axial and tangential turbulent normal stresses were found to be nearly equal to one another suggesting that isotropic turbulence models may be sufficient. It should be mentioned however, that the radial velocity fluctuation was not measured in this study, but that it may show larger anisotropy than the two components measured here, thereby negating the above comment. The maximum $\overline{u'w'}$ turbulent shear stress value was found to be much lower in the reacting flow case when compared to the isothermal flow case. This may be due to the larger value of kinematic viscosity(*i.e.* a factor of 10 - 15 larger due to increased temperature) in the reacting flow case causing increased turbulence

dissipation. Temperature measurements indicate that the peak occurs at the boundary of the central recirculation bubble and the high speed wall flow. This temperature peak location also corresponds to the location where turbulent stresses peak and suggests that gradient transport modeling may be appropriate to predict this flow field. It should be mentioned that these temperature measurements should be viewed as qualitative until they have been corrected for radiation losses and until the effects of probe induced flow disturbances can be quantified.

CONCLUSIONS

Successful two component laser Doppler velocimetry (LDV) and single point temperature measurements were made in the highly swirling flow field of a model dump combustor. The velocity measurements were possible only because forward scatter signal collection was used. The reacting flow case was found to have higher axial and tangential mean velocities throughout most of the flow field than the isothermal flow case due to heat release. A shorter, stronger central recirculation bubble was found in the reacting flow case due to the pressure gradients produced by these higher mean velocities. Turbulent normal stresses were found to be larger while the maximum value of the $\overline{u'w'}$ turbulent shear stress was found to be less in the reacting flow case when compared to the isothermal flow case. Maximum mean temperatures occurred were maximum turbulent stresses occurred thus suggesting that gradient transport modeling may be successful in predicting this flow field.

ACKNOWLEDGMENTS

The author would like to thank Dr. A. S. Nejad for use of his laboratory and equipment and for all the help he provided. Thanks are also due to Dr. S. A. Ahmed for help in operating the LDV data acquisition system and to Mr. C. Smith for all his technical support in the laboratory. This investigation was performed at Aeropropulsion and Power Directorate. Wright Laboratory (WL/POPT) under the Summer Faculty Research Program supported by AFOSR.

REFERENCES

1. Ahmed, S. A., Rose, A., and Nejad, A. S., (1992) "Three Component LDV Velocity Measurements in a Can Type Research Combustor for CFD Validation Part 1 Isothermal," Presented at the Int'l Gas Turbine and Aeroengine Congress and Exposition. Cologne, Germany, June 1-4.

2. Craig, R. R., Nejad, A. S., Hahn, E. Y. and Schwartzkopf, K. G. 1984 "A General Approach for Obtaining Unbiased LDV Data in Highly Turbulent Non-Reacting and Reacting Flows," AIAA Paper. No. 84-0366.

3. Meyers, J. F., (1988) "Laser Velocimeter Data Acquisition and Real Time Processing using a Microcomputer," Proceedings of the 4th International Symposium on Applications of Laser Anemometry to Fluid Mechanics, Lisbon, Portugal, p. 7.20.

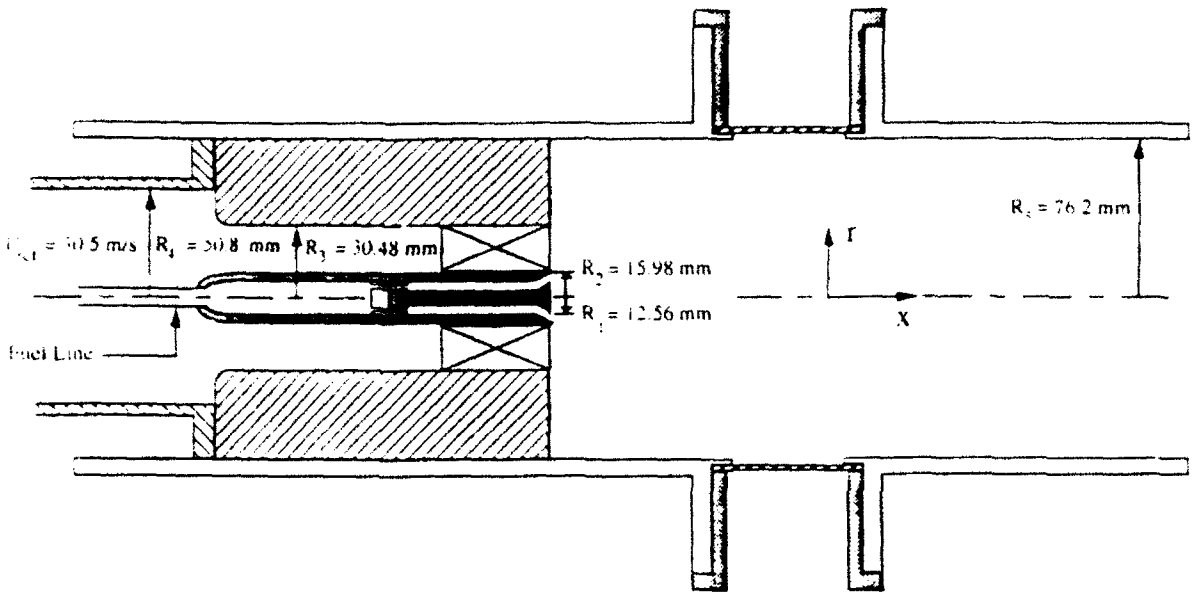


Figure 1. Model dump combustor geometry.

(Top View)

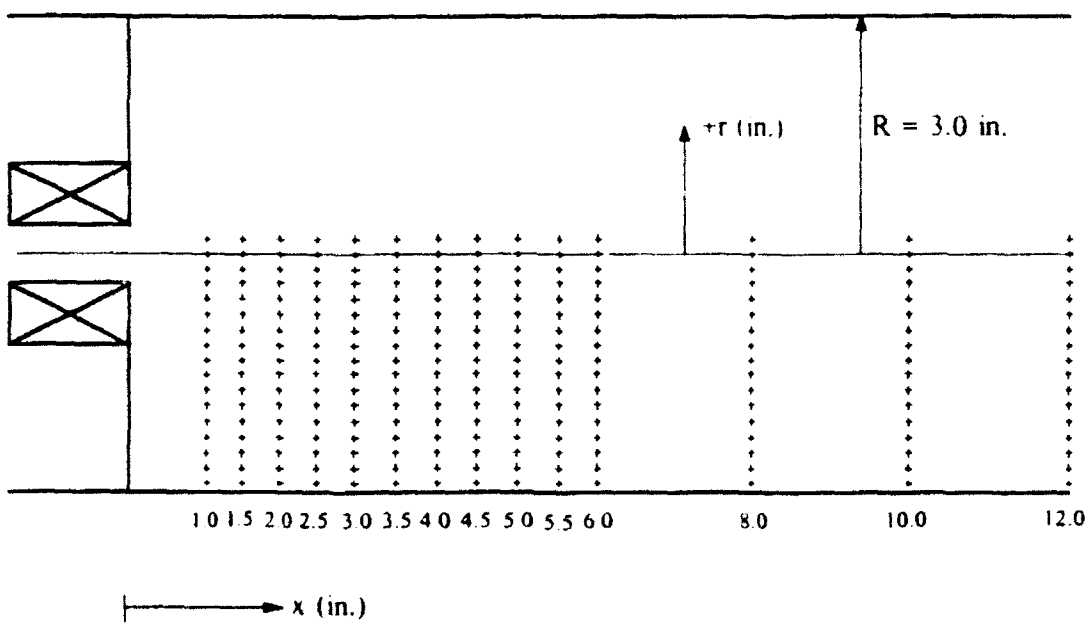


Figure 2. Measurement locations.

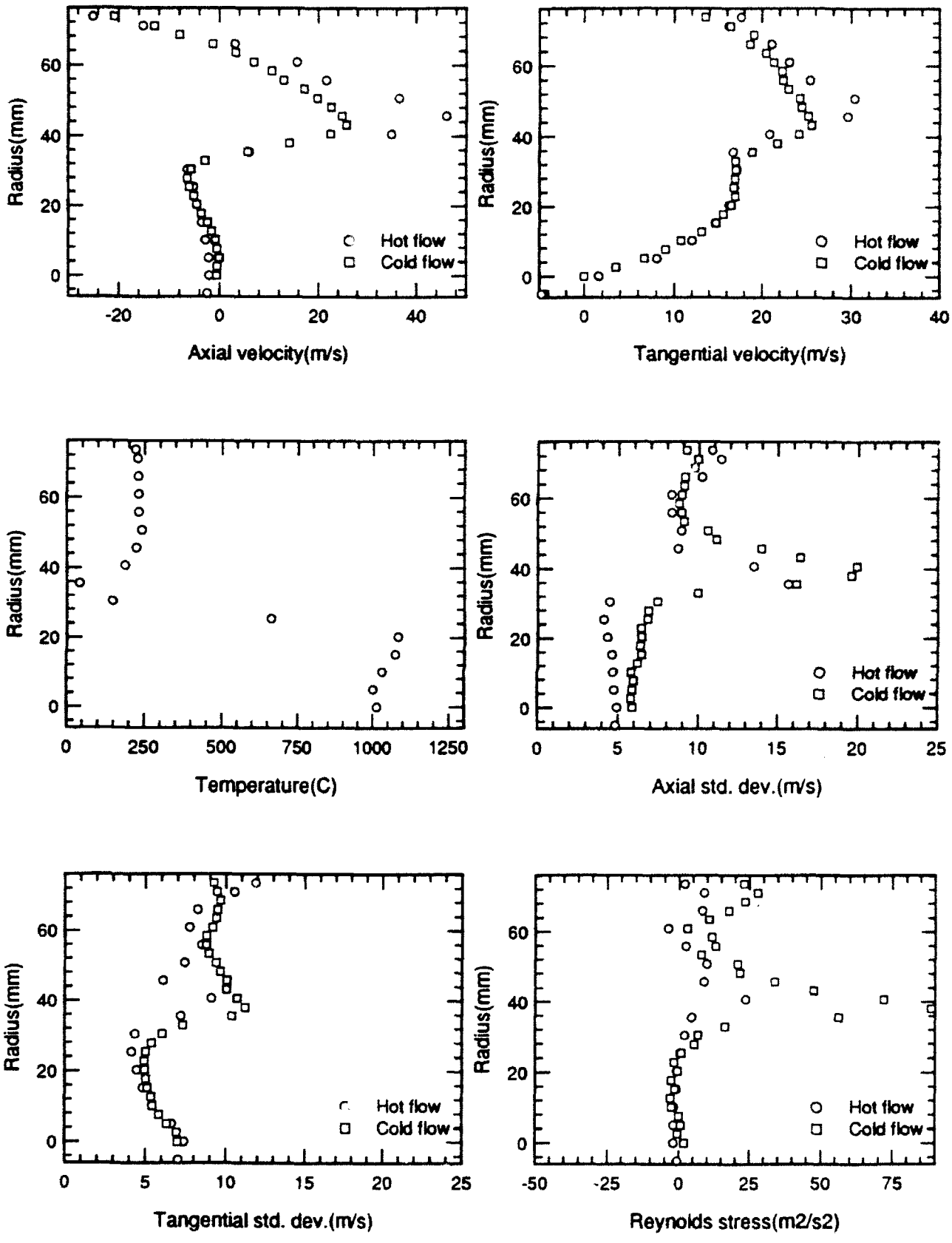


Figure 3. WL/RR Swirl Combustor, x = 25.4 mm (1 in.).

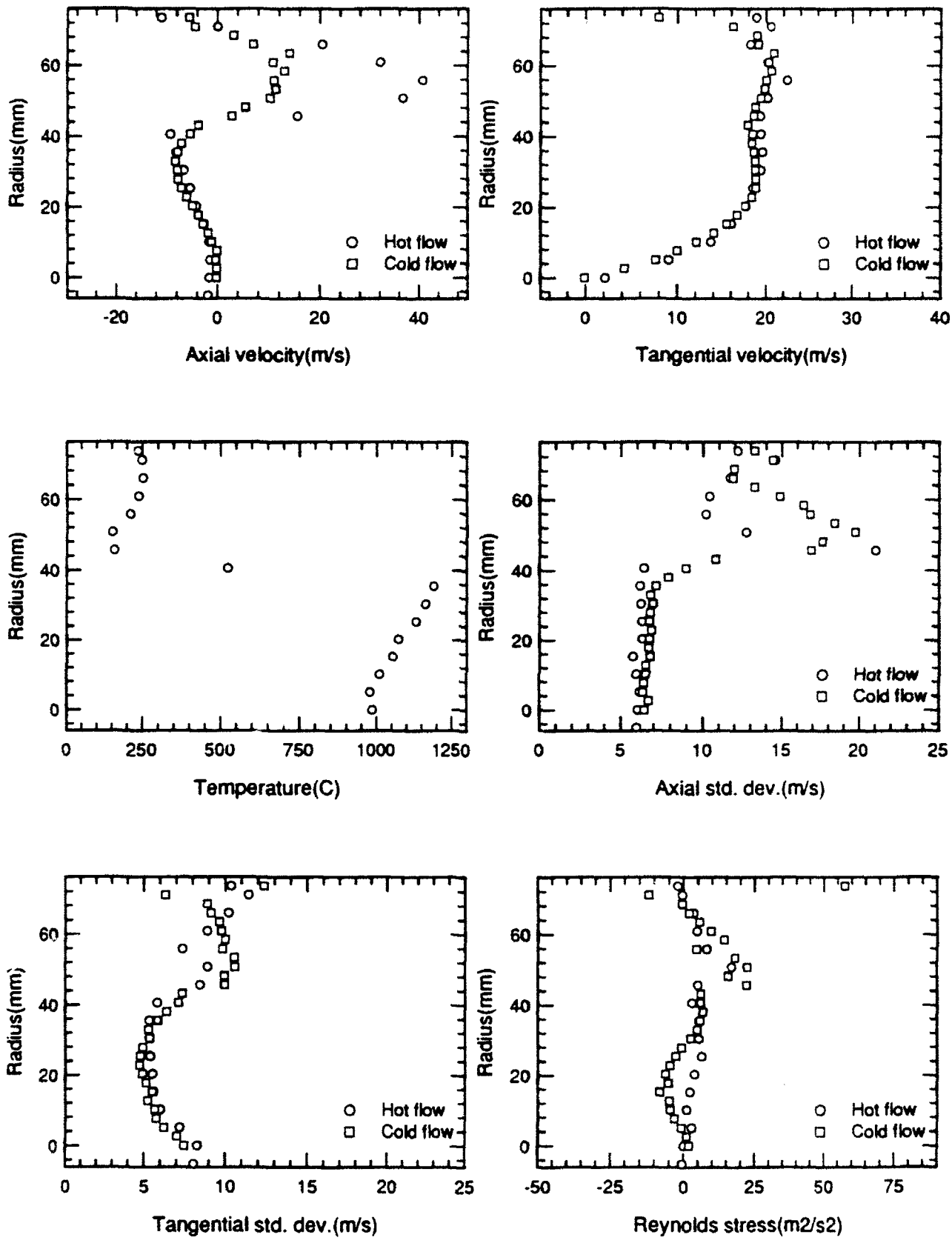


Figure 4. WL/RR Swirl Combustor, x = 38.1 mm (1.5 in.).

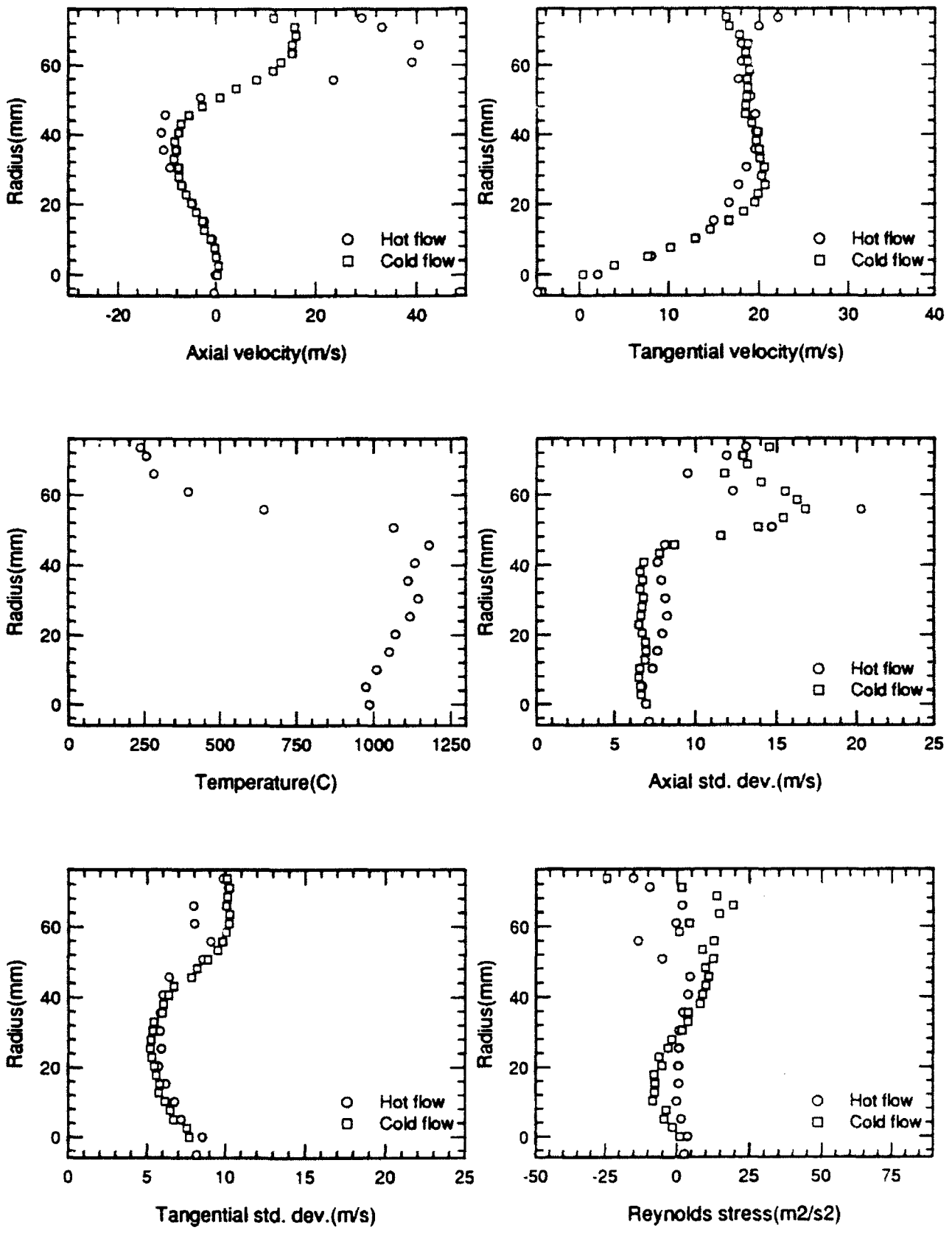


Figure 5. WL/RR Swirl Combustor, x = 50.8 mm (2 in.).

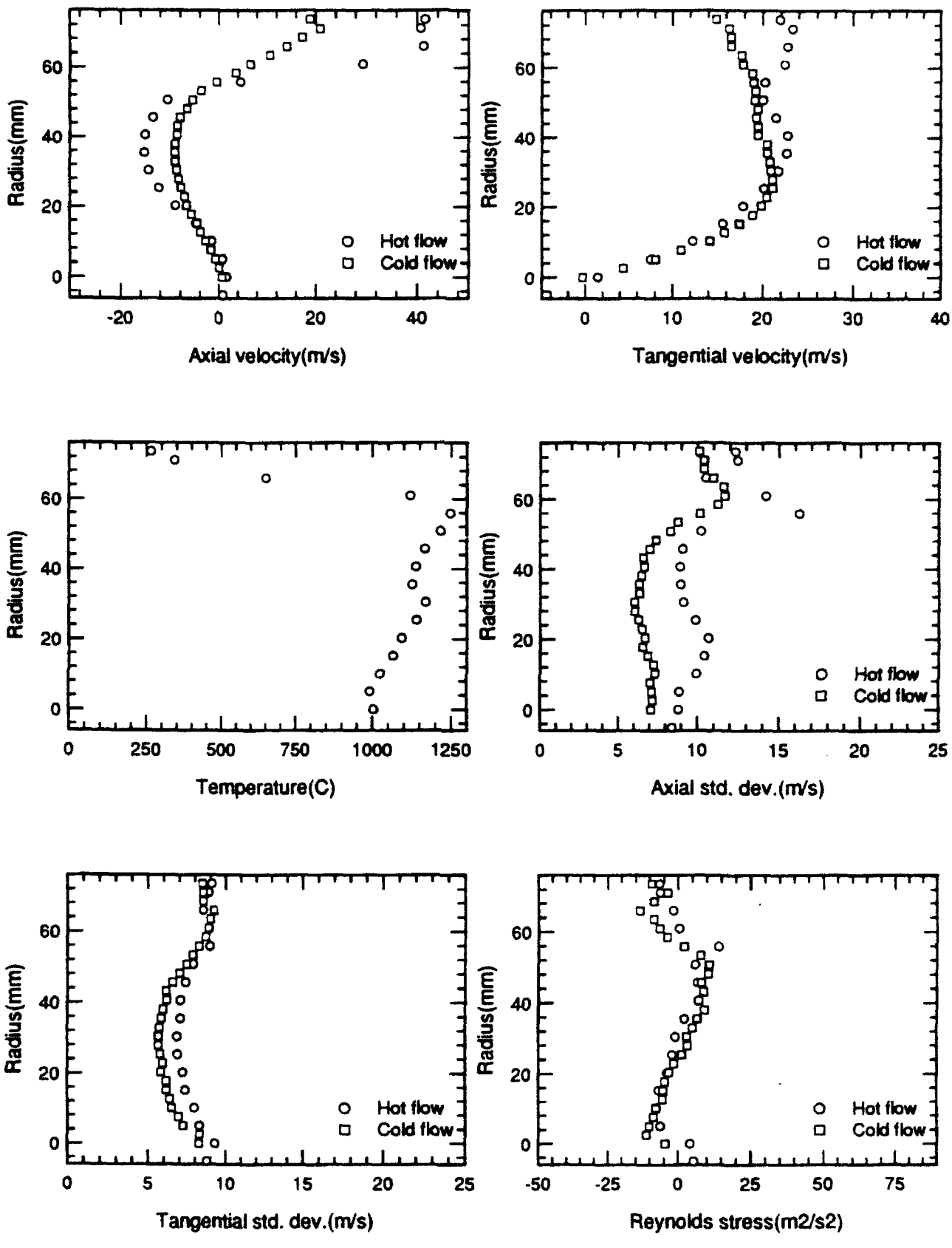


Figure 6. WL/RR Swirl Combustor, x = 63.5 mm (2.5 in.).

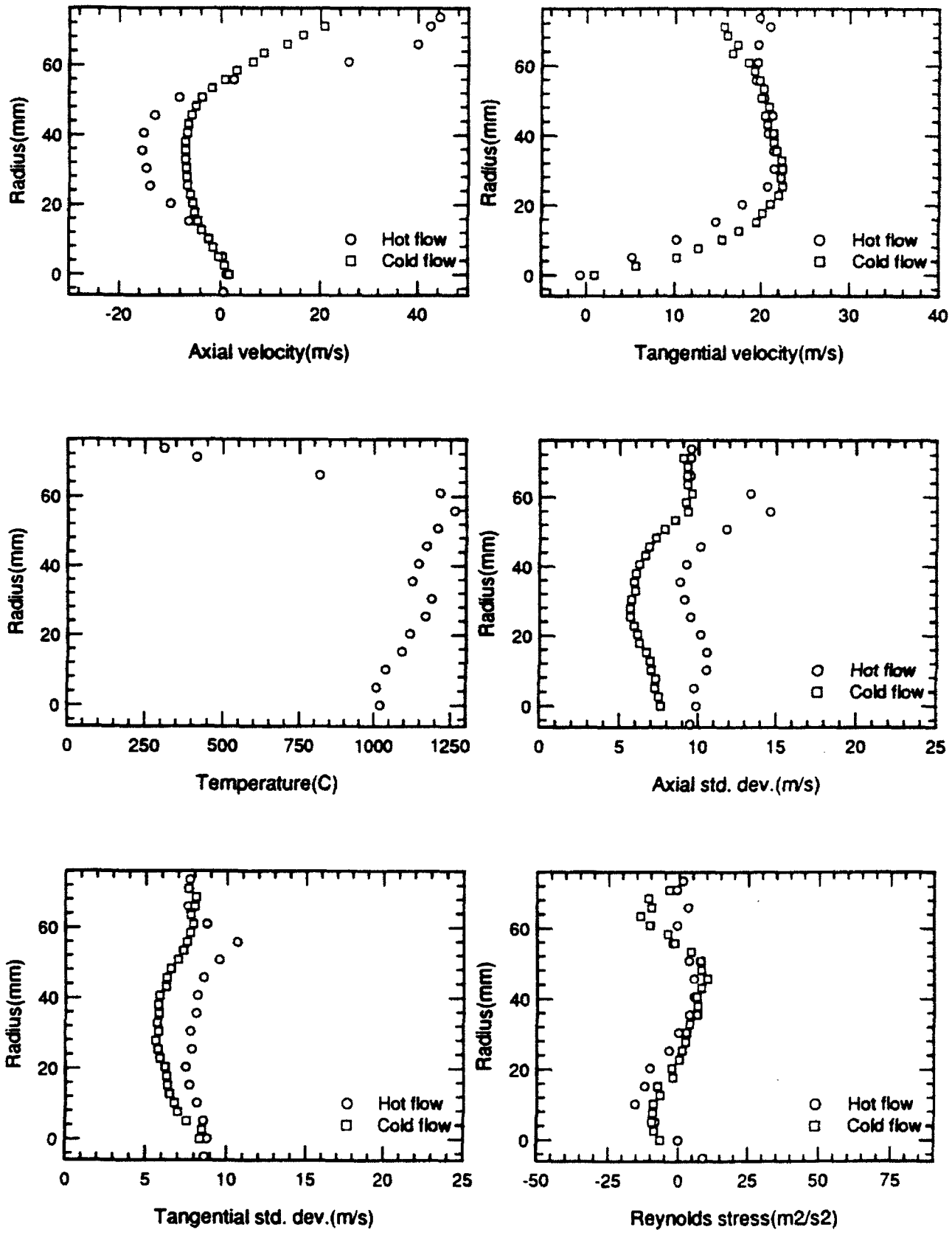


Figure 7. WL/RR Swirl Combustor, x = 76.2 mm (3 in.).

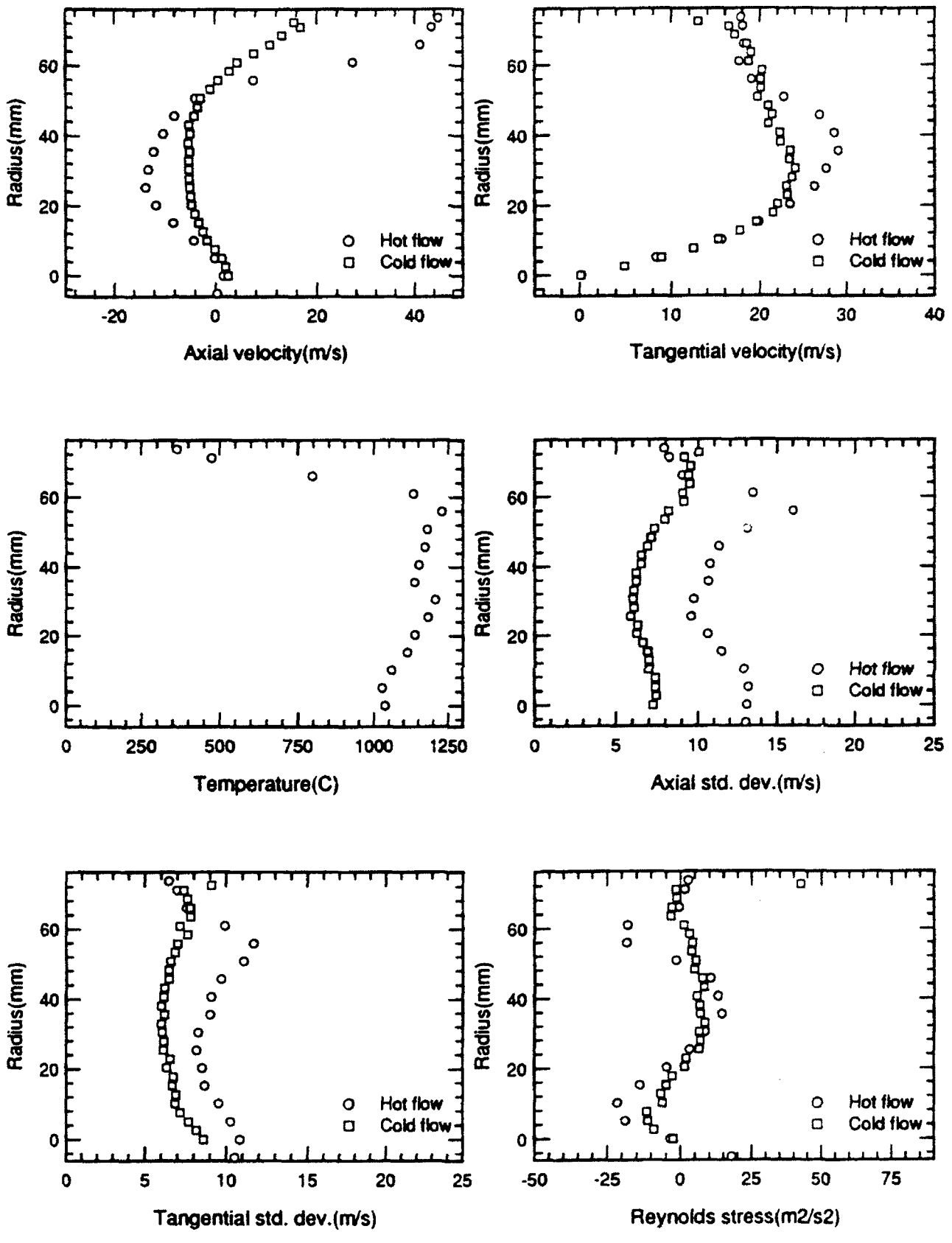


Figure 8. WL/RR Swirl Combustor, x = 88.9 mm (3.5 in.).

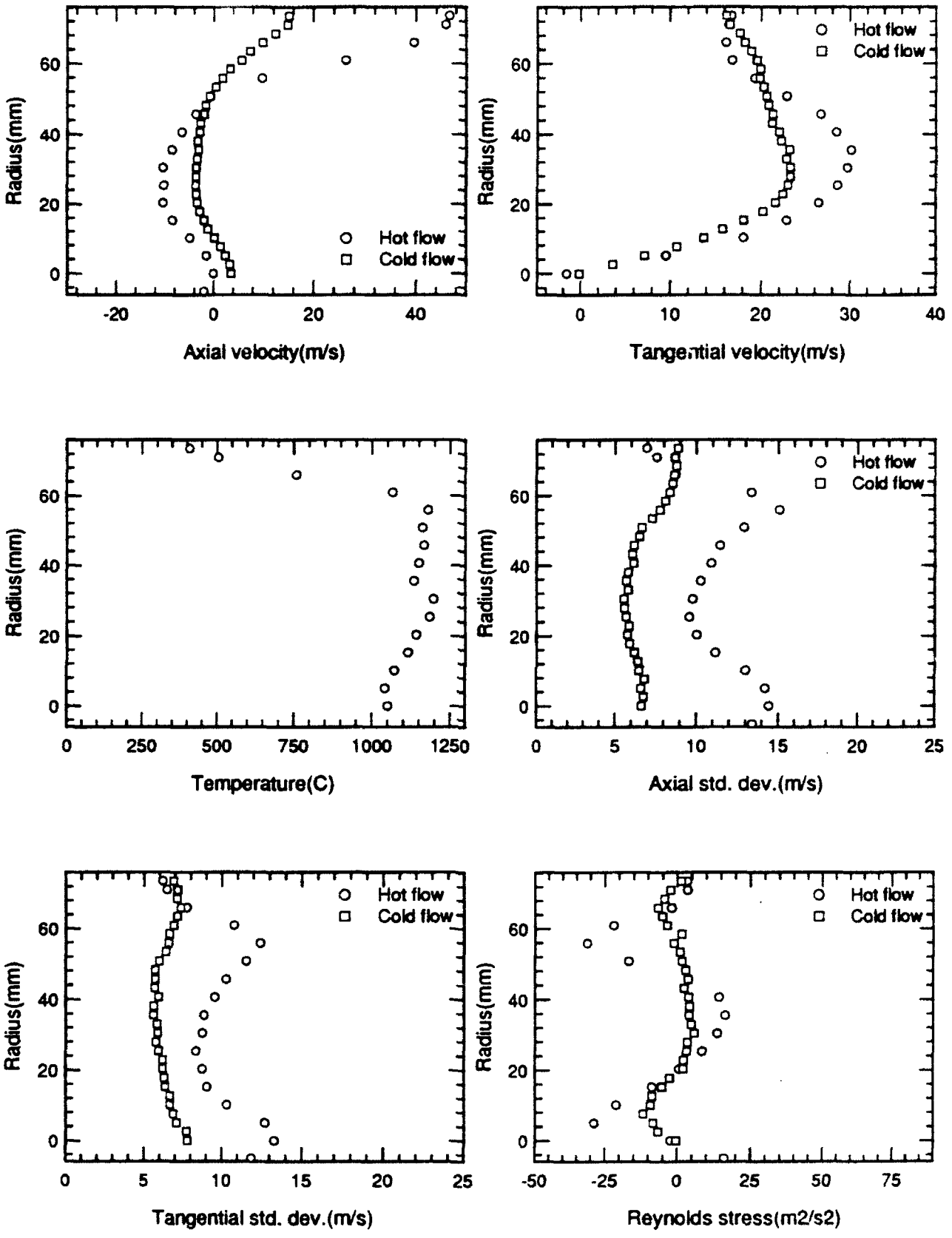


Figure 9. WL/RR Swirl Combustor, x = 101.6 mm (4 in.).

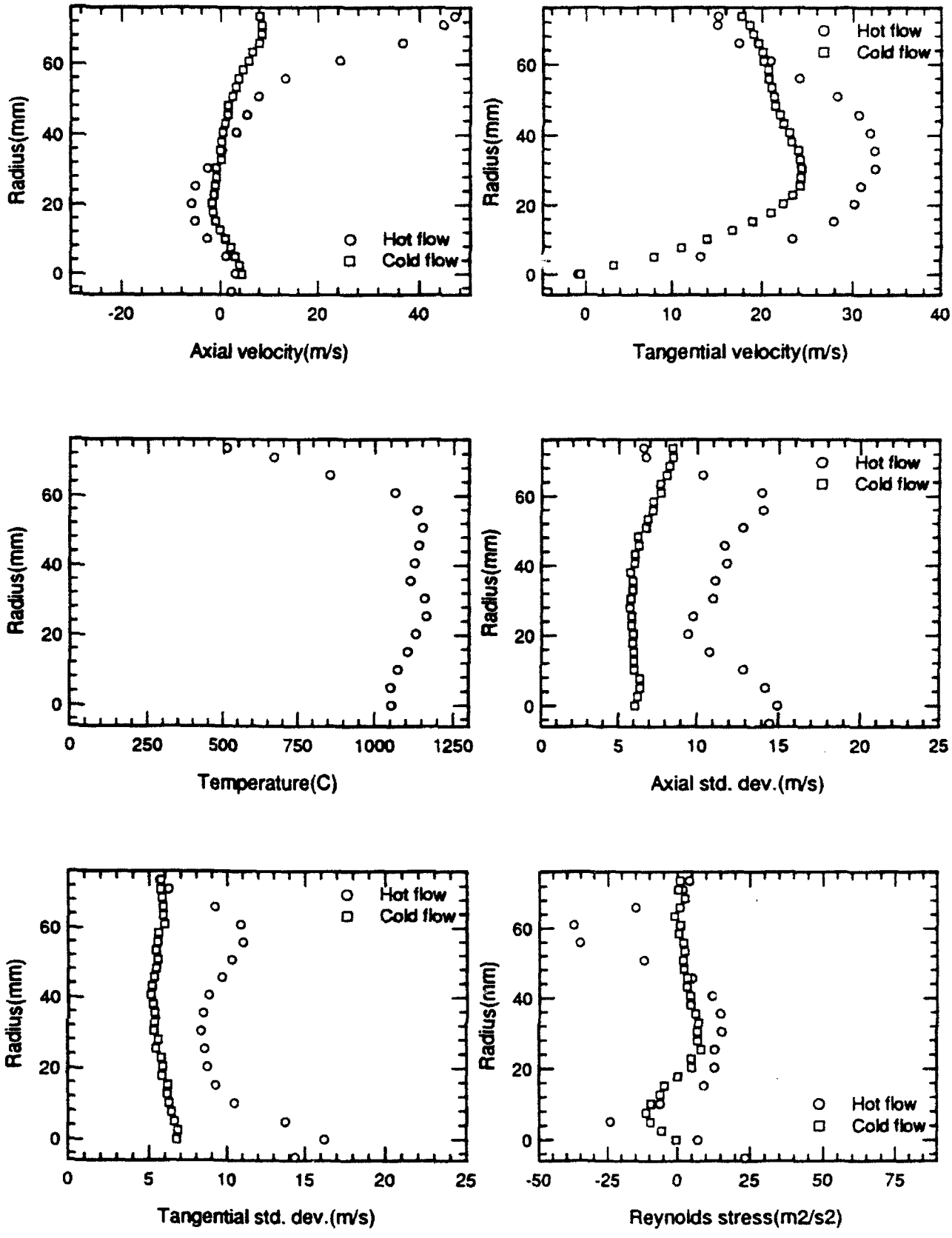


Figure 10. WL/RR Swirl Combustor, x = 127 mm (5 in.).

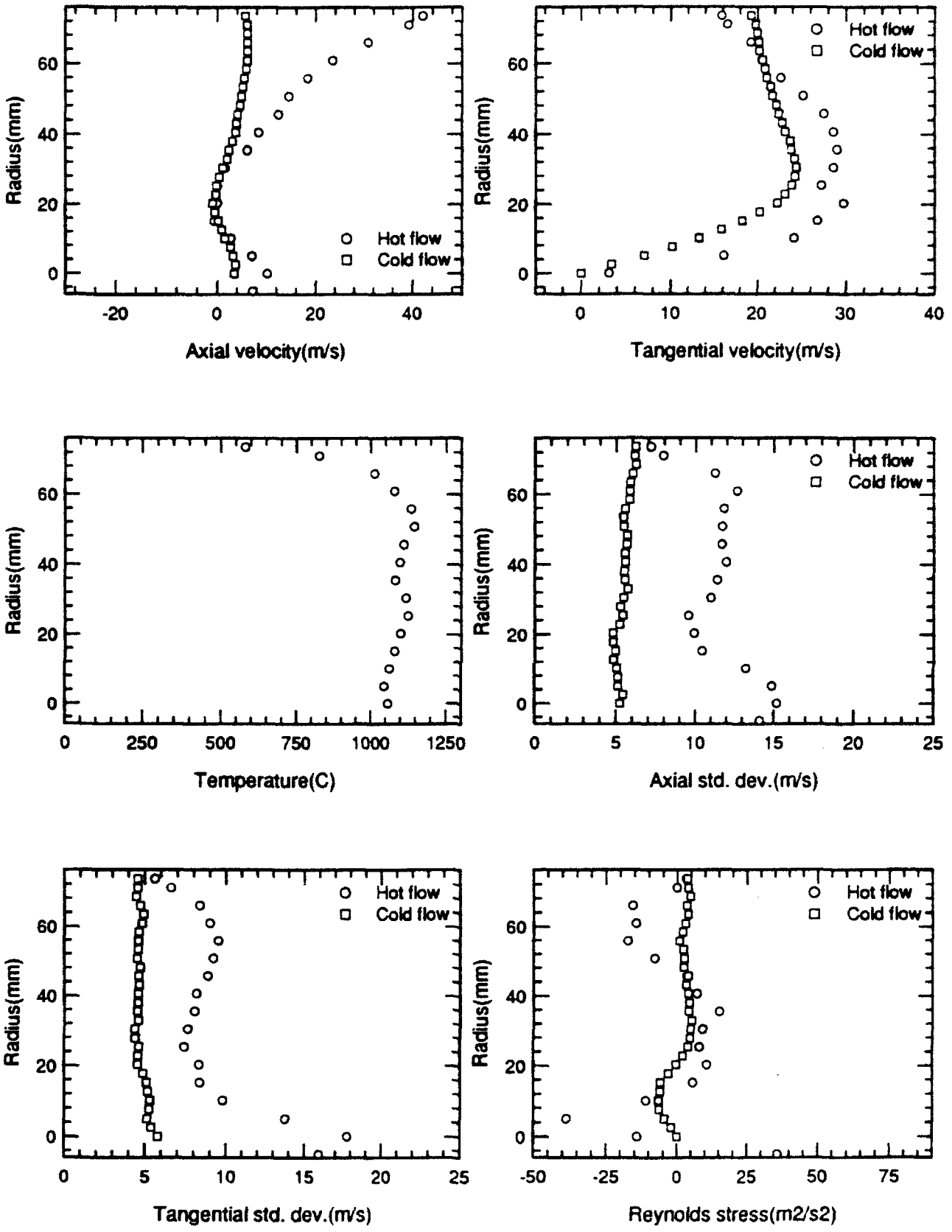


Figure 11. WL/RR Swirl Combustor, x = 152.4 mm (6 in.).

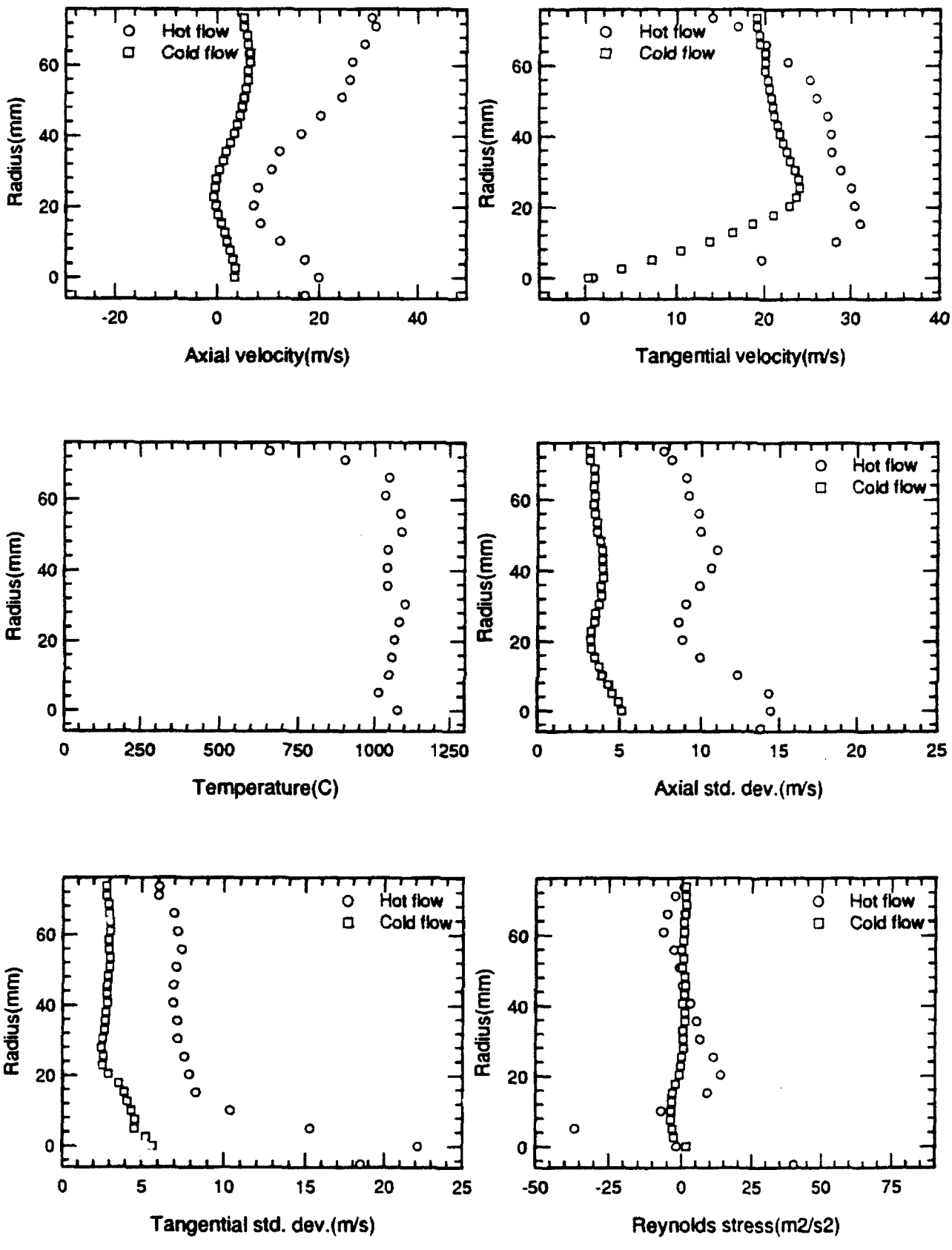


Figure 12. WL/RR Swirl Combustor, x = 203.2 mm (8 in.).

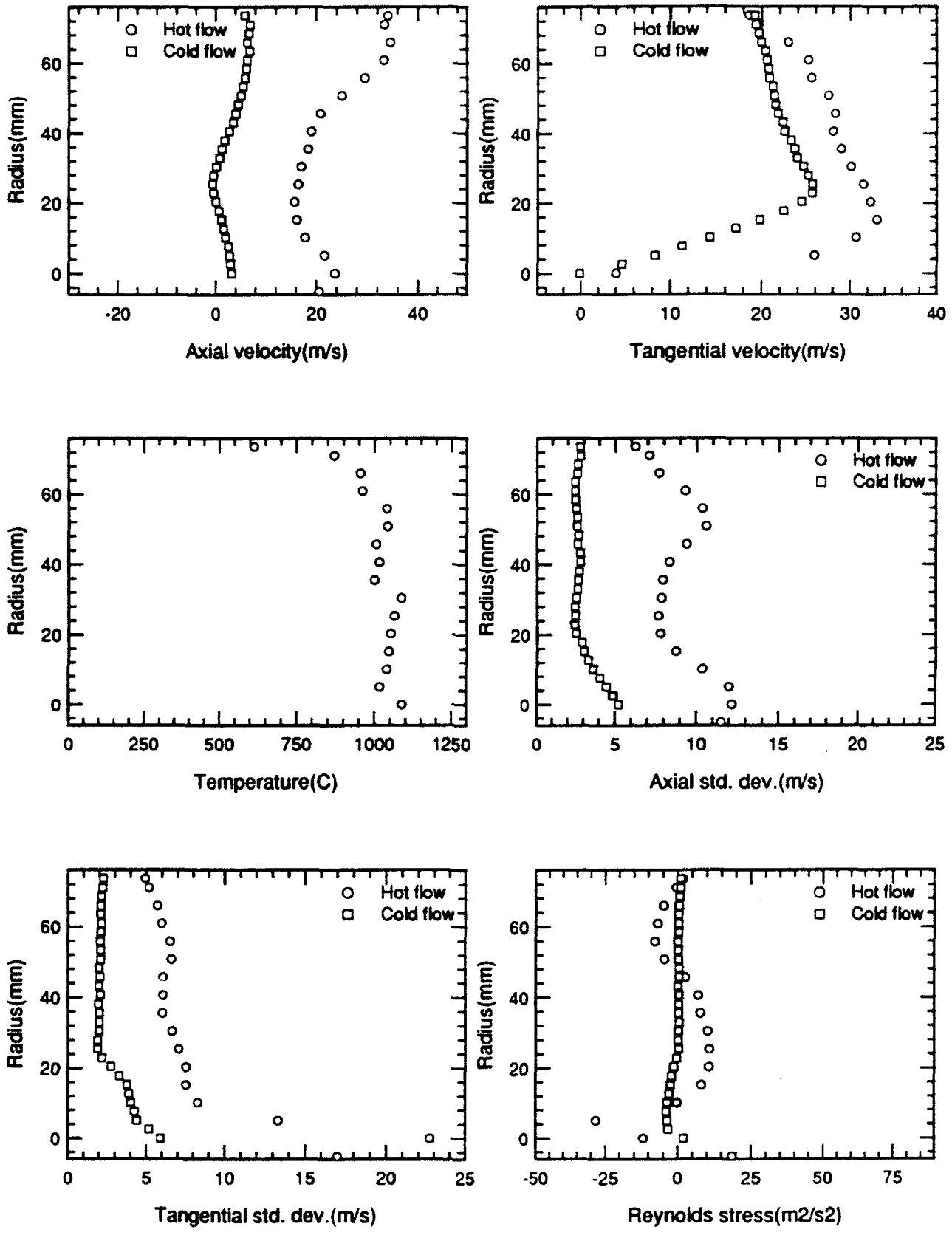


Figure 13. WL/RR Swirl Combustor, x = 254 mm (10 in.).

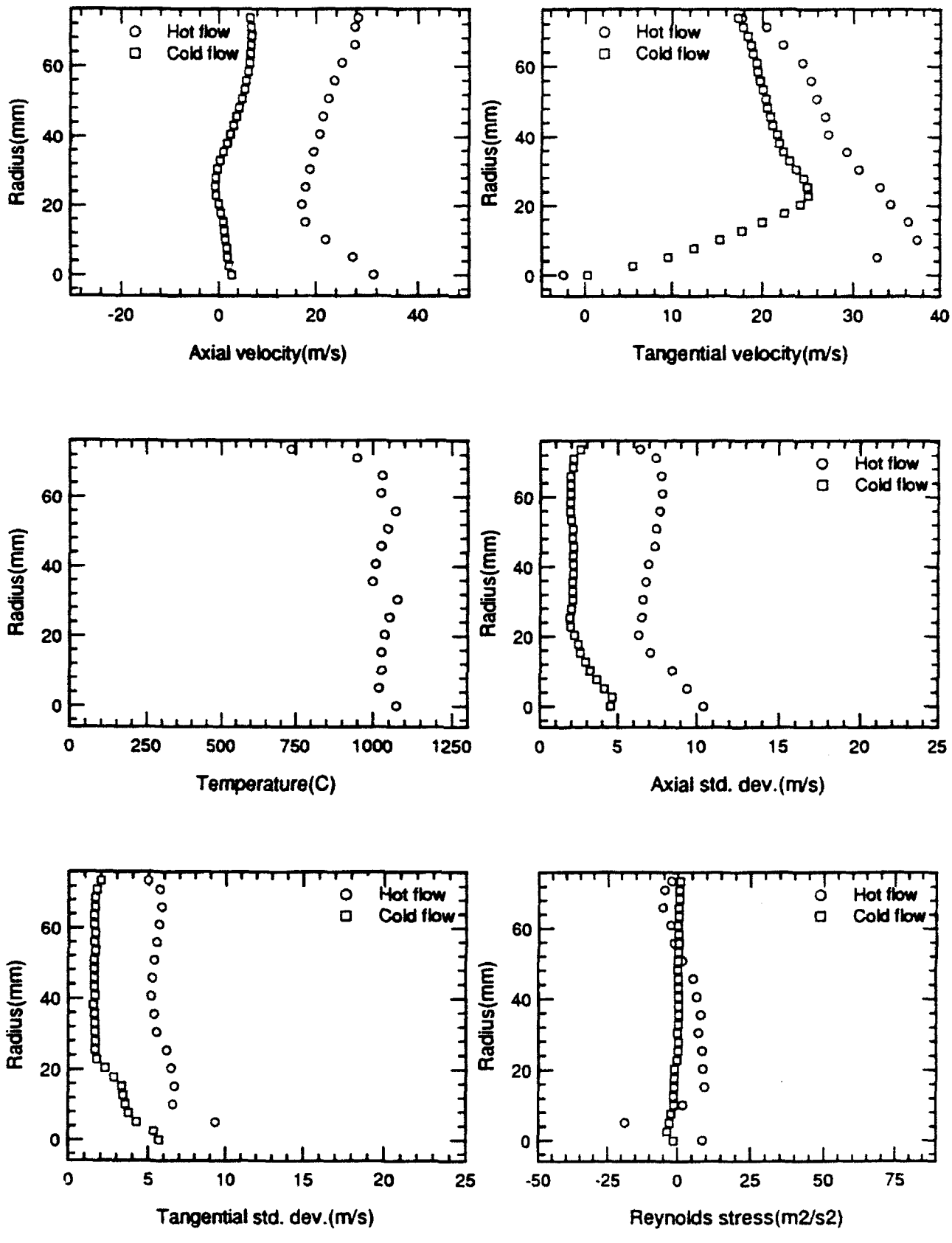


Figure 14. WL/RR Swirl Combustor, x = 304.8 mm (12 in.).

Effect of Aeroelasticity on the Measurement of the Indicial Response of an Airfoil

Gary M. Graham
Associate Professor
Department of Mechanical Engineering

Ohio University
261 Stocker
Athens, OH 45701

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

Effect of Aeroelasticity on the Measurement of the Indicial Response of an Airfoil

Gary M. Graham
Associate Professor
Department of Mechanical Engineering
Ohio University

Abstract

An analytical model for simulating the aeroelastic response of a tow tank test rig undergoing high inertial loading is described. The rig comprises an airfoil coupled to a drive shaft for pitching the airfoil, and whereon strain gauges are mounted for measuring the airfoil loading. The required input to the model is the torque imparted on the drive shaft for pitching the airfoil in a prescribed way. The model predicts the dynamic structural deformation of the test rig during the motion, as well as a "sensible" force at the strain gauges. The analytical sensible force is compared with actual strain gauge measurements for the case of the airfoil undergoing a small step change in angle of attack. Such a motion theoretically produces an airfoil indicial response. The analysis indicates that aeroelasticity has a significant effect on the strain gauge output. The present analysis combines the mode superposition method for dynamical systems undergoing forced motion with linear airfoil theory based on the convolution integral formulation for the loading on an airfoil in arbitrary motion.

Effect of Aeroelasticity on the Measurement of
the Indicial Response of an Airfoil

Gary M. Graham

INTRODUCTION

Background

The normal force response of a 2D NACA 0015 airfoil experiencing small and rapid "step" changes in angle of attack by rotation about the quarter chord has been measured recently in the Ohio University tow tank [1]. The motivation for these experiments was to study nonlinear airfoil behavior as defined in the theory of nonlinear mathematical modeling for aerodynamic systems [2]. Basically, the experiments involved strain gauge load cell measurements of the transient normal force loading on an airfoil undergoing a step change in angle of attack of approximately $\Delta\alpha = +1^\circ$ at a reduced pitch rate ($\dot{\alpha}b/U$, b=semichord) near 0.3. The angle of attack prior to the step onset (α_0) was steady and was varied over the range $2^\circ < \alpha_0 < 60^\circ$. In these experiments the test rig may experience high levels of inertial, as well as aerodynamic, loading due to the rapid starting and stopping required to impart the step. Therefore, an important issue is the degree to which aeroelastic reactions deform the structure thereby influencing the output of the strain gauge bridge. Knowledge of these reactions is useful in comparing these strain gauge data with classical indicial responses such as Wagner's function. This report describes an aeroelastic model of the O.U. tow tank load cell/airfoil test rig. The model is based on the mode superposition method for structural systems and classical linear airfoil theory.

Ohio University Tow Tank

A schematic of the O.U. tow tank is shown in Figure 1a. The facility consists of a large tank with a six inch chord NACA 0015 airfoil suspended vertically in the water with a submerged length of 42.0 inches. A carriage moves in translation along roller bearings fixed to I-beams spanning the tank at a speed of 2 ft/s, giving a Reynolds number near 95,000. The airfoil is driven in rotation via a drive shaft fixed to the airfoil quarter chord at one end, and coupled to a 3.5 hp stepper motor/gear box apparatus at the other end. Figure 1b is a schematic showing details of the drive shaft and airfoil. Shown here are dimensions in inches, and a numbering scheme (1 through 17) defined for the purpose of discretizing the mass of the structure as discussed in detail in a latter section. The drive shaft has a number of variations in cross section over the length of the shaft which must be considered in modeling the aeroelastic response of the structure. Near the middle of the drive shaft is a machined rectangular section for receiving a strain gauge load cell. The load cell is discretized into mass elements 5 through 9. The strain gauge circuit is electrically compensated to be sensitive to chord normal forces only. The upper most mass element 1 is made of steel while all other parts are aluminum. The modulus of elasticity (E) for steel was taken to be 28×10^6 psi, and for aluminum was measured to be 7.3×10^6 psi. The modulus of rigidity (G) was taken to be 10×10^6 psi and for aluminum 4×10^6 psi. The area moment of inertia of the airfoil about the chord line is 0.12 in^4 and the center of mass of the airfoil is located 1.335 inches aft of the pitch axis at the quarter chord. The polar moment of inertia of the airfoil about the pitch axis

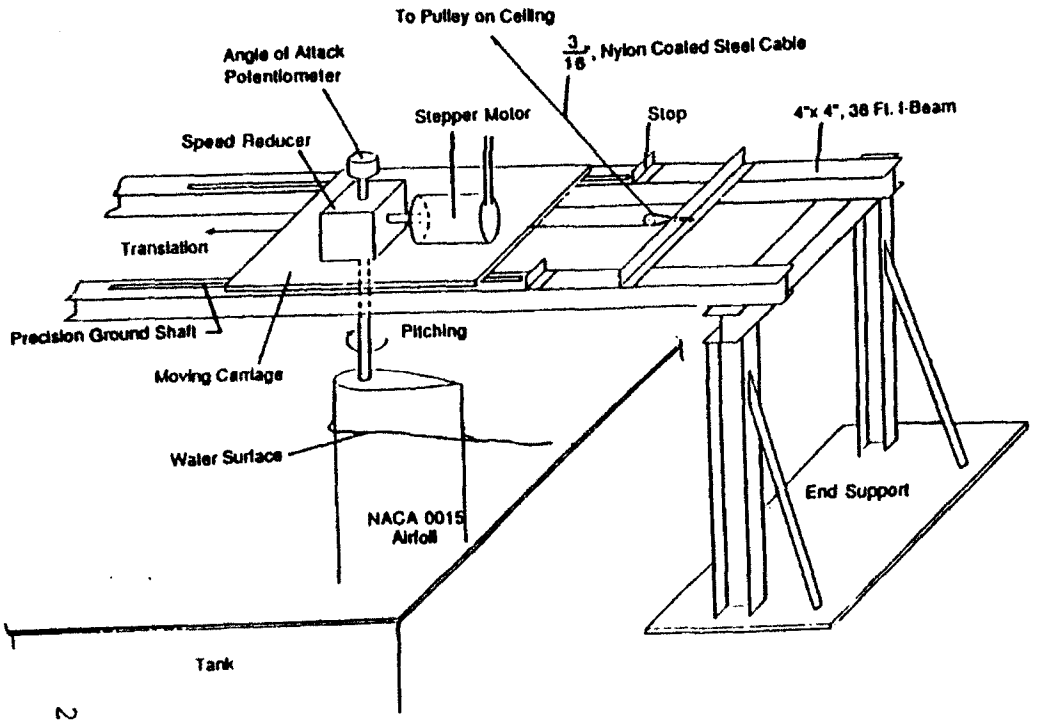


Figure 1a. Ohio University Tow Tank Facility.

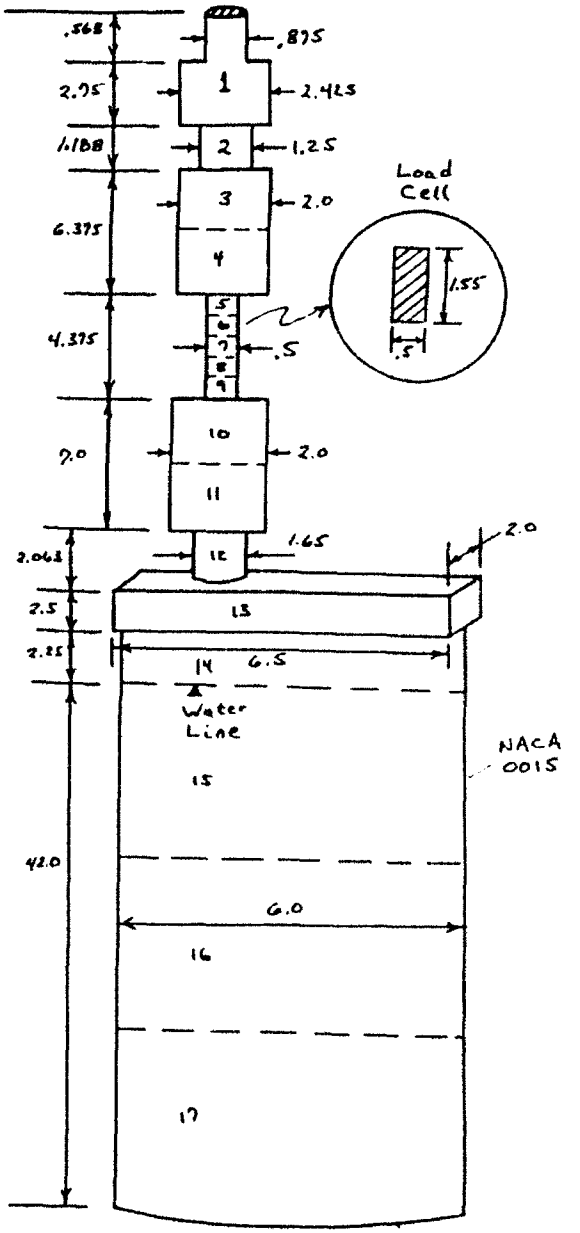


Figure 1b. Details of the Drive Shaft/Load Cell and Airfoil.

is 6.71 in^4. The densities of steel and aluminum were taken to be 0.3 and 0.1 lb_m/in^3. The mass of the airfoil per unit length is 0.1598 lb_m/in.

Classical Linear Normal Force Response

In an incompressible flow, the theoretical linear normal force coefficient response of a flat plate airfoil given an instantaneous step change in angle of attack by rotation about the quarter chord can be shown to be:

$$C_N(t) = C_{No} + \pi\Delta\alpha\,[\,\delta(t-0) + \frac{1}{2}\delta'(t-0)] + 2\pi\Delta\alpha\,[\,1 - 0.157e^{-.0455t} - 0.235e^{-.3t}\,] \qquad (1)$$

where C_{No} is the initial normal force, $\Delta\alpha$ is the step amplitude, and δ' is the time derivative of the Dirac delta function δ. The first two terms in brackets in Equation 1 are generalized functions [4] which describe the noncirculatory component of the loading, while the last group of terms is the circulatory component. The exponential terms arise from a two pole curve fit to Wagner's function [3]. Equation 1 is referred to as the indicial response (actually, the derivative of Eqn. 1 w.r.t. α) and is important in the convolution integral formulation for the loading on an airfoil in arbitrary motion.

Transient normal force responses of a NACA 0015 airfoil undergoing rapid small amplitude changes in angle of attack by rotation have been measured in the O.U. tow tank. Figure 2a shows angle of attack data for a typical run (small spikes are electrical noise). The onset angle is 2.09° and the step amplitude is approximately +1.3°. The motion resembles, to a reasonable approximation, a small amplitude ramp at a reduced rate near 0.3. Figure 2b is a comparison of the experimental normal force response to Equation 1. To facilitate the comparison, the coefficient 2π (for a flat plate airfoil) on the circulatory part in Equation 1 has been replaced by the experimentally measured local static normal force curve slope (for the present NACA 0015) measured at the onset angle in an independent test [1]. This substitution is necessary for the response of Equation 1 to approach the same steady state as the experimental response. The analysis which follows is designed to study the effects of aeroelasticity on the comparison of the experimental and theoretical responses of Figure 2b.

AEROELASTIC ANALYSIS

The present analysis is based on a combination of the mode superposition method for dynamic structural response to forced vibration, and linear airfoil theory formulated in terms of the convolution integral representation of the loading on an airfoil in arbitrary motion. The structure to be modeled has been shown in Figure 1b. Because the pitch axis does not coincide with the center of mass of the lower part of the structure (mass elements 13-17), it is necessary to consider the coupling between the normal and torsional vibration degrees of freedom.

System Representation

As illustrated in Figure 1b, the structure has been discretized into 17 mass elements which are considered to be concentrated at the centroid of each element. In the mode superposition method each mass communicates

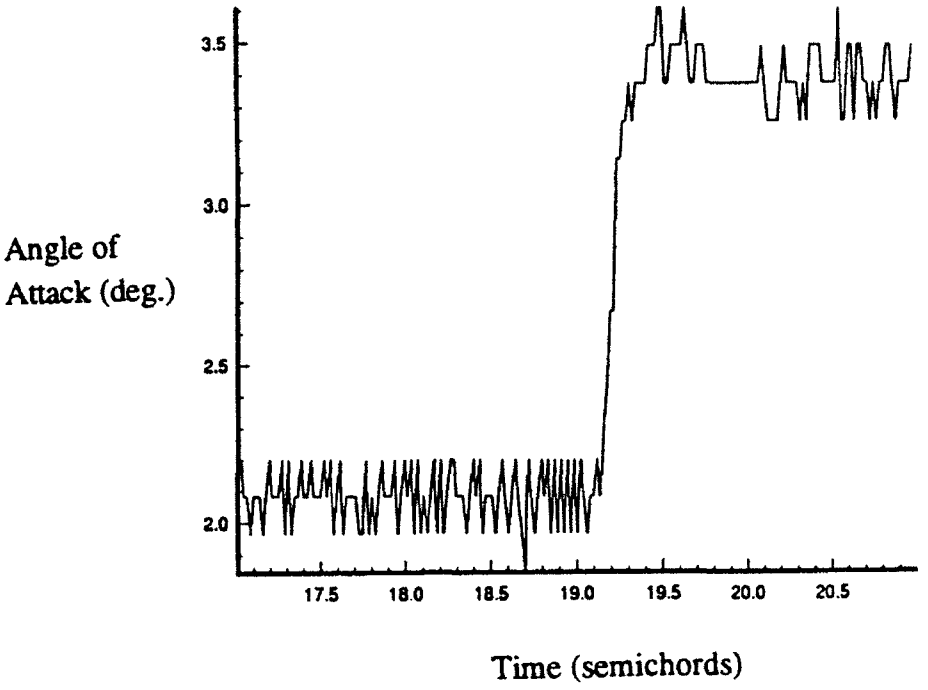


Figure 2a. Angle of Attack Data for Indicial Response Test.

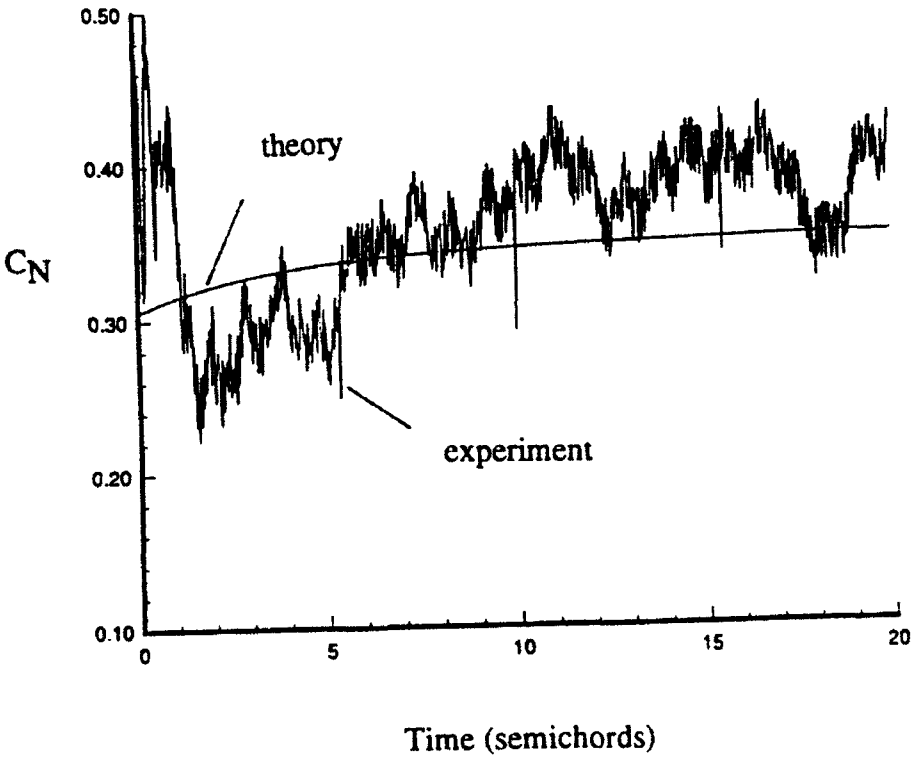


Figure 2b. Comparison of Normal Force Coefficient Data with Theoretical Respon

with neighboring masses through normal degree of freedom (NDOF) stiffness elements. The lowest three masses (15-17) represent the submerged portion of the airfoil, and masses 5 through 9 correspond to the rectangular cross section load cell. Masses are concentrated on the load cell to obtain good resolution of the deformation there, from which the "sensible" strain may be computed from the beam curvature. For the torsional degree of freedom (TDOF) the mass elements are replaced by polar mass moment of inertia elements which communicate with neighboring inertias through torsional stiffness elements. For the TDOF analysis, the structure is discretized in the same way as the NDOF for elements 1 through 12, however, elements 13 through 17 are lumped into a single inertia element giving a total of only 13 inertia elements. The reduction in elements in the TDOF is based on the fact that the torsional stiffness of both the mounting block (element 13) and airfoil are much larger than the drive shaft, and consequently the mounting block and entire airfoil experience nearly the same TDOF deflection. The resulting discretized structure may be described mathematically in terms of a diagonal mass matrix [M], a symmetric NDOF flexibility matrix [AN], a diagonal polar mass moment of inertia matrix [J], and a symmetric TDOF flexibility matrix [AT]. The flexibility coefficient AN(i,j) is by definition the deflection of mass i due to a unit force applied at mass j, while AT(i,j) gives the angular rotation of polar inertia i due to a unit torque applied to inertia j. The matrix [AN] has been computed using beam theory [5] and [AT] has been computed from elementary mechanics for shafts in torsion. The NDOF stiffness matrix [KN] and TDOF stiffness matrix [KT] are computed by inverting [AN] and [AT] and has been done using the commercially available code MATRIXX. Representative values for selected elements of the flexibility matrices are: AN(1,17) = 1.707E-4 in/lb_f , AT(1,13) = 1.017E-6 rad/lo_f-in. Values of the diagonal elements of [AN] and [AT] are given in Appendix A.

Eigenvalue Problem

The system matrices above have been used to solve the eigenvalue, or free vibration, problem to compute the natural frequencies and mode shape vectors of the system. This has been done in the usual way by assuming a harmonic solution in the unforced motion resulting in the frequency equations:

$$\left|\lambda_N[I] - [DN]\right| = 0 \tag{2a}$$

$$\left|\lambda_T[I] - [DT]\right| = 0 \tag{2b}$$

where the λ's are the inverse square of the natural frequencies given by $\lambda_N = 1/\omega_N^2$, $\lambda_T = 1/\omega_T^2$, and [I] is the identity matrix. The system matrices are given by [DN] = [AN][M] and [DT] = [AT][J]. The natural frequencies have been calculated using MATRIXX and the results for the two lowest natural frequencies for the NDOF and TDOF, respectively, are: ω_N = 7.91 hz and ω_T = 85.01 hz. Higher modal frequencies have been neglected in the present aeroelastic model. The justification for neglecting higher frequencies is based on two observations. First, oscilloscope traces of the strain gauge output, when the structure is excited with no water in the tank, indicate that the fundamental frequencies are dominant. Second, the computed response using fundamental frequencies only, simulates the experimental response with reasonable accuracy. Inclusion of higher frequencies increases the complexity of the algebra.

Given the natural frequencies, the normalized mode shape vectors corresponding to ω_N and ω_T are

computed using:

$$[\lambda_N[I] - [DN]]\{\phi_N\} = 0 \tag{3a}$$

$$[\lambda_T[I] - [DT]]\{\phi_T\} = 0 \tag{3b}$$

where $\{\phi_N\}$ and $\{\phi_T\}$ are the mode shape vectors in the NDOF and TDOF, respectively. These results are given in Appendix A, where $\{\phi_T\}$ is now extended to a 17x1 vector with ϕ_{T13} through ϕ_{T17} set to unity since, by assumption, the inertias associated with masses 13 through 17 undergo the same TDOF deflection. The resulting vector is interpreted as the TDOF mode shape of the inertias associated with the masses 1 through 17 illustrated in Figure 1b. This extension is necessary to make $\{\phi_T\}$ mathematically compatible with the 17x1 vector $\{\phi_N\}$ in certain matrix multiplications that arise in the aeroelastic vibration problem.

Structural Dynamics Model

The dynamic structural response has been modeled using concepts originating in the mode superposition method [6] and classical linear airfoil theory [3,5]. The present analysis, however, considers only the lowest (fundamental) frequency NDOF vibration mode and lowest TDOF mode and, as such, there is in reality no mode superposition. For modeling the solid body rotation of the test rig, a rotation DOF (RDOF) is introduced. The RDOF is used to simulate the change in angle of attack of the structure due to the rotation imparted by the stepper motor. Notice that the TDOF as defined above simulates only tosional deformations relative to the top of the structure.

In the NDOF and TDOF, the deflections of the structure are given by the normalized mode shape vectors multiplied by time dependent modal amplitudes. The total deflection of the structure is defined by:

$$\{v_N\} = \{\phi_N\}q(t) \tag{4a}$$

$$\{v_T\} = \{\phi_T\}g(t) \tag{4b}$$

$$\alpha = p(t) \tag{4c}$$

where $\{v_N\}$ and $\{v_T\}$ are the deflections in the NDOF and TDOF, respectively, and q, g and p are the time dependent amplitudes and are also the generalized coordinates of the system. The scalar RDOF motion variable α is simply the magnitude of the nominal angle of attack as the structure is pitched.

In the present analysis, the coordinate p allows the structure to undergo solid body rotation while the coordinates q and g measure deflections relative to the instantaneous position of the top of the structure (where the NDOF and TDOF deflections are zero). Substituting Equations 4a, 4b, and 4c into the system of equations of motion for the NDOF, TDOF, and RDOF yields, respectively:

$$[M]\{\phi_N\}\ddot{q} + [KN]\{\phi_N\}q = \{F\} \tag{5a}$$

$$[J]\{\phi_T\}\ddot{g} + [KT]\{\phi_T\}g = \{T\} \tag{5b}$$

$$J_T\ddot{p} + \pi L[\dot{p} + \tfrac{3}{8}\ddot{p}] = T_{IN} \tag{5c}$$

where the terms on the RHS are the forcing functions whose formulation is the subject of latter sections. Equation 5c represents a single equation in which J_T is the total mass moment of inertia of the structure, L is the submerged airfoil length, and T_{IN} is the input torque required to pitch the airfoil. In the experimental set up of Figure 1a, this would be the torque exerted by the stepper motor on the top of the airfoil drive shaft.

Equation 5a represents a system of 17 NDOF equations for the 17 NDOF masses (Figure 1b) and Equation 5b represents 13 TDOF equations for the 13 TDOF moments of inertia. Premultiplying Equation 5a by $\{\phi_N\}^T$ and Equation 5b by $\{\phi_T\}^T$ (before the extension of $\{\phi_T\}$ from a 13x1 to a 17x1 vector as described above) and performing the matrix multiplications results in the system of three generalized equations of motion [6]:

$$M\ddot{q} + KNq = F \tag{6a}$$

$$J\ddot{g} + KTg = T \tag{6b}$$

$$J_T \ddot{p} + \pi L[\dot{p} + \tfrac{3}{8}\ddot{p}] = T_{IN} \tag{6c}$$

where M ($=\{\phi_N\}^T[M]\{\phi_N\}$) is the generalized mass, KN ($=\{\phi_N\}^T[KN]\{\phi_N\}$) is the generalized NDOF stiffness, J ($=\{\phi_T\}^T[J]\{\phi_T\}$) is the generalized polar mass moment of inertia about the pitch axis, and KT ($=\{\phi_T\}^T[KT]\{\phi_T\}$) is the generalized TDOF stiffness. The forcing terms on the RHS of Equation 6a and 6b are the aeroelastic generalized force and moment given by $F = \{\phi_N\}^T\{F\}$ and $T = \{\phi_T\}^T\{T\}$. The solution of Equations 6a, 6b, and 6c for the generalized coordinates, q, g, and p subject to aeroelastic loading given by linear airfoil theory is described below.

Nondimensionalization and Sign Convention

In the generalized system of equations, the NDOF coordinate, q, has been nondimansionalized with respect to the airfoil semichord length, b (= 0.25 ft) and the dot superscripts indicate differentiation with respect to nondimensional time in semichords. The generalized TDOF coordinate, g, as well as the RDOF coordinate, p, are nondimensional by definition. Equation 6a is then rendered dimensionless by multiplying both sides by $(1/(\rho U^2 b^2)$, and Equations 6b and 6c are multiplied by $(1/(\rho U^2 b^3)$, where $\rho = 62.4\ lb_m/ft^3$, U = 2.03 ft/s. The resulting nondimensional generalized properties are M = 3.65, KN = 136.48, J = 5.93, and KT = 25644.0.

The sign convention adopted for the normal force and moment are the conventional ones used for airfoils in which a positive angle of attack (in the static sense) produces a positive normal force while a positive moment produces a nose-up torque. The generalized coordinates are defined such that a positive normal force tends to produce a positive change in q and its time derivatives (note this is the reverse of the usual definition used for airfoils in which the NDOF coordinate i.e. plunge is measured in the opposite direction from the normal force) while a positive moment tends to produce a positive change in g and p and their derivatives.

Aeroelastic Normal Forces

The normal force acting on the structure is decomposed into a rigid body force vector, $\{F_R\}$, associated with the RDOF, an aeroelastic force vector, $\{F_A\}$, due to time dependent aeroelastic deflections along the span of the

airfoil, and an inertial force vector, $\{F_I\}$, so that in Equation 6a: $\{F\} = \{F_R\} + \{F_A\} + \{F_I\}$.

The rigid body force vector is the ideal aerodynamic loading response to a step change in angle of attack due to rotation about the quarter chord. These aerodynamic forces are exerted only on the submerged part of the airfoil represented by masses 15, 16 and 17 of Figure 2a. For a step of magnitude $\Delta\alpha$ the vector $\{F_R\}$ is at any time, t, at or after the step given by (see also Equation 1):

$$F_{Ri} = 0 \qquad , i = 1, 2...,14$$

$$F_{Ri}(t) = \pi L_i \{ \dot{p}(t) + \frac{1}{2}\ddot{p}(t) + 2\int_0^t (\dot{p}(\tau) + \ddot{p}(\tau)) \Gamma(t-\tau)\, d\tau \} \quad , i = 15, 16, 17 \qquad (7)$$

where L_i is the length of airfoil element i divided by the semichord length (note $L_{15}=L_{16}=L_{17}=4.667$) and $\Gamma(t-\tau)$ is Wagner's function given by: $\Gamma(t-\tau) = 1 - .165\, e^{-.0455(t-\tau)} - .335\, e^{-.3(t-\tau)}$. Notice that $\{F_R\}$ is not a function of the generalized coordinates q and g since in the ideal response the loading is given by (rigid body) 2D airfoil theory alone. This does not say, however, that the present structure does not deform under this loading. Notice also that if p is a unit step function of amplitude $\Delta\alpha$ and L_i is unity, Equation 7 becomes identical to Equation 1.

The aeroelastic force vector, $\{F_A\}$, is also based on the convolution integral formulation for an airfoil in arbitrary motion, and again acts only on the submerged part of the airfoil. For the pitch axis at the quarter chord the loading is:

$$F_{Ai} = 0 \quad , i = 1, 2,...,14$$

$$F_{Ai}(t) = \pi L_i \{ -\phi_{Ni}\ddot{q}(t) + \phi_{Ti}(\dot{g}(t) + \frac{1}{2}\ddot{g}(t)) + 2\int_0^t [-\phi_{Ni}\ddot{q}(\tau) + \phi_{Ti}(\dot{g}(\tau) + \ddot{g}(\tau))]\, \Gamma(t-\tau)\, d\tau \} \qquad (8)$$

$$, i = 15, 16, 17$$

where ϕ_{Ni} is the value of the NDOF mode shape at element i, and ϕ_{Ti} is the value of the TDOF mode shape. The negative sign on the NDOF terms is due to the sign convention adopted.

The inertial loading arises from the fact that the centroids of masses 13 through 17 do not coincide with the pitch axis. Thus, angular acceleration about the pitch axis produces an inertial normal force which acts at the centroid. The inertial force vector is:

$$F_{Ii} = 0 \qquad , i = 1, 2,...,12$$

$$F_{Ii}(t) = m_i r_i \{ \phi_{Ti}\ddot{g}(t) + \ddot{p}(t) \} \quad , i = 13, 14,...,17 \qquad (9)$$

where m_i is the nondimensional mass and r_i is the distance from the pitch axis to the centroid divided by the semichord length. Notice in Equation 6c that $\ddot{p}$ is determined by the input torque T_{IN}, so that in Equation 9 the inertial loading increases with input torque.

Aeroelastic Moments

The TDOF system has been discretized into 13 inertia elements for the purposes of solving the eigenvalue problem and calculating the generalized inertia and TDOF stiffness. In the aeroelastic analysis the mode shape vector $\{\phi_T\}$ is now extended from a 13x1 vector to the 17x1 vector given in Appendix A. This requires that in the calculation of the generalized aeroelastic moment, the moment vector {T} be likewise extended to a 17x1. This extension does not affect the results for the TDOF generalized inertia and generalized stiffness of Equation 6b computed in the 13 element representation.

The moments acting on the inertias of Figure 1b are also decomposed into a rigid body moment vector, $\{T_R\}$, an aeroelastic moment vector, $\{T_A\}$, and an inertial moment vector, $\{T_I\}$, giving a total moment of: $\{T\} = \{T_R\} + \{T_A\} + \{T_I\}$. The rigid body moment and the aeroelastic moment act only on the submerged part of the airfoil. The expressions for these moments are simplified by the fact that the circulatory normal force acts at the quarter chord (for a flat plate airfoil) giving a zero moment arm in the present study. For a NACA 0015 airfoil the circulatory normal force (at $Re \sim 10^5$) acts within 2% fraction of chord from the quarter chord and on this basis has been neglected. The rigid body moment for rotation about the quarter chord is given by:

$$T_{Ri} = 0 \qquad , i = 1, 2,...,14$$
$$T_{Ri}(t) = -\pi L_i \{ \dot{p}(t) + \frac{3}{8}\ddot{p}(t) \} \quad , i = 15, 16, 17 \qquad (10)$$

The aeroelastic moment vector, $\{T_A\}$, acting on an airfoil in arbitrary motion with the pitch axis at the quarter chord is given by [3]:

$$T_{Ai} = 0 \qquad , i = 1, 2,...,14$$
$$T_{Ai}(t) = \pi L_i \{ \frac{1}{2}\phi_{Ni}\ddot{q}(t) - \phi_{Ti}(\dot{g}(t) + \frac{3}{8}\ddot{g}(t)) \} \quad , i = 15, 16, 17 \qquad (11)$$

The inertial moment vector, $\{T_I\}$, is given by:

$$T_{Ii} = 0 \qquad , i = 1,2,...12$$
$$T_{Ii}(t) = m_i r_i \phi_{Ni}\ddot{q}(t) - J_i \ddot{p}(t) \quad , i = 13, 14,...,17 \qquad (12)$$

where J_i is the (nondimensional) mass moment of inertia of element i. Notice again that the input torque, which determines $\ddot{p}$, influences the inertial loading.

The aeroelastic forces and moments have been formulated in terms of the three unknown generalized coordinates. The generalized force and moment may now be formulated by definition using the mode shapes. The resulting system of Equations 6a, 6b, and 6c may then be solved for q, g, and p subject to a prescribed input torque T_{IN}. As in the tow tank experiments described above, T_{IN} in essence "drives" the solution (or at least gets it started).

Input Torque

Of interest is the structural aeroelastic response to an ideal torque which, if the structure were perfectly rigid,

would produce a true step change in angle of attack; or in terms of the generalized coordinates, a step change in p. This torque is given by:

$$T_{IN}(t) = \Delta\alpha \left[\left(J_T + \frac{3}{8}\pi L\right)\delta'(t\text{-}0) + \pi L\delta(t\text{-}0) \right] \tag{13}$$

where $\Delta\alpha$ is step amplitude and δ' is the time derivative of the delta function δ. This fact can be deduced directly from Equation 6c. Substituting Equation 13 into Equation 6c and taking the Laplace transform [4] yields after rearranging: $P(s) = \frac{\Delta\alpha}{s}$. This result has the inverse transform $p(t) = \Delta\alpha\mu(t\text{-}0)$ where μ is the unit step function, which is the desired result. In the solution described below, Equation 13 has been used as the input torque. Notice, however, that any input torque, such as that actually imparted in the experiments or even harmonic motion, may be substituted into Equation 6c. For simulating the experimental torque wherein the variation in angle of attack is a small amplitude ramp (see Figure 2a), $T_{IN} = \dot{\alpha}\left[\left(J_T + \frac{3}{8}\pi L\right)(\delta(t\text{-}0) - \delta(t\text{-}\tau_s)) + \pi L(\mu(t\text{-}0) - \mu(t\text{-}\tau_s))\right]$, where $\dot{\alpha}$ is the pitch rate, and the ramp begins at $t = 0$ and ends at $t = \tau_s$. Using the data of Figure 2a yields approximately : $\dot{\alpha}b/U = 0.3$ and $\tau_s = 0.075$ semichords (= 0.009 sec in real time).

Solution of the Generalized Equations of Motion

Equations 6a, 6b, and 6c have been solved using the Laplace transform method to yield three algebraic equations which are linear in terms of the transformed generalized coordinates. The[illegible]uations are solved simultaneously for the transformed generalized coordinates as functions of the Laplace variable, s, alone. The inverse Laplace transform is performed to transform these solutions to the time domain to obtain p(t), q(t), and g(t). The Laplace transform of each of equations 6a, 6b, and 6c can be accomplished by hand using standard tables. To overcome the very cumbersome algebra required in the solution, the commercially available code MACSYMA has been used. MACSYMA is capable of symbolic mathematics required in the solution of the simultaneous equations in terms of the Laplace variable s. Other utility codes such as MATHMATICA also have this capability. The resulting equations of motion in the Laplace transform domain are quite long and are omitted here for brevity. Once the Laplace transforms have been done the solution proceeds by first rearranging, using MACSYMA, the system of three equations into the following form:

$$[C(s)]\{X(s)\} = \{R(s)\} \tag{14}$$

where [C(s)] is a 3x3 matrix and is a function of the Laplace variable s, and {X(s)} is given by:

$$\{X(s)\} = \begin{Bmatrix} Q(s) \\ G(s) \\ P(s) \end{Bmatrix} \tag{15}$$

where Q(s) is the Laplace transform of q(t), G(s) of g(t), and P(s) of p(t). The 3x1 vector {R(s)} contains all other terms which are not coefficients of Q, G, or P. The matrix [C(s)] is inverted symbolically to obtain the solution in the Laplace domain:

$$\{X(s)\} = [C(s)]^{-1} \{R(s)\} \qquad (16)$$

MACSYMA returns the solution for each generalized coordinate in the Laplace transform domain as a quotient in the form of a a sixth order polynomial in s divided by a seventh order polynomial in s. The inverse Laplace transform is accomplished by decomposing the quotients using partial fractions into functions which can be inverted by hand. In the partial fraction decomposition MACSYMA has been used to find the roots of the seventh order denominator. There are three real roots and two complex conjugate roots. The real roots give the rigid body response while the imaginary roots give the oscillatory response. The solution for each generalized coordinate in the time domain (t = nondimensional time based on semichords) is of the form:

$$A - Be^{-bt} - Ce^{-ct} + e^{-dt}[\, D\sin(\omega_1 t) + E\cos(\omega_1 t)\,] + e^{-ft}[\, F\sin(\omega_2 t) + G\cos(\omega_2 t)\,] \qquad (17)$$

where: b = 0.290179
c = 0.045387
d = 0.436706
f = 0.685316
ω_1 = 2.290063 (reduced frequency based on semichord = $\omega b/U$)
ω_2 = 47.1939 " " " "

The values of the other constants for each of the three generalized coordinates are given in the following Table.

Table 1. Constants in the Generalized Coordinate Solutions for an Instantaneous Step Input.

	q(t)	g(t)	p(t)
A	0.007404	-2.9937E-5	0.017453
B	-0.001866	-4.6738E-7	-2.0388E-8
C	-0.001197	-1.64493E-7	-3.3237E-8
D	0.013803	-1.4371E-5	-9.5866E-9
E	0.006941	1.9559E-5	-2.1811E-9
F	-7.8039E-4	-3.6745E-4	1.8772E-8
G	-0.01124	1.0682E-5	-2.8508E-7

where again it is noted that the results for q(t) are nondimensionalized based on the semichord and g and q are in radians. Notice that in Equation 17, ω_1 is associated with the NDOF and ω_2 is due to the TDOF.

Calculation of the Sensible Normal Force at the Load Cell

The purpose of this exercise is to determine the effects of aeroelasticity on the output of the strain gauge load cell used in the experiments described above. In these tests the strain gauge output is interpreted ideally as that due to a moment exerted by a normal force applied at the midspan of the submerged part of the airfoil (point

loads were applied here to calibrate the strain gauge bridge). In reality, however, this is not the case since the strain gauge output is determined solely by the local beam curvature at the cell. This curvature is due, not to aerodynamic loading alone, but rather the total aeroelastic structural response.

From elementary beam theory the moment at the center of the load cell (centroid of element 7) is related to the beam curvature by:

$$T_{LC}(t) = EI \frac{d^2\{v_N(t)\}}{dx^2}\Big)_7 \tag{18}$$

where x is measured along the span. The second derivative has been computed numerically from the NDOF deflections of elements 5 through 9 (inclusive) using a five point numerical derivative given by Richardson's extrapolation method [7]. A time dependent "sensible" normal force, N, acting at the center of the submerged portion of the airfoil midspan which produces the same moment as that given by Equation 18 may then be defined for comparison with experimental normal force data, which have been interpreted in the same way (i.e. as acting at the midspan). For this purpose, the moment arm between the mass 7 and the midspan is 37.0 in. These results are presented in the form of a normal force coefficient defined as:

$$C_N = C_{No} + \frac{N}{\rho U^2 b L} \tag{19}$$

where C_{No} is normal force at the origin of the motion, ρ is the density of water, and L is the submerged airfoil length (42.0 in), U and b have been given above.

RESULTS

The means chosen for comparison with experimental responses has been to cast the present analytical results in terms of a sensible normal force at the load cell as described in the previous section. The sensible force is determined by the input torque appearing in Equation 5c. The input corresponding to an instantaneous step change in angle of attack as well as a torque for producing a small amplitude ramp motion, as in the case of the tow tank experiments of Reference[1], are considered below.

Generalized Coordinates for an Instantaneous Step

Of interest is the aeroelastic response to a unit step input of amplitude $\Delta\alpha = +1^\circ$ at a step onset angle of zero. Notice that while the present model is sufficiently general to simulate any prescribed motion through an appropriate choice of input torque, it is equally possible to obtain the sensible force response to an arbitrary angle of attack variation through a convolution integral of the force response to a unit step input(assuming of course the structure is linear).

Figure 3a shows the RDOF generalized coordinate p(t). Ideally the variation in p(t) should be a unit step function with no oscillatory part, recall p(t) is simply the angle of attack. The oscillations in Figure 3a are very small and are due to numerical noise. The scale on the ordinate does not change due to the fact that variations in p(t) from the steady state are less than 1E-6 rad. (see Table 1) . The correct steady state value of

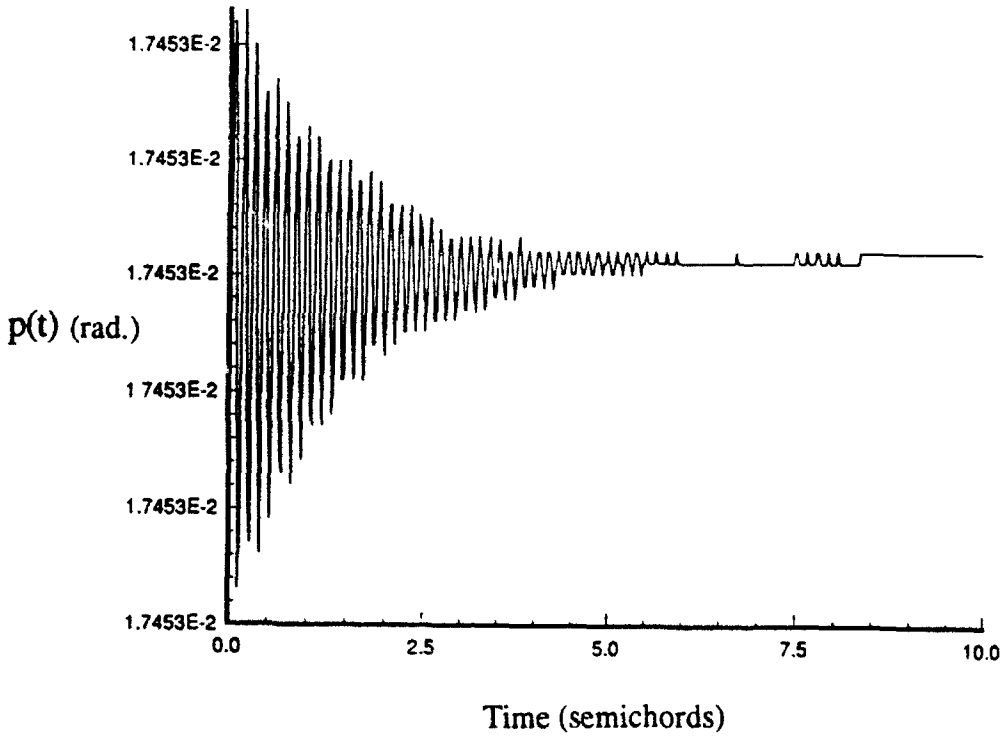


Figure 3a. Results for RDOF Generalized Coordinate.

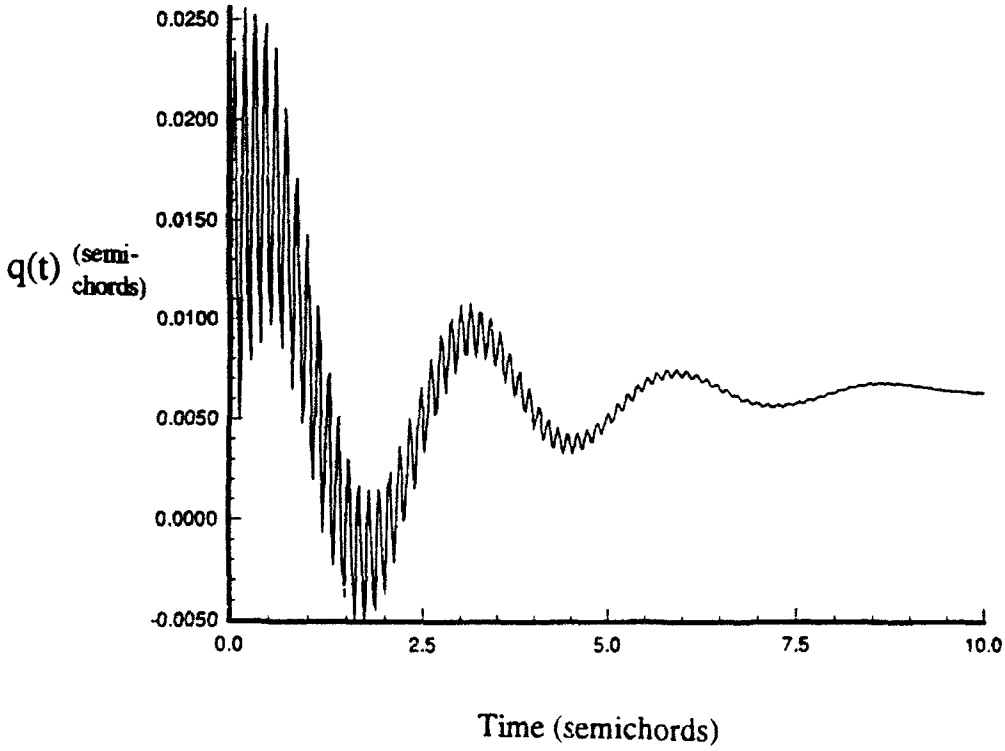


Figure 3b. Results for NDOF Generalized Coordinate.

0.01745 rad. (1°) is observed.

Figure 3b shows the results for the NDOF coordinate, q(t), expressed in fractions of semichord. The low frequency oscillation is due to direct aeroelastic reaction in the NDOF, while the high frequency oscillation is due to the coupling with the TDOF. Most of the oscillatory response has decayed beyond 2.5 semichords of travel after the step onset. The steady state value of q(t) determines the steady state normal force via Equation 18. As shown below, the steady value of q(t) yields the correct steady force.

The generalized coordinate for the TDOF is shown in Figure 3c. The magnitude of the TDOF oscillation is small and the proper steady value of zero is observed. Zero torsion in the steady state is due to the fact that the steady normal force acts at the quarter chord giving a zero moment arm on the present rig.

Sensible Force at the Load Cell for an Instantaneous Step

Figure 4 is a comparison of the theoretical response of Equation 1 and the aeroelastic sensible force of the present analysis. The results indicate that aeroelastic effects due to both the direct NDOF response (low frequency) and the NDOF response due to coupling with the TDOF (high frequency) are significant shortly after the step onset. The aeroelastic effects lead to a classical underdamped response and measurable differences with the theory. The capability of predicting the total structural response gives rise to the interesting possibility for developing a rational to correct experimental strain gauge data accordingly and extract therefrom the aerodynamic component of interest.

Comparison of the Sensible Normal Force with Strain Gauge Data

Figure 5a shows a comparison of the present analysis with strain gauge data taken in the Ohio University tow tank. These data are the same as in Figure 2b and are for a step height of 1.3° and a step onset angle of 2.09°. From static tests C_{No} (at 2.09°) in Equation 19 was found to be 0.22. The theoretical response for this case has been shown in Figure 2b. Also noteworthy is the fact that the present analysis assumes a steady state normal force curve slope of 2π /rad (see Equation 7) whereas the experimental value is near 8.0 /rad--this would become more apparent at larger values of time as steady state is approached. Nevertheless, there is sufficient agreement between the analysis and experiment to conclude that the rise in the strain gauge data shortly after the step, followed by an underdamped oscillation is due to aeroelasticity. Such a conclusion is not a priori possible since such behavior could very well arise from aerodynamics alone, particularly when considering that the noncirculatory loading at the step origin is theoretically large and positive. Figures 5b and 5c are magnified views of Figure 5a. In 5b below 1.0 semichords, the actual data indicate a lower frequency of vibration than the analysis. This might be expected since the actual structure is undoubtedly less stiff than the analytical model due to deformations in the I-beams, gear box, etc. The high frequency oscillations in the actual data appear to damp-out faster than the analysis which may be due to the fact that no structural damping is put into the analytical model.

Comment on Apparent Mass

Apparent mass and apparent inertia arise from second order noncirculatory terms in the normal force and moment. These terms introduce sufficient "added" mass and inertia to cause a significant shift in the vibrational frequencies of the structure. More specifically, the NDOF fundamental frequency is reduced from

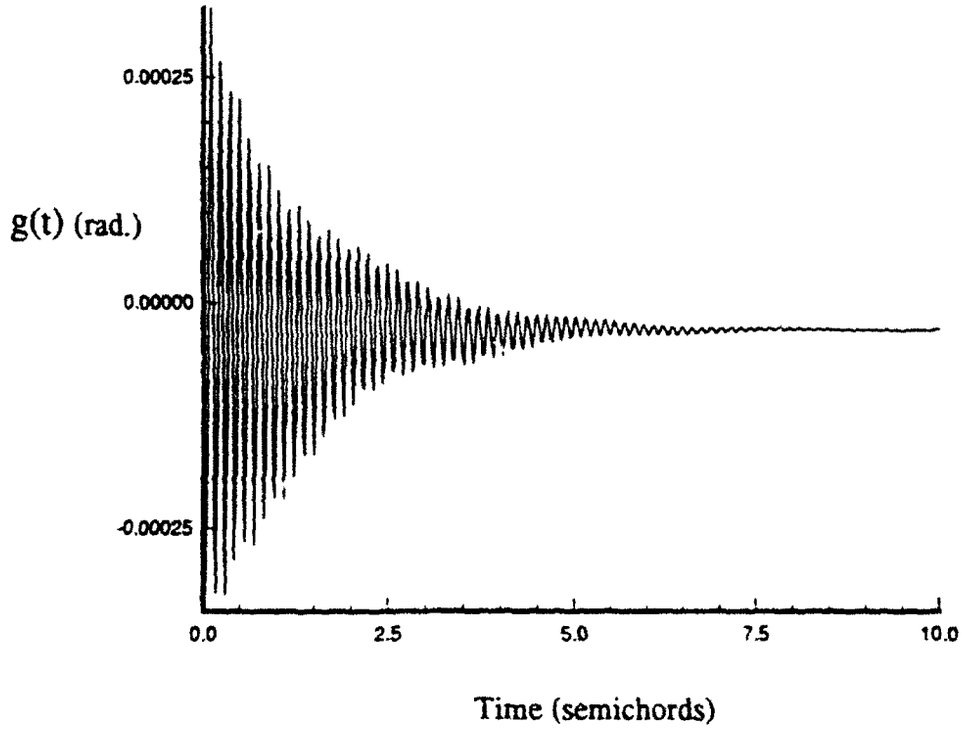


Figure 3c. Results for TDOF Generalized Coordinate.

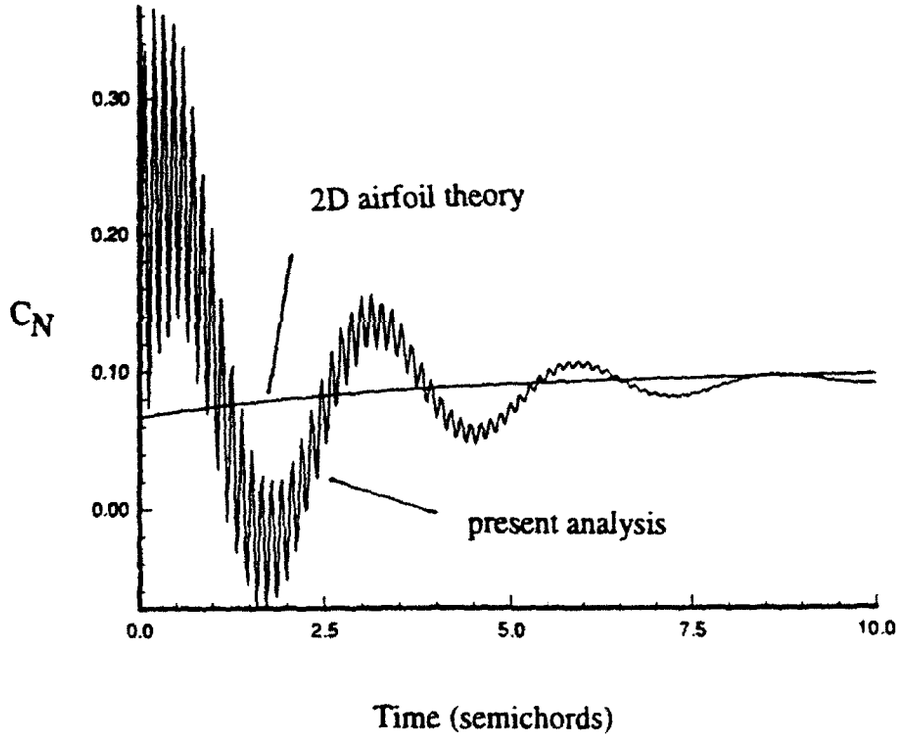


Figure 4. Comparison of the Aeroelastic Response for $\Delta\alpha$=+1° with the Theoretical Response (step onset angle = 0°).

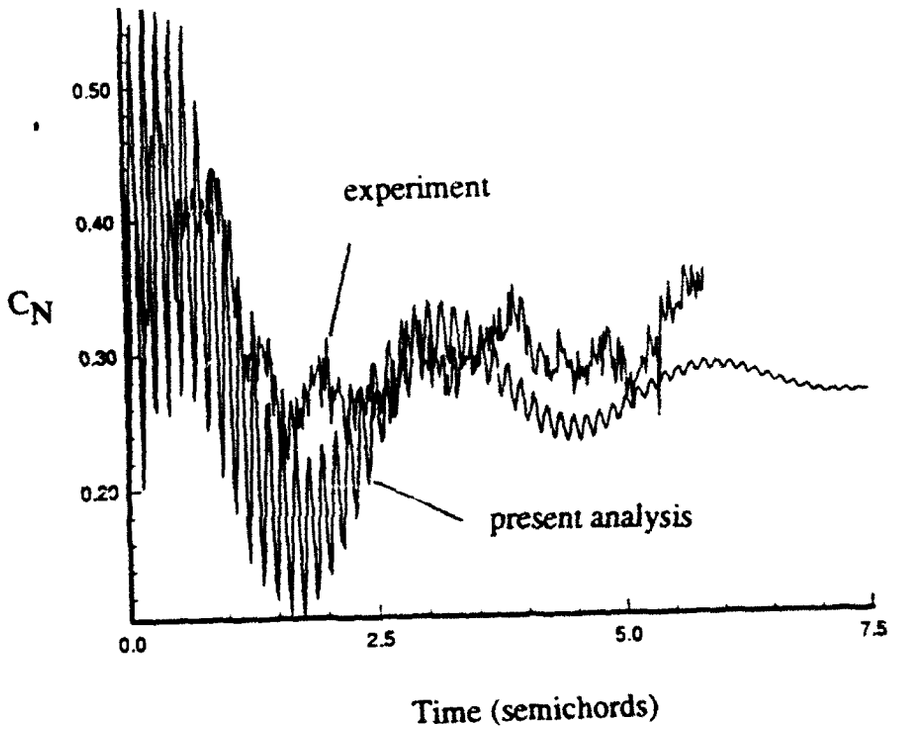


Figure 5a. Comparison of Normal Force Coefficient Predicted from Present Analysis wi
Experimental Data (step onset angle = 2°, $\Delta\alpha$ = 1.3°).

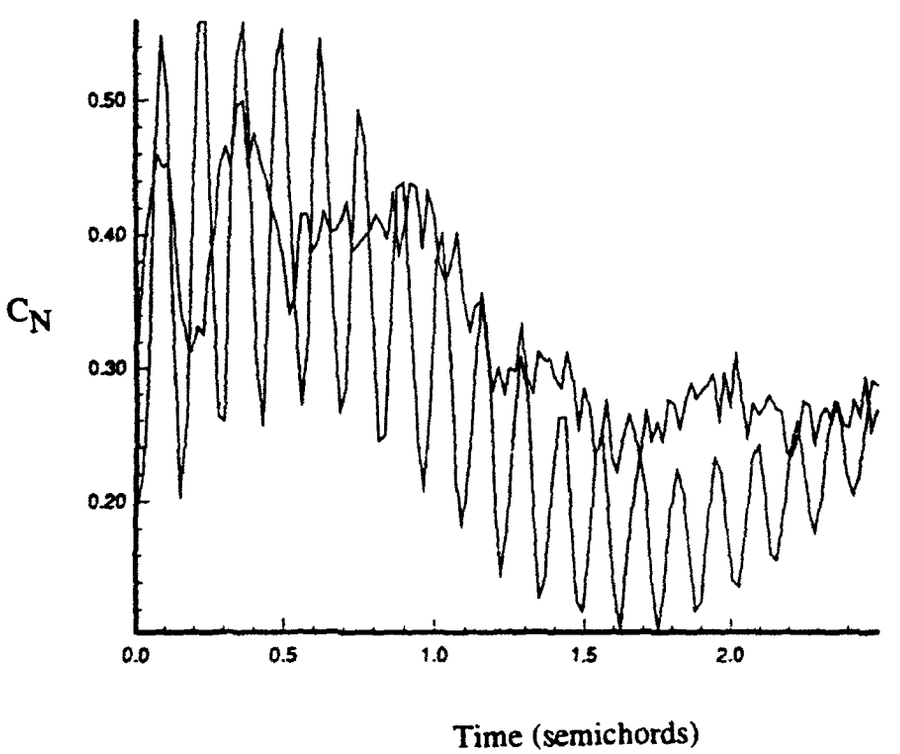


Figure 5b. Magnified View of Figure 5a.

7.91 hz (with no water in the tank) to 2.92 hz in the aeroelasticity problem, and the TDOF is reduced from 85.01 hz to 60.09 hz. It may be possible to use Fourier analysis to provide an indirect measurement of apparent mass. This may be a useful technique in the nonlinear aerodynamics problem where, to the author's knowledge, the effect of apparent mass has not been widely studied.

CONCLUSION

The present analysis provides a framework for studying aeroelastic effects in experimental nonlinear aerodynamics. Additional experiments are needed for direct measurement of system structural properties such as flexibility, natural frequency, and material properties for incorporating into the analysis. A goal of this work is ultimately to develop a method for correcting strain gauge data for aeroelasticity in the measurement of indicial responses.

ACKNOWLEDGEMENTS

The author expresses his gratitude to the engineers and staff of the Design Prediction Group at Wright Laboratory for their assistance and friendship during the 1992 AFOSR Summer Faculty Research Program.

REFERENCES

[1]Graham, G. M., Islam, M. and Jenkins, J. E.,"Nonlinear Normal Force Indicial Responses for a 2-D Airfoil," Proceedings AIAA Atmospheric Flight Mechanics Conf., New Orleans, La., Aug. 12-14, 1991 .

[2]Tobak, M. and Schiff, L.B., "Aerodynamic Mathematical Modeling-Basic Concepts," AGARD Lecture Series No.114, Paper 1, March 1981.

[3]Bisplinghoff, R.L. and Ashley, H., *Principles of Aeroelasticity*, Dover Publications Inc., New York, 1962.

[4]Kaplan, W.K., *Advanced Calculus*, Addison-Wesley Publishing Co., 2nd ed., pp. 515-521.

[5]Fung, Y.C, *An Introduction to the Theory of Aeroelasticity*, Dover Publications, New York, 1969.

[6]Clough, R.W. and Penzien, J., *Dynamics of Structures*, McGraw Hill Publisher, 1975.

[7]Burden, R.L., Farres, J.D., and Reynolds, A.C., *Numerical Analysis*, Prindle, Weber & Schmidt Boston, Ma., 1980, p.185.

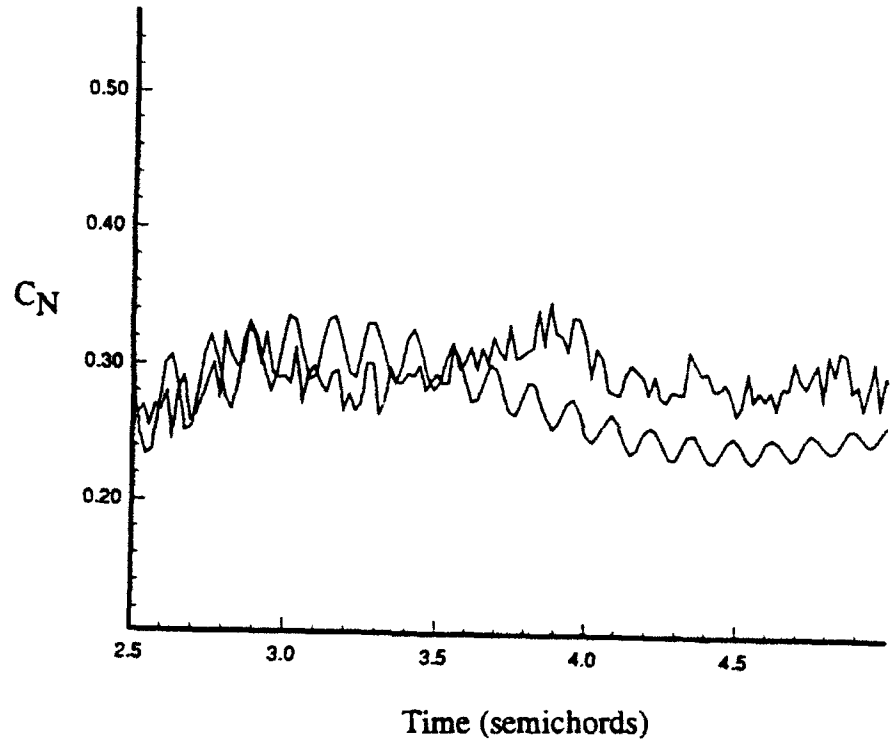


Figure 5c. Magnified View of Figure 5a.

APPENDIX A . Mode Shape Vectors and Diagonal Elements of Flexibility Matrices

Mass Element-i	$\phi_N(i)$	$\phi_T(i)$	AN(i,i) (in/lb$_f$)	AT(i,i) (rad/lb$_f$-in)
1	1.801E-3	3.923E-2	1.730E-6	1.017E-6
2	4.494E-3	6.470E-2	8.944E-6	1.678E-6
3	1.177E-2	9.832E-2	3.007E-5	2.551E-6
4	2.494E-2	0.1178	1.020E-4	3.058E-6
5	3.470E-2	0.2079	1.780E-4	5.402E-6
6	3.897E-2	0.3687	2.194E-4	9.584E-6
7	4.590E-2	0.5294	2.674E-4	1.377E-5
8	5.276E-2	0.6900	3.233E-4	1.795E-5
9	6.045E-2	0.8507	3.881E-4	2.213E-5
10	8.188E-2	0.9417	5.962E-4	2.450E-5
11	0.1182	0.9629	1.077E-3	2.506E-5
12	0.1486	0.9867	1.598E-3	2.569E-5
13	0.1751	1.0	2.124E-3	2.569E-5
14	0.2036	1.0	2.773E-3	———
15	0.3328	1.0	6.070E-3	———
16	0.6416	1.0	1.854E-2	———
17	1.0	1.0	2.775E-2	———

A STUDY OF VIRTUAL REALITY AND ITS APPLICATION TO AVIONICS

Elmer A. Grubbs
Assistant Professor
Department of Electrical Engineering

New Mexico Highlands University
Las Vegas, New Mexico 87701

Final Report for:
Summer Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

August 1992

A STUDY OF VIRTUAL REALITY AND ITS APPLICATION TO AVIONICS

Elmer A. Grubbs
Assistant Professor
Department of Electrical Engineering
New Mexico Highlands University

Abstract

A literature search on the topic of virtual reality was performed. Books and magazine articles on the topic were read and relevant ones saved. A review of work categories in Wright Laboratory's Avionics Directorate was performed and work areas were matched to possible virtual reality applications. A commercial program, "Virtual Reality Studio" was used to design several possible applications, including building design, visual programming, virtual office, and data visualization. A flicker free 3-D television system using standard television equipment with no hardware modification was designed, built and tested. Two patent disclosures were filed for 3-D television and computer monitor viewing.

A STUDY OF VIRTUAL REALITY AND ITS APPLICATION TO AVIONICS

Elmer A. Grubbs

INTRODUCTION

Virtual Reality (VR) provides the capability to simulate in three dimensions nearly anything imaginary or real. VR provides the user with the ability to interact with the simulated environment and other individuals using the simulation. VR simply involves computer software and input/output devices that directly interact with our senses, producing an artificial world for us to interact in. Depending upon the particular application involved, VR is also referred to as immersion simulation, artificial reality, telepresence, virtual world, and virtual environment.

The state of the art is such that the virtual world is not completely believable. The objects represented do not have intricate structure, and are very simple and cartoon-like. Interacting with the objects may not occur in real time, i.e. there is a noticeable delay in the movements represented in the simulation space. In addition, the VR equipment is bulky and expensive. There is little in the way of tactile feedback; for example one can't feel an object as it's picked up.

The limitations to VR are being worked on by several research and development organizations. The technology for correcting most of the VR limitations already exists. However, research has been conducted by a relatively small group of scientists, some who are not familiar with the breadth of the available hardware and software solutions. For example, position tracking has traditionally been done since the 1970s with a magnetic device used. This approach has a very slow response time. There is no technical barrier that limits the ability to sense position information orders of magnitude faster. Unfortunately, many researchers still use the magnetic devices not knowing that during '70s only magnetic devices were available but

today new technology devices exist and can be used. Except for the tactile feedback problem, the other limitations to using VR have solutions only a few years away.

There are many areas in the Wright Laboratory Avionics Directorate that can use VR technology. These areas will be addressed in the body of this report.

DISCUSSION

A detailed literature review was conducted during the first several weeks of this assignment. Searching Computer Select (CS), the Defense Technical Information Center (DTIC), the WPAFB technical library index, and the Colorado Alliance for Research Libraries (CARL) technical database resulted in a great deal of information. This information was organized and categorized. There are a broad range of issues, technical and otherwise, which are addressed in the literature. This report will concentrate only on the technical aspects of this emerging technology.

The hardware for a VR system consists of a computer system and input/output (I/O) devices. Input devices consist of devices such as joysticks, position sensors, television cameras, voice recognition systems, data gloves, and body suits. Output devices are devices such as a head mounted displays (HMD), television/computer screens, 3-D glasses, stereo sound equipment, haptic displays, and computer speech systems. In short, anything that makes the computer easier to interface with using any of the human senses can be considered an I/O device to a VR system.

The software for a VR system consists of programs such as 3-D rendering, position and view calculation, I/O device interfaces, algorithms for computing constraints caused by physical laws, and motion simulation. Please refer to figure one for a block diagram of a general VR system. At this time there are VR systems available using current technology prices costing about $50,000. This price is expected to drop significantly, even as the technology improves. For example, Texas Instruments (TI) is expected to soon introduce a consumer

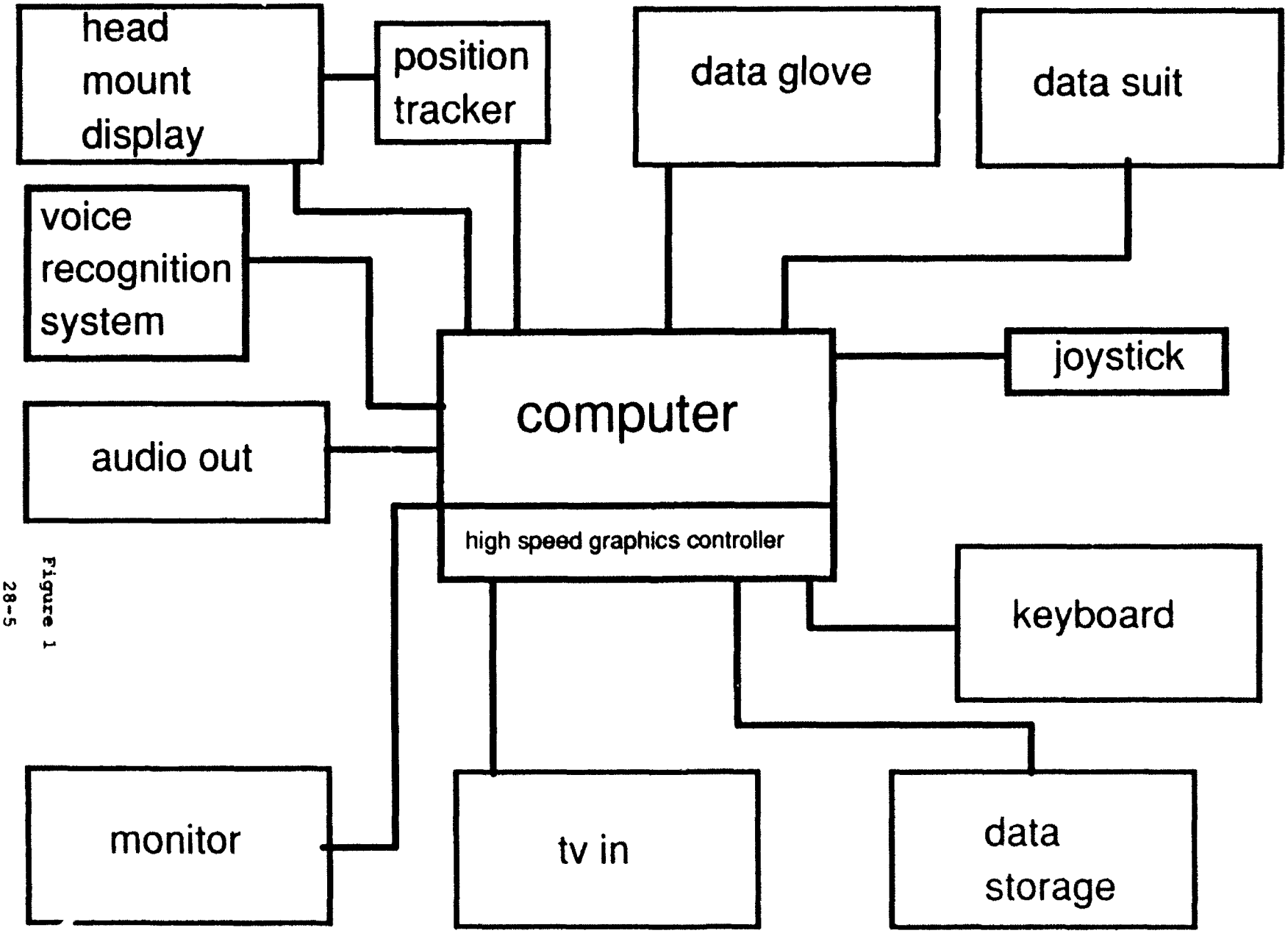

head mount display
position tracker
data glove
data suit
voice recognition system
computer
joystick
audio out
high speed graphics controller
keyboard
monitor
tv in
data storage

Figure 1
28-5

product costing about $400.

There are several interesting applications of VR that are in use or being designed at this time. Several are listed below, and Appendix 1 contains a more complete list. At the University of North Carolina a VR system was prototyped for simulating the interactions between molecules. This system allows scientists to place molecules together in the correct orientation to visualize the molecular bonding. In Japan, a VR system is in use that allows customers to design a kitchen themselves and then interact and visualize their design in a three dimensional simulation. Finally, the Department of Defense has contemplated a VR system that allows a simulation of a military battle. Multiple users fight the battle in computer generated tanks and aircraft.

Another aspect of virtual reality studied this summer is the concept of three-dimensional (3-D) television. Recently, with the advent of liquid crystal diode (LCD) glasses and the emergence of virtual reality, the concept of 3-D TV has been resurrected. One of the two popular approaches is as follows. Two cameras are used, one representing left eye information, and the other representing right eye information. The left eye information is sent to the TV during the first field of a frame, and the right eye information is sent during the second, or interlaced, field. During the time while the left eye information is present, the left lens of the LCD glasses is open and the right lens of the LCD glasses is closed. During the time while the right eye information is displayed, the right eye lens is open and the left eye lens is closed. Because of the fact that the field rate is 30 HZ, each LCD lens is turning on and off at 30 HZ. This causes an annoying flicker in the picture, which is unacceptable for general use. The second approach modifies the TV receiver so that the field rate is increased to 60 HZ, thus eliminating the flicker. This approach solves one problem, but introduces another since the TV receiver needs to be modified. The high cost of this modification and the additional circuitry eliminates this alternative approach for general public use.

During the summer, circuitry was designed and built that eliminates the

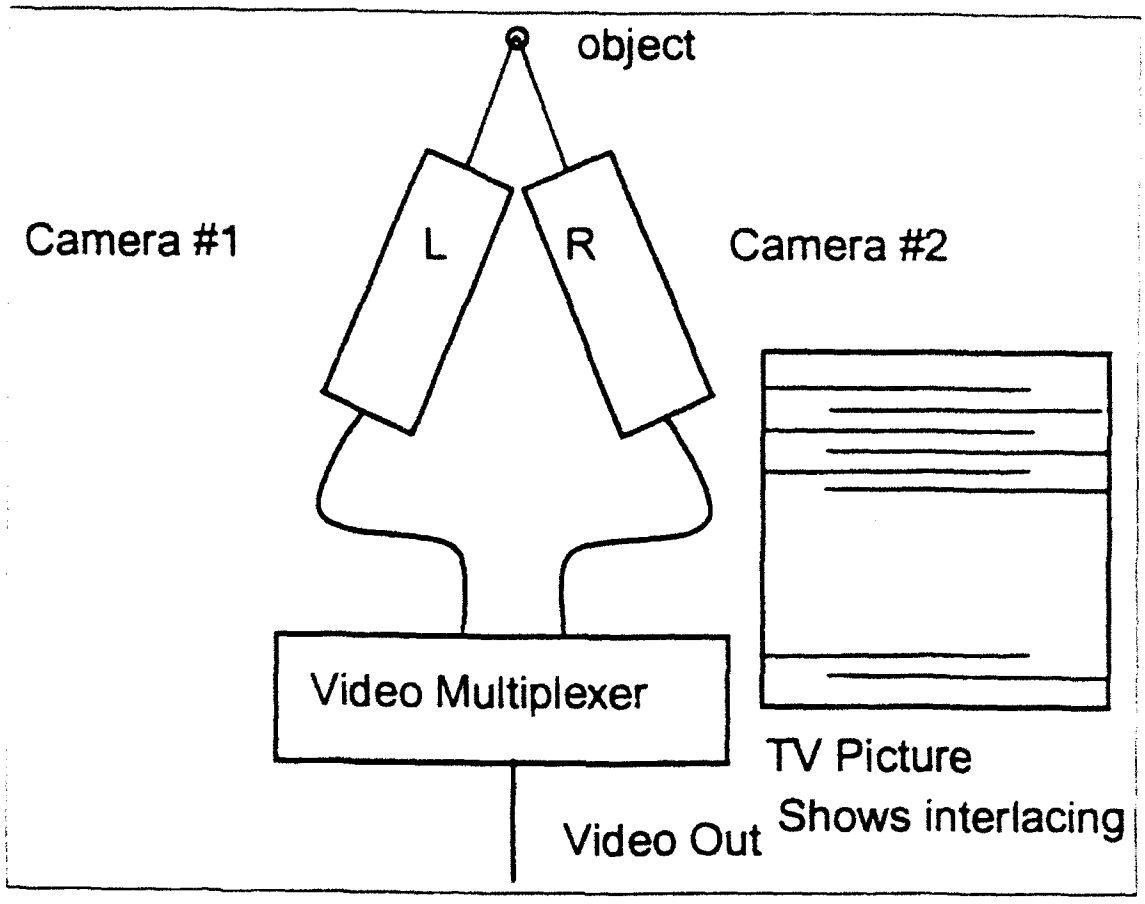
30 HZ flicker. This was accomplished with no internal modifications or connections to the TV receiver! This approach is a tremendous cost saver for VR researchers who use three dimensional TVs or computer monitors. In addition, this approach opens up a mass market in three dimensional commercial TV broadcasting using TVs already installed in the homes of viewers. Viewers would only need to purchase a pair of LCD glasses coupled with a small amount of electronic circuitry. Nothing needs to be done with the TV.

Recognizing the vast economic and military potential of this new approach to 3-D viewing of TV and computer monitors, the Air Force started the Patent Application process. Two Air Force Disclosure and Record of Invention Forms were submitted to the Air Force Patents Division. Please refer to figure 2 for a block diagram of the general approach used here. Further written information is not available at this time, due to the patent disclosure process.

The Avionics Directorate of the Wright Laboratory (WL/AA) is organized as three technology divisions, System Avionics (WL/AAA), Mission Avionics (WL/AAR), and electronic warfare (WL/AAW). In addition WL/AA has several support divisions such as Management Operations and Technical Services divisions and various support areas, including safety, security, supply, meteorology, etc. In the recommendations portion of this paper, I will outline several areas in which virtual reality could specifically be applied.

RESULTS

As has been previously mentioned, two invention disclosures on the subject of 3-D TV and computer monitor viewing have been submitted. Using the expansive resources of WL/AA, the inventions have been reduced to practice by their construction and demonstration. The demonstrations were so impressive that the Air Force is not hesitating in submitting the formal patent applications to the US Patent Office. The Air Force has recognized the importance of these inventions to both the military and private US industry. In short, the research conducted at WL/AA allow the owner of a personal


object
Camera #1
L
R
Camera #2
Video Multiplexer
TV Picture
Shows interlacing
Video Out

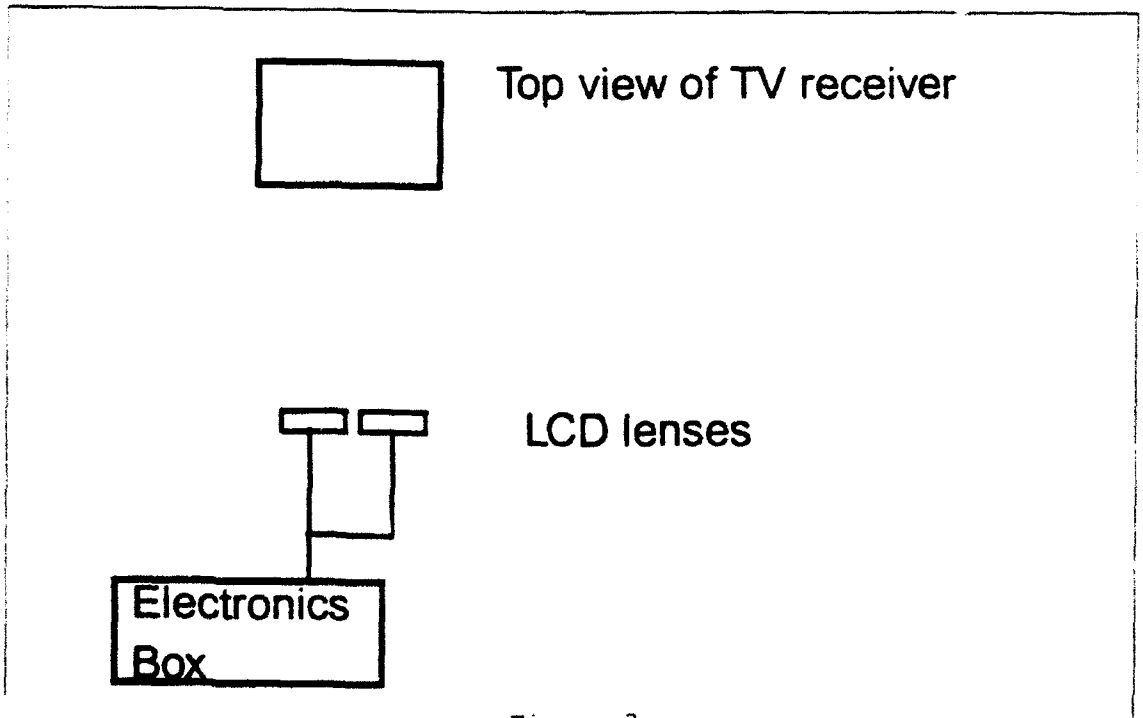

Top view of TV receiver
LCD lenses
Electronics
Box

Figure 3

computer system or a television set to view 3-D pictures, with no attachments or modifications to the existing monitor. All that is needed are LCD glasses coupled with inexpensive electronics. The 3-D pictures are free of the annoying flicker that usually accompanies 3-D pictures in these environments.

In addition to the 3-D VR work, a PC computer was programmed to illustrate several possibilities of VR. A commercial software product called "Virtual Reality Studio" was used. This inexpensive piece of software allows the user to set up his/her own virtual world. The user creates and edits three dimensional building blocks. For example, some of the building blocks consisted of rectangles, circles, cubes and pyramids of various sizes and colors.

The first program developed is a model of the Avionics Directorate's primary building, Building 620. This model allows the user to approach the building, enter any one of three doors, take the elevator to the third floor, look out the windows, enter offices and laboratories, or just walk around the building. Additionally the user can go outside of Bldg. 620 and enter a neighboring temporary building. This VR model simulates a three dimensional building in two dimensions, but a real virtual reality program would allow the user to actually feel as though he's entering and looking around the building. This model demonstrates the potential VR has for laboratory and building lay-out and design.

The second VR program developed simulates an office. Upon entering the office, the user sees a file cabinet, a bookcase with several books in it, a desk with a computer on it and a VCR and television set. The user can open the file cabinet and look at the files inside. He/She can take a book from the shelf and read it, or use the VCR to look at a helicopter. All of this represents data normally available on a computer, but allows the user to access the data in the same fashion that he/she would in the real world. There is no need to learn a series of special computer commands.

The third VR program represents a forest of data or measurements. The

user enters and sees a group of brown trees of various sizes. Looking from a distance he/she can see that the height of the trees generally is continuous from one side of the forest to the other side. However he/she notices that one tree is extremely small and another tree is very tall. The user can move closer and click on the small tree. A message comes up explaining what parameter the tree represents. This is useful for looking at a large amount of data, and seeing which values are not generally consistent. The user also notes different colored leaves on some of the trees. Moving over and clicking on a green leaf, he/she can see additional information concerning a parameter associated with the particular "tree."

The fourth program represents a visual programming language. This program demonstrates the possible of using VR in software design and post deployment support. The concept involves a user, who upon entering would see a collection of flow charting symbols, such as rectangles, squares, and decision diamonds. The software engineer/programmer lays out a logical flow chart for the function of the program by moving around and joining together the various symbols. The software engineer enters text into the symbols as required, then pulls a lever called create program. The program produces an output in the programmer's choice of language that performs the functions as outlined in the flow chart.

RECOMMENDATIONS

This sections outlines some the possible uses for VR in the Avionics Directorate. The section includes recommendations for using VR technology. The list presented here is not meant to be exhaustive, only representative of what is possible using this new technology.

Performance monitoring of system parameters: In software monitoring a person could be placed in a virtual world of trees representing the various parameters of the avionics software. As the aircraft went through various maneuvers, the parameters could be monitored to see if any went out of bounds during any part of the flight. As various conditions occurred, the leaves or

fruit on a tree could change color, alerting the user to these conditions.

Software development: Software could be developed using a visual 3-D programming language. The software engineer can arrange programming symbols in 3-D space to construct a program. In addition; the engineer can follow the flow of data through the software design, watching the data change size and shape as it is manipulated.

Navigation: A user could enter a virtual world representing the area above the earth where the aircraft was flying. The user could see "safe" corridors ahead indicating which direction to fly. Targets and threats would be represented in three dimensions augmented by other necessary data.

Packaging: Packaging of various instruments could be designed and viewed in cyberspace prior to actually assembling the device. This approach will result in errors being discovered ahead of time.

Avionics Simulation: VR can eliminate the need to have simulation domes, physical cockpits, and actual aircraft assets. The entire aircraft, its environment, and avionics can be modelled using VR. In addition, the physical test consoles and support equipment used in the simulation can be excluded by using VR technology.

Battle Management: VR technology can allow missions/battles to be viewed in virtual space. Commanders could view the action from different perspectives. For example, commanders will be able to view the battle from the perspective of the fighting units or from the enemy's point of view. Commanders could communicate with any of his assets by reaching out and "touching" the proper vehicle, unit, or aircraft.

Unmanned and Robotic vehicles: A pilot could fly an aircraft or cruise missile remotely using VR. With VR technology the pilot would be able to see and feel only a virtual representation of the cockpit, the instruments, and the view outside the vehicle.

Facilities: Office and lab space could all be designed in cyberspace ahead of time. In fact the entire facility/lab could exist in cyberspace, where workers could perform their duties remotely. This approach would

eliminate the time and money spent commuting with an automobile.

Because of the tremendous potential of this technology, I recommend WL/AAAF purchase a commercial (VR) development system. This system would serve as both an educational and prototyping medium. Individuals will be able to gain actual experience with VR technology and be able to test examples of how VR technology might be used in solving specific avionics problems.

CONCLUSIONS

A great deal of valuable research and development (R&D) work was accomplished this summer. The work allowed for a broad understanding of VR with respect to avionics. The R&D has led to new ideas in the development and application of VR technology. Two patent applications in the area of 3-D display of information have resulted from the R&D. VR technology can be applied to a wide variety of avionics design and post deployment activities. In fact, the development and application of VR technology can not only enhance America's military superiority, but can contribute greatly to both the commercial and academic sectors of the United States. The Avionics Directorate is in an excellent position to solve existing problems with VR and refine VR technology for transition to Air Force users. Therefore, I believe that it is critical that VR tools be introduced into the Avionics Directorate so that work can begin. VR technology is an important technology that neither the Air Force nor the US can afford to ignore.

CENTERS OF RESEARCH

Artificial Reality Corporation

Myron Krueger

Boeing Advanced Technology Center

Human Resources Engineering Division - Armstrong Laboratory - U.S. Air Force

MIT, MIT Media Lab

Marvin Minsky Thomas Sheridan

NASA - Ames Research Center

Stephen Ellis Robert Welch Elizabeth Wenzel

Telepresence Research

Scott Fisher Brenda Laurel

University of North Carolina - Chapel Hill

Frederick Brooks Henry Fuchs Warren Robinett

University of Washington - Human Interface Technology Lab

Meredith Bricken William Bricken Thomas Furness

VPL Research

Jaron Lanier

VENDORS

Ascension Technology Corp.	Motion Trackers	
P.O. Box 527 Burlington, VT 05402		(802) 655-7879
Covox, Inc.	Music/Voice I/O	
675 Conger St. Eugene, OR 97402		(503) 342-1271
Crystal River Engineering	High Speed DSP	
12350 Wards Ferry Rd. Groveland, CA 95321		(209) 962-6382
Digital Image Design, Inc.	3-D Computer Graphics	
170 Claremont Ave.,Suite 6 N.Y.,N.Y. 10027		(212) 222-5236
Dimension International	VR Computer System for PC's	
Berkshire, England		(44) 734 810077
Division Ltd.	VR System	
Bristol, England		(44) 454 324527
Evans & Sutherland	Computer Image Generator	
600 Komas Dr. Salt Lake City, UT 94158		(801) 582-5847
EXOS Inc.	Data Glove with Tactile Feedback	
8 Blanchard Rd. Burlington, MA 01803		(617) 229-2075
Fake Space Labs	Alternate to HMD/Remote Camera	
935 Hamilton Ave. Menlo Park, CA 94025		(415) 688-1940
Focal Point 3-D Audio	3-D sound	
1402 Pine Ave. Suite 127 Niagra Falls, NY 14301		(416) 963-9188
LEEP Systems, Inc.	HMD/Optics/Cameras	
241 Crescent St. Waltham, MA 02154		(617) 647-1395
Logitech	3-D Mouse/Tracker	
6505 Kaiser Dr. Fremont, CA 94555		(510) 795-8500
Polhemus Inc	Position Trackers	
P.O. Box 560 Colchester, VT 05446		(802) 655-3159

Sense8 Corp. 3-D Graphics
1001 Bridgeway, Suite 477 Sausalito,CA 94965 (415) 331-6318

Shooting Star Technology Position Tracker
1921 Holdom Ave. Burnaby,BC Canada V5B 3W4 (604) 298-8574

SimGraphics Engineering Corp. 3-D Mouse/Software
1137 Huntington Dr. South Pasadena, CA 91030 (213) 255-0900

StereoGraphics Corp. LCD Glasses 3-D Computer/Television
2171 E. Francisco Blvd. San Rafael, CA 94901 (415) 459-4500

StrayLight Corp. VR System for PC's
150 Mt. Bethel Rd. Warren, NJ 07059 (908) 580-0086

Telepresence Research Mobile Robot
635 High St. Palo Alto, CA 94301 (415) 325-8951

Virtual Technologies Data Glove/Suit/Feedback
P.O. Box 5984 Stanford, CA 94309 (415) 599-2331

The Vivid Group VR Software + Video

317 Adelaide St. W. Suite 302 Toronto,On Canada (416) 340-9290

VPL Research Inc. Data Glove/Suit/VR systems/HMD/3-D Sound
656 Bair Island Rd. Redwood City, CA 94063 (415) 361-1710

VREAM Inc. VR System for PC's
2568 N. Clark St. Suite 250 Chicago, IL 60614 (312) 477-0425

W Industries Ltd. VR Systems/Data Glove/Tactile Feedback/Etc
3 Oswin Rd. Leicester LE3 1HR, England (44) 533 542127

Xtensory Inc. Controller/Tactile Feedback/Software
140 Sunridge Dr. Scotts Valley, CA 95066 (408) 439-0600

EVENTS

Virtual Reality Showcase	February
Virtual Reality International	April
International Conference on Cyberspace	May
Medicine Meets Virtual Reality	June
SIGGRAPH	July
Virtual Reality	September
Calgary Virtual Reality Conference	October

References

1. AI EXPERT, August 1991, pp.26-39

2. AI EXPERT, August 1992, pp.22-29

3. AI EXPERT Special Report, July 1992

4. ASEE PRISM, "The Next Best Thing to Being There", May 1992, pp.26-29

5. BYTE, April 1992, pp.135-150, 175-182

6. Communications of the ACM, "Exploring Virtual Worlds with Tom Furness", July 1991

7. Computergram International, "Texas, Simgraphics, Abrams to Team on Home Virtual Reality System", March 25, 1992

8. Digital Media: A Seybold Report, "Stepping into Virtual Reality", Caruso, Aug 1991

9. Edventure Holdings, "Virtual Reality: Spreadsheets for Industry", Oct 8, 1990

10. EXE, "Virtual Worlds", Roth, A., Dec 1991

11. Patricia Seybold's Office Computing Report, "Cyberspace: Reality is No Longer Enough", Oct 1990

12. PIXEL, No. 6, 1991

13. Presence, MIT Press, Winter 1992

14. Science, "Looking Glass Worlds", Peterson, I., Jan 4, 1992, pp.8-15

15. Virtual Reality, Howard Rheingold, Summit, 1991

CURRENT AND FUTURE APPLICATIONS OF VR TECHNOLOGY

1. A system that allows a surgeon to enter a body virtually to perform an operation using microminiature surgical tools.

2. An air traffic control system that allows the operator to see the aircraft in three dimensions and communicate with an aircraft by reaching out in cyberspace and grabbing it.

3. Systems that allow a user to remotely control a robot or combat vehicle as though he was actually present at the site.

4. Tools which allow young students to design and interact with their own virtual worlds.

5. Games that allow multiple users to interact with each other.

6. Systems that allow a user to see in three dimensions the person they are talking to via telephone

7. A maintenance system that allows a technician to take three dimensional visual parts descriptions with him to work on an aircraft.

8. A medical system that allows a doctor to see inside a patient to locate tumors or other problems

9. A virtual office where records are stored in virtual file cabinets, books and reports in virtual bookcases, and videos and records can be played.

10. A virtual school where students attend from a distance without leaving their own homes