REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructi... the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, includi... Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budge...

AFRL-SR-BL-TR-00-

0733

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. R
	December, 1996	

4. TITLE AND SUBTITLE
1996 Summer Research Program (SRP), Summer Faculty Research Program (SFRP), Final Reports, Volume 5B, Wright Laboratory

5. FUNDING NUMBERS
F49620-93-C-0063

6. AUTHOR(S)
Gary Moore

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research & Development Laboratories (RDL)
5800 Uplander Way
Culver City, CA 90230-6608

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research (AFOSR)
801 N. Randolph St.
Arlington, VA 22203-1977

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The United States Air Force Summer Research Program (USAF-SRP) is designed to introduce university, college, and technical institute faculty members, graduate students, and high school students to Air Force research. This is accomplished by the faculty members (Summer Faculty Research Program, (SFRP)), graduate students (Graduate Student Research Program (GSRP)), and high school students (High School Apprenticeship Program (HSAP)) being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates, Air Force Air Logistics Centers (ALC), and other AF Laboratories. This volume consists of a program overview, program management statistics, and the final technical reports from the SFRP participants at the Wright Laboratory.

14. SUBJECT TERMS
Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student

15. NUMBER OF PAGES

16. PRICE CODE

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

UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1996

SUMMER FACULTY RESEARCH PROGRAM FINAL REPORTS

VOLUME 5B

WRIGHT LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way

Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major Linda Steel-Goodwin

Program Manager, RDL
Scott Licoscos

Program Administrator, RDL
Johnetta Thompson

Program Administrator, RDL
Rebecca Kelly

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1996

20010321 072

AFRL-01-06-1288

PREFACE

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

Due to its length, Volume 5 is bound in three parts, 5A, 5B and 5C. Volume 5A contains #1-24. Volume 5B contains reports #25-48 and 5C contains #49-70. The Table of Contents for Volume 5 is included in all parts.

This document is one of a set of 16 volumes describing the 1996 AFOSR Summer Research Program. The following volumes comprise the set:

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

SFRP FINAL REPORT TABLE OF CONTENTS i-xii

1. INTRODUCTION 1

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM 2

3. RECRUITING AND SELECTION 3

4. SITE VISITS 4

5. HBCU/MI PARTICIPATION 4

6. SRP FUNDING SOURCES 5

7. COMPENSATION FOR PARTICIPATIONS 5

8. CONTENTS OF THE 1996 REPORT 6

APPENDICIES:

A. PROGRAM STATISTICAL SUMMARY A-1

B. SRP EVALUATION RESPONSES B-1

SFRP FINAL REPORTS

SRP Final Report Table of Contents

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR Richelle M Allen-King	Washington State University , Pullman , WA Reduction Kinetics in a Batch Metallic Iron/Water System:Effect of Iron/Water Exposure	AL/EQC	2- 1
DR Anthony R Andrews	Ohio University , Athens , OH Investigation of the Electrochemiluminescent Properties of Several Natural & Synthetic Compounds	AL/EQC	2- 2
DR Deborah L Armstrong	Univ of Texas at San Antonio , San Antonio , TX Development of A primary Cell Culture Preparation for Studying Mechanisms Governi ng Circadian Rhyth	AL/CFTO	2- 3
DR Robert L Armstrong	New Mexico State University , Las Cruces , NM Microparticle Bioluminescence	AL/CFD	2- 4
DR Maureen E Bronson	Wilkes Univ School of Pharmacy , Wilkes-Barre , PA Lack of Effect of UltraWideband Radiation on Pentylenetetrazol-Induced Convulsions in Rats	AL/OER	2- 5
DR Marc L Carter, PhD, PA	University of South Florida , Tampa , FL Assessment of the Reliability of Ground-Based Observers for the Detection of Aircraft	AL/OEO	2- 6
DR Jer-Sen Chen	Wright State University , Dayton , OH A Study of Data Compression Based on Human Visual Perception	AL/CFHV	2- 7
DR Cheng Cheng	Johns Hopkins University , Baltimore , MD Sequential Optimization Algorithm for Personnel Assignmt Based on Cut-Off Profiles & Rev of Brogden	AL/HRM	2- 8
DR Elizabeth T Davis	Georgia Institute of Tech , Atlanta , GA Perceptual Issues in Virtual Environments & Other Simulated Displays	AL/CFHP	2- 9
DR Keith F Eckerman	Univ of Tennessee , Knoxville , TN	AL/OEB	2- 10
DR Paul A Edwards	Edinboro Univ of Pennsylvania , Edinboro , PA A Viartion Fuel Identification- Neural Network Analysis of the Concentration of Benzene and Naphtha	AL/EQC	2- 11

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR Randolph D Glickman	Univ of Texas Health Science Center , San Antonio , TX A Study of Oxidative Reaactions Mediated by Laser-Excited Ocular Melanin	AL/OEO	2- 12
DR Ellen L Glickman-Weiss	Kent State University , Kent , OH The Effect of Short Duration Respiratory Musculature Training on Tactical Air Combat	AL/CFTF	2- 13
DR Irwin S Goldberg	St. Mary's Univ of San Antonio , San Antonio , TX Development of a Physiologically-Based Pharmacokinetic Model for the Uptake of Volatile Chcicals dur	AL/OES	2- 14
DR Robert J Hirko	University of Florida , Gainesville , FL Investigation of The Suitability of Tactile and Auditory Stimuli for use in Brain Actuated Control	AL/CFHP	2- 15
ISU VPP Acct4212313(Dooley)	Iowa State University , Ames , IA Determination of the Influence of Ultrawideband Exposure of Rts During Early Pregnancy on Pregnancy	AL/OER	2- 16
DR Andrew E Jackson	Arizona State University , Tempe , AZ A Description of Integrated Joint Use Initiatives to Satisfy Customer Reqmts Across Govt Academia	AL/HRA	2- 17
DR John E Kalns	Ohio State University , Columbus , OH	AL/AOHR	2- 18
DR Nandini Kannan	Univ of Texas at San Antonio , San Antonio , TX Modeling Decompression Sickness Using Survival Analysis Techniques	AL/CFTS	2- 19
DR Antti J Koivo	Purdue Research Foundation , West Lafayette , IN Skill Evaluation of Human Operators	AL/CFBA	2- 20
DR Suk B Kong	Incarnate Word College , San Antonio , TX Aromatic Hydrocabon Components in Diesel, Jet-A And JP-8 Fuels	AL/OEA	2- 21
DR Xuan Kong	Northern Illinois University , De Kalb , IL Mental Workload Classification via Physiological Signal Processing: EOG & EEG Analyses	AL/CFHP	2- 22

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR Charles S Lessard	Texas A & M Univ-College Station , College Station , TX Preliminary Studies of Human Electroencephalogram (EEG) Correlates of GzAcceleration Tolerance	AL/CFTO	2- 23
DR Audrey D Levine	Utah State University , Logan , UT Biogeochemical Assessment of Natural Attenuation of JP-4 Contaminated Ground Water	AL/EQC	2- 24
DR David A Ludwig	Univ of N.C. at Greensboro , Greensboro , NC The Illusion of Control & Precision Associated w/Baseline Comparisons	AL/AOCY	2- 25
DR Robert G Main	Cal State Univ, Chico , Chico , CA Designing Instruction For Distance Learning	AL/HRT	2- 26
DR Phillip H Marshall	Texas Tech University , Lubbock , TX Time to Contact Judgments in The Presence of Static and Dynamic Objects: A Preliminary Report	AL/HRM	2- 27
MS Sandra L McAlister	Stonehill College , North Easton , MA	AL/AO	2- 28
MR Bruce V Mutter	Bluefield State College , Bluefield , WV Environmental Cost Analysis:Calculating Return on Investment for Emerging Technologies	AL/EQP	2- 29
DR Sundaram Narayanan	Wright State University , Dayton , OH Java-Based Application of the Model-View-Controller Framwork in Developing Interfaces to interactive	AL/HRT	2- 30
DR Karl A Perusich	Purdue University , South Bend , IN Examing Alternate Entry Points in a Problem Using Fuzzy Cognitive Maps	AL/CFHI	2- 31
DR Judy L Ratliff	Murray State Univ , Murray , KY A Study of The Ability of Tunicates to be used as Global Bioindicators	AL/EQC	2- 32
DR Paul D Retzlaff	Univ of Northern Colorado , Greeley , CO Computerized Neuropsychological Assessment of USAF Pilots	AL/AOCN	2- 33

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR William G Rixey	University of Houston , Houston , TX The use of Solid-Phase Microextraction (SPME) for the low level Detection of BTEX and PAHs In Aqueou	AL/EQC	2- 34
DR Ali M Sadegh	CUNY-City College , New York , NY Investigation of Neck Models for Predicting Human Tolerance to Accelerations	AL/CFBE	2- 35
DR Kandasamy Selvavel	Claflin College , Orangeburg , SC Truncated Bivariate Exponential Models	AL/AOEP	2- 36
DR Barth F Smets	University of Connecticut , Storrs , CT Biodegradation of 2-4-DNTand 2,6-DNT in Mixed Culture Aerobic Fluidized Bed Reactor and Chemostat	AL/EQC	2- 37
DR Mary Alice Smith	University of Georgia , Athens , GA A Study of Apoptosis During Limb Development	AL/OET	2- 38
DR Daniel P Smith	Utah State University , Logan , UT Bioremediation & its Effect on Toxicity	AL/EQW	2- 39
MR. Joseph M Stauffer	Indiana State University , Terre Haute , IN Joint Corrections for Correlation Coefficients	AL/HRMA	2- 40
DR William B Stavinoha	Univ of Texas Health Science Center , San Antonio , TX Studies to Identify Characterisstic Changes in the Urine Following Ingestion of Poppy seed	AL/AOT	2- 41
DR William A Stock	Arizona State University , Tempe , AZ Application of Meta-Analysis to Research on Pilot Training	AL/HRA	2- 42
DR Nancy J Stone	Creighton University , Omaha , NE Engagement, Involvement, and Self-Regualted Leaarnign Construct and Measurement Development to Asses	AL/HRT	2- 43
DR Brenda M Sugrue	Univ of Northern Colorado , Greeley , CO Aptitude-Attribute Interactions in Test Performance	AL/HRTI	2- 44

SRP Final Report Table of Contents

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR Stephen A Truhon	Winston-Salem State University, Winston-Salem, NC Mechanical Specialties in the U.S. Air Force: Accession Quality & Selection Test Validity	AL/HRM	2 - 45
DR Mariusz Ziejewski	North Dakota State University, Fargo, ND Validation of the Deformable Neck Model for A +Gz Acceleration	AL/CFBV	2 - 46

SRP Final Report Table of Contents

Author	University/Institution Report Title	Phillips Laboratory Directorate	Vol-Page
DR Graham R Allan	New Mexico Highlands University, Las Vegas, NM Temporal and Spatial Characterization of a Synchronously-Pumped Periodically-Poled Lithium Niobate Optical	PL/LIDN	3 - 1
DR Brian P Beecken	Bethel College, St. Paul, MN Testing of a Dual-Band Infrared Focal Plane Array & An Infrared Camera Sys	PL/VTRP	3 - 2
DR Mikhail S Belen'kii	Georgia Inst of Technology, Atlanta, GA Tilt Sensing Technique w/Small Aperture Beam & Related Physical Phenomena	PL/LIG	3 - 3
DR Asoke K Bhattacharyya	Lincoln University, Jefferson City, MO Part A: Effect of Earth's Surface & Loss on the Resonant Frequencies of Buried Objects	PL/WSQ	3 - 4
DR Joseph M Calo	Brown University, Providence, RI Transient Studies of the Effects of Fire Suppressants in a Well-Stirred Combustor	PL/GPID	3 - 5
DR James J Carroll	Youngstown State University, Youngstown, OH Examination of Critical Issues in the use of (178) hf For High Energy Density Applications	PL/WSQ	3 - 6
DR Soyoung S Cha	Univ of Illinois at Chicago, Chicago, IL A Study on Hartmann Sensor Application to Flow Aero-Optics Investigation Through Tomographic Recons	PL/LIMS	3 - 7
DR Tsuchin Chu	Southern Illinois Univ-Carbondale, Carbondale, IL	PL/RKS	3 - 8
DR Kenneth Davies	Univ of Colorado at Boulder, Boulder, CO Studies of Ionospheric Electron contents and High-Frequency Radio Propagation	PL/GPIM	3 - 9
DR Judith E Dayhoff	Univ of Maryland, College Park, MD Dynamic Neural Networks: Prediction of an Air Jet Flowfield	PL/LIMS	3 - 10
DR Ronald R DeLyser	University of Denver, Denver, CO Analysis of Complex Cavities Using the Finite Difference Time Domain Method	PL/WSTS	3 - 11
DR Andrew G Detwiler	S Dakota School of Mines/Tech, Rapid City, SD Evaluation of Engine-Related Factors Influencing Contrail Prediction	PL/GPAB	3 - 12
DR Itzhak Dotan	The Open University of Israel, Tel-Aviv Israel Studies of Ion-Molecule Reaction Rates at Very High Temperatures	PL/GPID	3 - 13

Author	University/Institution Report Title	Phillips Laboratory Directorate	Vol-Page
DR Omar S Es-Said	Loyola Marymount University, Los Angeles, CA On the Matis Selection of Durable Coatings for Cryogenic Engineer Technology	PL/RKE	3 - 14
DR Jeffrey F Friedman	University of Puerto Rico, Mayaguez, PR Testing the Frozen Screen Model of Atmospheric Turbulence	PL/LIMI	3 - 15
DR John A Guthrie	University of Central Oklahoma, Edmond, OK Ultrawide-Band Microwave Effects Testing on an Electronic System	PL/WSMA	3 - 16
DR George W Hanson	Univ of Wisconsin - Milwaukee, WI A Volumetric Eigenmode Expansion Method for Dielectric Bodies	PL/WSQ	3 - 17
DR Mayer Humi	Worcester Polytechnic Inst., Worcester, MA Wavelets and Their Applications to the Analysis of Meteorological Data	PL/GPAA	3 - 18
DR Christopher H Jenkins	S Dakota School of Mines/Tec, Rapid City, SD Shape Control of An Inflated Thin Circular Disk	PL/VT	3 - 19
DR Dikshitulu K Kalluri	University of Lowell, Lowell, MA Electromagnetic Wave Transformation in a Two-Dimensional-Space-Varying and Time-Varying Magnetoplasm	PL/GPIA	3 - 20
DR Aravinda Kar	University of Central Florida, Orlando, FL Thick Section Cutting w/Chemical Oxygen-Iodine Laser & Scaling Laws	PL/LIDB	3 - 21
DR Spencer P Kuo	Polytechnic University, Farmingdale, NY Theory of Electron Acceleration by HF-Excited Langmuir Waves	PL/GPI	3 - 23
DR Andre Y Lee	Michigan State University, East Lansing, MI Characterization Methods for Adhesion Strength Between Polymers & Ceramics	PL/RKS	3 - 24
DR Bruce W Liby	Manhattan College, Riverdale, NY Acousto-Optic Retro-Modulator	PL/VTRA	3 - 25
DR Feng-Bao Lin	Polytechnic Inst of New York, Brooklyn, NY Structural Ballistic Risk Assessment-Fracture Modeling	PL/RKEM	3 - 26
DR M Arfin K Lodhi	Texas Tech University, Lubbock, TX Theory, Modeling & Analysis of AMTEC	PL/VTP	3 - 27

SRP Final Report Table of Contents

Author	University/Institution Report Title	Phillips Laboratory Directorate	Vol-Page
DR Ronald A Madler	Embry-Riddle Aeronautical University, Prescott, AZ Estimating the Area of Artificial Space Debris	PL/WSAT	3 - 28
DR Carlos A Ordonez	University of North Texas, Denton, TX Boundary Conditions at A Plasma-Facing Surface	PL/WSQA	3 - 29
DR Michael J Pangia	Georgia Southwestern Coll, Americus, GA Further Analysis of Kilohertz Order Waves Associated with Electron Beam Operations on STS46	PL/GPSG	3 - 30
DR Ronald M Pickett	University of Lowell, Lowell, MA Temporal-Displacement Stereograms of the Ionosphere: An Exploration of Their Utility in the Analysis of Equatorial Emission Depletion Bands	PL/GPIA	3 - 31
DR Edgar Sanchez-Sinencio	Texas A&M Univ-College Station, College Station, TX Low Voltage Analog Circuit Design for Radiation Tolerance	PL/VTER	3 - 32
DR Joseph C Slater	Wright State University, Dayton, OH Smart Structure/Actuator Modeling 7 Design for the Integrated Ground Demonstration Lab	PL/VTI	3 - 33
DR Ashok Srivastava	Louisiana State University, Baton Rouge, LA Modeling of Total Dose Response of SOI N-MOSFETS for Low Power CMOS Circuits	PL/VTER	3 - 34
DR James M Stiles	University of Kansas, Lawrence, KS The Potential Applications of Super-Resolution & Array Processing to Space-Based Radars	PL/VTRA	3 - 35
DR Charles M Swenson	Utah State University, Logan, UT Balloon Launch Retromodulator Experiment	PL/VTRA	3 - 36
DR Miguel Velez-Reyes	University of Puerto Rico, Mayaguez, PR Regularization Methods for Linear and Nonlinear Retrieval Problems	PL/GPAS	3 - 37

SRP Final Report Table of Contents

Author	University/Institution Report Title	Rome Laboratory Directorate	Vol-Page
DR A F Anwar	University of Connecticut, Storrs, CT A Study of Quantum Wells Formed in AlxGa1-xAsySb1-y/InzGa1-zAs/AlxGa1-AsySb1-y Heterostructures	RL/ER	4 - 1
DR Ercument Arvas	Syracuse University, Syracuse, NY An Assessment of the Current State of the Art of Stap from an Electromagnetics Point of View	RL/OCSS	4 - 2
DR Ahmed E Barbour	Georgia Southern University, Statesboro, GA Formal Verification Using ORA Larch/VHDL Theorem Prover	RL/ERDD	4 - 3
DR Milica Barjaktarovic	Wilkes University, Wilkes Barre, PA Formal Specification and Verification of Missi Architecture Using Spin	RL/C3AB	4 - 4
DR Daniel C Bukofzer	Cal State Univ, Fresno, Fresno, CA Performance Analysis & Simulation Results of Delay & Spread Spectrum Modulated Flip Wave-Signal Gene	RL/C3BA	4 - 5
DR Xuesheng Chen	Wheaton College, Norton, MA Optical and Non-Destructive Methods to Determine the Composition and Thickness of an INxGA1-xAS/INP	RL/ERX	4 - 6
DR Jun Chen	Rochester Inst of Technology, Rochester, NY A Study of Optoelectronic Feedback-Sustained Pulsation of Laser Diodes at 1300 nm & 780 nm	RL/OCPA	4 - 7
DR Everett E Crisman	Brown University, Providence, RI Evaluation of Semiconductor Configurations as Sources for Optically Induced Microwave Pulses	RL/ERAC	4 - 8
DR Digendra K Das	SUNYIT, Utica, NY Techniques for Determining of the Precision of Reliability Predictions and Assessments.	RL/ERSR	4 - 9
DR Matthew E Edwards	Spelman College, Atlanta, Ga The Analysis of PROFILER for Modeling the Diffusion of Aluminum-Copper on a Silicon Substrate	RL/ERDR	4 - 10
DR Kaliappan Gopalan	Purdue University - Calumet, Hammond, IN Speaker Identification & Analysis of Stressed Speech	RL/IRAA	4 - 11
DR Joseph W Haus	Rensselaer Polytechnic Institute, Troy, NY Mode-Locked Laser Models and Simulations	RL/OCPA	4 - 12

SRP Final Report Table of Contents

Author	University/Institution Report Title	Rome Laboratory Directorate	Vol-Page
DR James P LeBlanc	New Mexico State University, Las Cruces, NM Multichannel Autoregressive Modeling & Spectral Estimation Methods for Airborne Radar Environment	RL/OCSS	4 - 13
DR David J McLaughlin	Northeastern University, Boston, MA A Review of Microwave Terrain Clutter Measurements at Bistatic	RL/ERCS	4 - 14
DR Hrushikesh N Mhaskar	Cal State Univ, Los Angeles, Los Angeles, Ca Neural Beam Steering & Direction Finding	RL/ERAA	4 - 15
DR Ronald W Noel	Rensselaer Polytechnic Institute, Troy, NY A Low Dimensional Categorization Technique for C Source Code	RL/C3CA	4 - 16
DR Jeffrey B Norman	Vassar College, Poughkeepsie, NY Frequency Response of Semiconductor Photorefractive Matls: ZnTe:Mn:V,GaAs:Cr,&CdMnTe:V	RL/OCPA	4 - 17
DR Glenn E Prescott	University of Kansas Center for Research, Lawrence, KS Rapid Prototyping of Software Radio Sys Using Field Programmable Gate Arrays & DSP Microprocessors	RL/C3BB	4 - 18
DR Mark R Purtill	Texas A&M Univ-Kingsville, Kingsville, TX A Network Flow Heuristic for Graph Mapping	RL/C3CB	4 - 19
DR Mysore R Rao	Rochester Inst. Of Technology, Rochester, NY Dectection of Concealed Objects in Images: Investigation into Wavelet Transform Based Object Isolation Techniques	RL/OCSM	4 - 20
DR Scott E Spetka	SUNY of Tech Utica, Utica, NY Integrating a Multimedia Database & WWW Indexing Tools	RL/IRD	4 - 21
DR Gang Sun	University of Massachusetts-Boston, Boston, MA Confined Optical Phonon Modes in Si/ZnS Superlattices	RL/EROC	4 - 22

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Mohammad S Alam	Purdue University, Fort Wayne, IN Fast Infrared Image Registration and High Resolution Reconstruction for Real Time Applications	WL/AAJT	5 - 1
DR Dominick Andrisani II	Purdue University, West Lafayette, IN A Fast Fourier Transform Analysis of Pilot Induced Oscillations	WL/FIGC	5 - 2
DR Pnina Ari-Gur	Western Michigan University, Kalamazoo, MI Texture and Microstructure of Hot Rolled Ti-6A1-4V	WL/MLLN	5 - 3
DR James D Baldwin	University of Oklahoma, Norman, OK Statistical Analysis of Fatigue Crack Growth Rate Data for 7075-T6 Aluminum Damaged by Prior Corrosion	WL/FIB	5 - 4
DR Armando R Barreto	Florida International Univ, Miami, FL Deconvolution of The Space-Time Radar Spectrum	WL/AAMR	5 - 5
MR Larry A Beardsley	Univ of Texas at Austin, Austin, TX The Use of Wavelets and Neural Networks in Data Compression, Data Fusion and Their Effects on Target Identification	WL/MNGA	5 - 6
DR Raj K Bhatnagar	University of Cincinnati, Cincinnati, OH Variable Width Template Construction for ATR with HRR Data	WL/AACR	5 - 7
DR Alley C Butler	University of Cincinnati, Cincinnati, OH Importance of Current Crowding and Self-Heating in a CdS/LaS Cold Cathode	WL/MLIM	5 - 9
DR Reaz A Chaudhuri	University of Utah, Salt Lake City, UT A Novel Compatibility/Equilibrium Based Iterative Post-Processing Approach for Axismmetric Brittle	WL/MLBM	5 - 11
DR Julian Cheung	New York Inst. Of Technology, New York, NY New Techniques for Non-Cooperative Target Identification	WL/AACT	5 - 12
DR Milton Cone	Embry-Riddle Aeronautical University, Prescott, AZ Of Match Maker and Metrics	WL/AACF	5 - 13
DR Robert R Criss	Randolph-Macon Woman's College, Lynchburg, VA Optical Studies of Two Novel Electro-Explosive Devices	WL/MNMF	5 - 14

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Robert J DeAngelis	Univ of Nebraska - Lincoln, Lincoln, NE Granin Size Effects in the Determination of X-Ray Pole figures and Orientation Distribution Function	WL/MNM	5 - 15
DR Yujie J Ding	Bowling Green State University, Bowling Green, OH Investigation of Photoluminescence Intensity Saturation and Decay, and Nonlinear Optical Devices in Semiconductor Structures	WL/AADP	5 - 16
DR Gregory S Elliott	Rutgers State Univ of New Jersey, Piscataway, NJ Laser Based Diagnostic Techniques for Combustion and Compressible Flows	WL/POPT	5 - 17
DR Altan M Ferendeci	University of Cincinnati, Cincinnati, OH Vertical 3-D Interconnects for Multichip Modules	WL/AADI	5 - 18
DR Dennis R Flentge	Cedarville College, Cedarville, OH Kinetic Studies of the Thermal Decomposition of Demnum and X-1P Using the System for Thermal Diagnostic Studies (STDS)	WL/POSL	5 - 19
DR Himansu M Gajiwala	Tuskegee University, Tuskegee, AL Novel Approach for the Compressive Strength Improvement of Rigid Rod Polymers	WL/MLBP	5 - 20
DR Allen G Greenwood	Mississippi State University, Mississippi State, MS A Framework for Manufacturing-Oriented, Design-Directed Cost Estimation	WL/MTI	5 - 21
DR Rita A Gregory	Georgia Inst of Technology, Atlanta, GA Affects of Int'l Quality Standards on Bare Base Waste Disposal Alternatives	WL/FIVC	5 - 22
DR Michael A Grinfeld	Rutgers University, Piscataway, Piscataway, NJ Mismatch Stresses, Lamellar Microstructure & Mech	WL/MLLM	5 - 23
DR Awatef A Hamed	University of Cincinnati, Cincinnati, OH Inlet Distortion Test Considerations for High Cycle Fatigue in Gas Turbine Engines	WL/FIM	5 - 24
DR Stewart M Harris	SUNY Stony Brook, Stony Brook, NY Compositional Modulation During Epitaxial Growth of Some III-V Heterostructures	WL/MLPO	5 - 25
DR Larry S Helmick	Cedarville College, Cedarville, OH Effect of Humidity on Wear of M-50 Steel with a Krytox Lubricant	WL/MLBT	5 - 26
DR Kenneth L Hensley	University of Oklahoma, Norman, OK Hyperbaric Oxygen Effects on the Postischemic Brain	MED/SGP	5 - 27

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Iqbal Husain	University of Akron, Akron, OH Fault Analysis & Excitation Requirements for Switched Reluctance Starter-Generators	WL/POOC	5 - 28
DR David W Johnson	University of Dayton, Dayton, OH In Situ Formation of Standards for the Determination of Wear Metals in Perfluoropolyalkylether Lubricating Oils	WL/MLBT	5 - 29
DR Marian K Kazimierczuk	Wright State University, Dayton, OH Aircraft Super Capacitor Back-Up System	WL/POOC	5 - 30
DR Edward T Knobbe	Oklahoma State University, Stillwater, OK Corrosion Resistant Sol-Gel Coatings for Aircraft Aluminum Alloys	WL/MLBT	5 - 31
DR Michael C Larson	Tulane University, New Orleans, LA Cracks at Interfaces in Brittle Matrix Composites	WL/MLLM	5 - 32
DR Douglas A Lawrence	Ohio University, Athens, OH Analysis & Design of Gain Scheduled Missile Autopilots	WL/MNAG	5 - 33
DR Junghsen Lieh	Wright State University, Dayton, OH Determination of 3D Deformations, Forces and Moments of Aircraft Tires with a Synchronized Optical and Analog System	WL/FIVM	5 - 34
DR Chun-Shin Lin	Univ of Missouri - Columbia, Columbia, MO Neural Network Technology for Pilot-Vehicle Interface & Decision Aids	WL/FIGP	5 - 35
DR Zongli Lin	SUNY Stony Brook, Stony Brook, NY Control of Linear Sys with Saturating Actuators with Applications to Flight Control Systems	WL/FI	5 - 36
DR Kuo-Chi Lin	University of Central Florida, Orlando, FL Study on Dead Reckoning Translation in High Level Architecture	WL/AASE	5 - 37
DR James S Marsh	University of West Florida, Pensacola, FL A Conceptual Model for Holographic Reconstruction & Minimizing Aberrations During Reconstruction	WL/MNSI	5 - 38
DR Paul Marshall	University of North Texas, Denton, TX Computational Studies of the Reactions of CH3I With H and OH	WL/MLBT	5 - 39

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Hui Meng	Kansas State University, Manhattan, KS Investigation of Holographic PIV and Holographic Visualization techniques for Fluid Flows and Flames	WL/POSC	5 - 40
DR Douglas J Miller	Cedarville College, Cedarville, OH Band Gap Calculations on Oligomers with an All-Carbon Backbone	WL/MLBP	5 - 41
DR Ravi K Nadella	Wilberforce University, Wilberforce, OH Hydrogen & Helium Ion Implantations for Obtaining High-Resistance Layers in N-Type 4H Silicon Carbide	WL/MLPO	5 - 42
DR Krishna Naishadham	Wright State University, Dayton, OH Hydrogen & Helium Ion Implantations for Obtaining High-Resistance Layers in N-Type 4H Silicon	WL/MLPO	5 - 43
DR Timothy S Newman	Univ of Alabama at Huntsville, Huntsville, All A Summer Faculty Project for Anatomical Feature Extraction for Registration of Multiple Modalities of Brain MR	WL/AACR	5 - 44
DR Mohammed Y Niamat	University of Toledo, Toledo, OH FPGA Implementation of the Xpatch Ray Tracer	WL/AAST	5 - 45
DR James L Noyes	Wittenberg University, Springfield, OH The Development of New Learning Algorithms	WL/AACF	5 - 46
DR Anthony C Okafor	University of Missouri - Rolla, Rolla, MO Assessment of Developments in Machine Tool Technology	WL/MTI	5 - 47
DR Paul D Orkwis	University of Cincinnati, Cincinnati, OH Assessing the Suitability of the CFD++Algorithm for Advanced Propulsion Concept simulations	WL/POPS	5 - 48
Dr Robert P Penno	University of Dayton, Dayton, OH Grating Lobes in Antenna Arrays	WL/AAMP	5 - 49
DR George A Petersson	Wesleyan University, Middletown, CT Absolute Rates for Chemical Reactions	WL/MLBT	5 - 50
DR Mohamed N Rahaman	University of Missouri - Rolla, Rolla, MO Effect of Solid Solution Additives on the Densification & Creep of Granular Ceramics	WL/MLLN	5 - 51

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Martin Schwartz	University of North Texas, Denton, TX AB Initio Modeling of the Enthalpies of Formation of Fluorocarbons	WL/MLBT	5 - 52
DR Thomas E Skinner	Wright State University, Dayton, OH A Method for Studying Changes in Tissue Energetics Resulting from Hyperbaric Oxygen Therapy	MED/SGP	5 - 53
DR Marek Skowronski	Carnegie Melon University, Pittsburgh, PA Investigation of Structural Defects in 4H-SiC Wafers	WL/MLPO	5 - 54
DR Grant D Smith	Univ of Missouri - Columbia, Columbia, MO Theoretical Investigation of Phthalocyanin Dimers	WL/MLPJ	5 - 55
DR James A Snide	University of Dayton, Dayton, OH Aging Aircraft: Preliminary Investigation of Various Materials and Process Issues	WL/MLLP	5 - 56
DR Yong D Song	North Carolina A & T State University, Greensboro, NC Memory-Base Control Methodology with Application to EMRAAT Missile	WL/MNAG	5 - 57
DR Raghavan Srinivasan	Wright State University, Dayton, OH Microstructural Development During Hot Deformation	WL/MLIM	5 - 58
DR Janusz A Starzyk	Ohio University, Athens, OH Feature Selection for ATR Neural Network Approach	WL/AACA	5 - 59
DR Alfred G Striz	University of Oklahoma, Norman, OK On Multiobjective Function Optimization in Engineering Design	WL/FIB	5 - 60
DR Barney E Taylor	Miami Univ - Hamilton, Hamilton, OH Optical and Electro-Optical Studies of Polymers	WL/MLBP	5 - 61
DR Joseph W Tedesco	Auburn University, Auburn, AL Effects of Airblast Characteristics on Structural Response	WL/MNSA	5 - 62
DR Scott K Thomas	Wright State University, Dayton, OH The Effects of Curvature on the Performance of a Spirally-Grooved Copper-Ethanol Heat Pipe	WL/POOS	5 - 63
DR James P Thomas	University of Notre Dame, Notre Dame, IN Subcritical Crack Growth of Ti-6Al-4V Under Ripple Loading Conditions	WL/MLLN	5 - 64
DR Karen A Tomko	Wright State University, Dayton, OH Grid Level Parallelization of an Implicit Solution of the 3D Navier-Stokes Equations	WL/FIM	5 - 65

SRP Final Report Table of Contents

Author	University/Institution Report Title	Arnold Engineering Development Center Directorate	Vol-Page
DR Saad A Ahmed	King Fahd Univ of Petroleum & Minerals, Saudi, Arabia Turbulence Statistics & Energy Budget of a Turbulent Shear Layer	AEDC	6 - 1
DR Csaba A Biegl	Vanderbilt University, Nashville, TN Turbine Engine Blade Vibration Analysis System	AEDC	6 - 2
DR Frank G Collins	Tennessee Univ Space Institute, Tullahoma, TN Laser Vapor Screen Flow Visualization Technique	AEDC	6 - 3
DR Randolph S Peterson	The University of the South, Sewanee, TN	AEDC	6 - 4
DR Robert L Roach	Tennessee Univ Space Institute, Tullahoma, TN A Process for Setting Up Computation of Swirling Flows in the AEDC H-3 Heater	AEDC	6 - 5

Author	University/Institution Report Title	U.S. Air Force Academy Directorate	Vol-Page
DR Ryoichi Kawai	Univ of Alabama at Birmingham, Birmingham, AL A Massively Parallel Ab Initio Molecular Dynamis Simulation of Polymers & Molten Salts	USAFA	6 - 6

SRP Final Report Table of Contents

Author	University/Institution Report Title	Air Logistic Centers Directorate	Vol-Page
DR Sandra A Ashford	University of Detroit Mercy, Detroit, MI Evaluation of Current Jet Engine Performance Parameters Archive, Retrieval and Diagnostic System	OCALC	6 - 7
MR Jeffrey M Bigelow	Oklahoma Christian Univ of Science & Art, Oklahoma City, OK Enhancing Tinker's Raster-to-Vector Capabilities	OCALC	6 - 8
DR K M George	Oklahoma State University, Stillwater, OK A Computer Model for Sustainability Ranking	OCALC	6 - 9
DR Jagath J Kaluarachichi	Utah State University, Logan, UT Optimal Groundwater Management Using Genetic Algorithm	OCALC	6 - 10

INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from $20,000 to $25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The numbers of projected summer research participants in each of the three categories and SREP "grants" are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest, and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1996 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After

completion of the 1990 contract, RDL (in 1993) won the recompetition for the basic year and four 1-year options.

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

YEAR	SRP Participation, by Year			TOTAL
	SFRP	GSRP	HSAP	
1979	70			70
1980	87			87
1981	87			87
1982	91	17		108
1983	101	53		154
1984	152	84		236
1985	154	92		246
1986	158	100	42	300
1987	159	101	73	333
1988	153	107	101	361
1989	168	102	103	373
1990	165	121	132	418
1991	170	142	132	444
1992	185	121	159	464
1993	187	117	136	440
1994	192	117	133	442
1995	190	115	137	442
1996	188	109	138	435

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years. Since then, the number of funded positions has remained fairly constant at a slightly lower level.

3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 52-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous individual requesters (over 1000 annually).

RDL placed advertisements in the following publications: *Black Issues in Higher Education, Winds of Change,* and *IEEE Spectrum.* Because no participants list either *Physics Today* or *Chemical & Engineering News* as being their source of learning about the program for the past several years, advertisements in these magazines were dropped, and the funds were used to cover increases in brochure printing costs.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools. High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1996 are shown in this table.

1996 Applicants and Participants			
PARTICIPANT CATEGORY	TOTAL APPLICANTS	SELECTEES	DECLINING SELECTEES
SFRP	572	188	39
(HBCU/MI)	(119)	(27)	(5)
GSRP	235	109	7
(HBCU/MI)	(18)	(7)	(1)
HSAP	474	138	8
TOTAL	1281	435	54

4. SITE VISITS

During June and July of 1996, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MIs)

Before 1993, an RDL program representative visited from seven to ten different HBCU/Mis annually to promote interest in the SRP among the faculty and graduate students. These efforts were marginally effective, yielding a doubling of HBCI/MI applicants. In an effort to achieve AFOSR's goal of 10% of all applicants and selectees being HBCU/MI qualified, the RDL team decided to try other avenues of approach to increase the number of qualified applicants. Through the combined efforts of the AFOSR Program Office at Bolling AFB and RDL, two very active minority groups were found, HACU (Hispanic American Colleges and Universities) and AISES (American Indian Science and Engineering Society). RDL is in communication with representatives of each of these organizations on a monthly basis to keep up with the their activities and special events. Both organizations have widely-distributed magazines/quarterlies in which RDL placed ads.

Since 1994 the number of both SFRP and GSRP HBCU/MI applicants and participants has increased ten-fold, from about two dozen SFRP applicants and a half dozen selectees to over 100 applicants and two dozen selectees, and a half-dozen GSRP applicants and two or three selectees to 18 applicants and 7 or 8 selectees. Since 1993, the SFRP had a two-fold applicant

increase and a two-fold selectee increase. Since 1993, the GSRP had a three-fold applicant increase and a three to four-fold increase in selectees.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates. This year, 5 HBCU/MI SFRPs declined after they were selected (and there was no one qualified to replace them with). The following table records HBCU/MI participation in this program.

SRP HBCU/MI Participation, By Year				
YEAR	**SFRP**		**GSRP**	
	Applicants	**Participants**	**Applicants**	**Participants**
1985	76	23	15	11
1986	70	18	20	10
1987	82	32	32	10
1988	53	17	23	14
1989	39	15	13	4
1990	43	14	17	3
1991	42	13	8	5
1992	70	13	9	5
1993	60	13	6	2
1994	90	16	11	6
1995	90	21	20	8
1996	119	27	18	7

6. SRP FUNDING SOURCES

Funding sources for the 1996 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1996 SRP selected participants are shown here.

1996 SRP FUNDING CATEGORY	SFRP	GSRP	HSAP
AFOSR Basic Allocation Funds	141	85	123
USAF Laboratory Funds	37	19	15
HBCU/MI By AFOSR (Using Procured Addn'l Funds)	10	5	0
TOTAL	188	109	138

SFRP - 150 were selected, but nine canceled too late to be replaced.
GSRP - 90 were selected, but five canceled too late to be replaced (10 allocations for the ALCs were withheld by AFOSR.)
HSAP - 125 were selected, but two canceled too late to be replaced.

7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

1996 SRP Associate Compensation

PARTICIPANT CATEGORY	1991	1992	1993	1994	1995	1996
Faculty Members	$690	$718	$740	$740	$740	$770
Graduate Student (Master's Degree)	$425	$442	$455	$455	$455	$470
Graduate Student (Bachelor's Degree)	$365	$380	$391	$391	$391	$400
High School Student (First Year)	$200	$200	$200	$200	$200	$200
High School Student (Subsequent Years)	$240	$240	$240	$240	$240	$240

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of $50/day for faculty and $40/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and

GSRP associates, 65 % (194 out of 297) claimed travel reimbursements at an average round-trip cost of $780.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-five percent (85 out of 188) of faculty associates took orientation trips at an average cost of $444. By contrast, in 1993, 58 % of SFRP associates took orientation visits at an average cost of $685; that was the highest percentage of associates opting to take an orientation trip since RDL has administered the SRP, and the highest average cost of an orientation trip. These 1993 numbers are included to show the fluctuation which can occur in these numbers for planning purposes.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

In 1996, RDL implemented direct deposit as a payment option for SFRP and GSRP associates. There were some growing pains. Of the 128 associates who opted for direct deposit, 17 did not check to ensure that their financial institutions could support direct deposit (and they couldn't), and eight associates never did provide RDL with their banks' ABA number (direct deposit bank routing number), so only 103 associates actually participated in the direct deposit program. The remaining associates received their stipend and expense payments via checks sent in the US mail.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

8. CONTENTS OF THE 1996 REPORT

The complete set of reports for the 1996 SRP includes this program management report (Volume 1) augmented by fifteen volumes of final research reports by the 1996 associates, as indicated below:

1996 SRP Final Report Volume Assignments

LABORATORY	SFRP	GSRP	HSAP
Armstrong	2	7	12
Phillips	3	8	13
Rome	4	9	14
Wright	5A, 5B	10	15
AEDC, ALCs, WHMC	6	11	16

APPENDIX A – PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP associates represented 169 different colleges, universities, and institutions, GSRP associates represented 95 different colleges, universities, and institutions.

B. States Represented

SFRP -Applicants came from 47 states plus Washington D.C. and Puerto Rico. Selectees represent 44 states plus Puerto Rico.

GSRP - Applicants came from 44 states and Puerto Rico. Selectees represent 32 states.

HSAP - Applicants came from thirteen states. Selectees represent nine states.

Total Number of Participants	
SFRP	188
GSRP	109
HSAP	138
TOTAL	435

Degrees Represented			
	SFRP	GSRP	TOTAL
Doctoral	184	1	185
Master's	4	48	52
Bachelor's	0	60	60
TOTAL	188	109	297

SFRP Academic Titles	
Assistant Professor	79
Associate Professor	59
Professor	42
Instructor	3
Chairman	0
Visiting Professor	1
Visiting Assoc. Prof.	0
Research Associate	4
TOTAL	188

Source of Learning About the SRP		
Category	Applicants	Selectees
Applied/participated in prior years	28%	34%
Colleague familiar with SRP	19%	16%
Brochure mailed to institution	23%	17%
Contact with Air Force laboratory	17%	23%
IEEE Spectrum	2%	1%
BIIHE	1%	1%
Other source	10%	8%
TOTAL	100%	100%

APPENDIX B – SRP EVALUATION RESPONSES

1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

Evaluation Group	Responses
SFRP & GSRPs	275
HSAPs	113
USAF Laboratory Focal Points	84
USAF Laboratory HSAP Mentors	6

All groups indicate unanimous enthusiasm for the SRP experience.

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below:

A. Better preparation on the labs' part prior to associates' arrival (i.e., office space, computer assets, clearly defined scope of work).

B. Faculty Associates suggest higher stipends for SFRP associates.

C. Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same argument was used to raise the faculty and graduate student participation time a few years ago.)

2. 1996 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 84 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

		How Many Associates Would You Prefer To Get ? (% Response)											
		SFRP				GSRP (w/Univ Professor)				GSRP (w/o Univ Professor)			
Lab	Evals Recv'd	0	1	2	3+	0	1	2	3+	0	1	2	3+
AEDC	0	-	-	-	-	-	-	-	-	-	-	-	-
WHMC	0	-	-	-	-	-	-	-	-	-	-	-	-
AL	7	28	28	28	14	54	14	28	0	86	0	14	0
FJSRL	1	0	100	0	0	100	0	0	0	0	100	0	0
PL	25	40	40	16	4	88	12	0	0	84	12	4	0
RL	5	60	40	0	0	80	10	0	0	100	0	0	0
WL	46	30	43	20	6	78	17	4	0	93	4	2	0
Total	**84**	**32%**	**50%**	**13%**	**5%**	**80%**	**11%**	**6%**	**0%**	**73%**	**23%**	**4%**	**0%**

LFP Evaluation Summary. The summarized responses, by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
   a. Time available for evaluation of applications:
   b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5. a. Benefits of associate's work to laboratory:
   b. Benefits of associate's work to Air Force:
6. a. Enhancement of research qualifications for LFP and staff:
   b. Enhancement of research qualifications for SFRP associate:
   c. Enhancement of research qualifications for GSRP associate:
7. a. Enhancement of knowledge for LFP and staff:
   b. Enhancement of knowledge for SFRP associate:
   c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10. a. Your working relationship with SFRP:
    b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

(Continued on next page)

12. Quality of program literature for associate:
13. a. Quality of RDL's communications with you:
    b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Table B-3. Laboratory Focal Point Reponses to above questions

	AEDC	*AL*	*FJSRL*	*PL*	*RL*	*WHMC*	*WL*
# Evals Recv'd	0	7	1	14	5	0	46
Question #							
2	-	86 %	0 %	88 %	80 %	-	85 %
2a	-	4.3	n/a	3.8	4.0	-	3.6
2b	-	4.0	n/a	3.9	4.5	-	4.1
3	-	4.5	n/a	4.3	4.3	-	3.7
4	-	4.1	4.0	4.1	4.2	-	3.9
5a	-	4.3	5.0	4.3	4.6	-	4.4
5b	-	4.5	n/a	4.2	4.6	-	4.3
6a	-	4.5	5.0	4.0	4.4	-	4.3
6b	-	4.3	n/a	4.1	5.0	-	4.4
6c	-	3.7	5.0	3.5	5.0	-	4.3
7a	-	4.7	5.0	4.0	4.4	-	4.3
7b	-	4.3	n/a	4.2	5.0	-	4.4
7c	-	4.0	5.0	3.9	5.0	-	4.3
8	-	4.6	4.0	4.5	4.6	-	4.3
9	-	4.9	5.0	4.4	4.8	-	4.2
10a	-	5.0	n/a	4.6	4.6	-	4.6
10b	-	4.7	5.0	3.9	5.0	-	4.4
11	-	4.6	5.0	4.4	4.8	-	4.4
12	-	4.0	4.0	4.0	4.2	-	3.8
13a	-	3.2	4.0	3.5	3.8	-	3.4
13b	-	3.4	4.0	3.6	4.5	-	3.6
14	-	4.4	5.0	4.4	4.8	-	4.4

3. 1996 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 257 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from 1 (below average) to 5 (above average) - by Air Force base results and over-all results of the 1996 evaluations are listed after the questions.

1. The match between the laboratories research and your field:
2. Your working relationship with your LFP:
3. Enhancement of your academic qualifications:
4. Enhancement of your research qualifications:
5. Lab readiness for you: LFP, task, plan:
6. Lab readiness for you: equipment, supplies, facilities:
7. Lab resources:
8. Lab research and administrative support:
9. Adequacy of brochure and associate handbook:
10. RDL communications with you:
11. Overall payment procedures:
12. Overall assessment of the SRP:
13. a. Would you apply again?
    b. Will you continue this or related research?
14. Was length of your tour satisfactory?
15. Percentage of associates who experienced difficulties in finding housing:
16. Where did you stay during your SRP tour?
    a. At Home:
    b. With Friend:
    c. On Local Economy:
    d. Base Quarters:
17. Value of orientation visit:
    a. Essential:
    b. Convenient:
    c. Not Worth Cost:
    d. Not Used:

SFRP and GSRP associate's responses are listed in tabular format on the following page.

Table B-4. 1996 SFRP & GSRP Associate Responses to SRP Evaluation

	Arnold	Brooks	Edwards	Eglin	Griffis	Hanscom	Kelly	Kirtland	Lackland	Robins	Tyndall	WPAFB	average
# res	6	48	6	14	31	19	3	32	1	2	10	85	257
1	4.8	4.4	4.6	4.7	4.4	4.9	4.6	4.6	5.0	5.0	4.0	4.7	4.6
2	5.0	4.6	4.1	4.9	4.7	4.7	5.0	4.7	5.0	5.0	4.6	4.8	4.7
3	4.5	4.4	4.0	4.6	4.3	4.2	4.3	4.4	5.0	5.0	4.5	4.3	4.4
4	4.3	4.5	3.8	4.6	4.4	4.4	4.3	4.6	5.0	4.0	4.4	4.5	4.5
5	4.5	4.3	3.3	4.8	4.4	4.5	4.3	4.2	5.0	5.0	3.9	4.4	4.4
6	4.3	4.3	3.7	4.7	4.4	4.5	4.0	3.8	5.0	5.0	3.8	4.2	4.2
7	4.5	4.4	4.2	4.8	4.5	4.3	4.3	4.1	5.0	5.0	4.3	4.3	4.4
8	4.5	4.6	3.0	4.9	4.4	4.3	4.3	4.5	5.0	5.0	4.7	4.5	4.5
9	4.7	4.5	4.7	4.5	4.3	4.5	4.7	4.3	5.0	5.0	4.1	4.5	4.5
10	4.2	4.4	4.7	4.4	4.1	4.1	4.0	4.2	5.0	4.5	3.6	4.4	4.3
11	3.8	4.1	4.5	4.0	3.9	4.1	4.0	4.0	3.0	4.0	3.7	4.0	4.0
12	5.7	4.7	4.3	4.9	4.5	4.9	4.7	4.6	5.0	4.5	4.6	4.5	4.6
Numbers below are percentages													
13a	83	90	83	93	87	75	100	81	100	100	100	86	87
13b	100	89	83	100	94	98	100	94	100	100	100	94	93
14	83	96	100	90	87	80	100	92	100	100	70	84	88
15	17	6	0	33	20	76	33	25	0	100	20	8	39
16a	-	26	17	9	38	23	33	4	-	-	-	30	
16b	100	33	-	40	-	8	-	-	-	-	36	2	
16c	-	41	83	40	62	69	67	96	100	100	64	68	
16d	-	-	-	-	-	-	-	-	-	-	-	0	
17a	-	33	100	17	50	14	67	39	-	50	40	31	35
17b	-	21	-	17	10	14	-	24	-	50	20	16	16
17c	-	-	-	-	10	7	-	-	-	-	-	2	3
17d	100	46	-	66	30	69	33	37	100	-	40	51	46

4. 1996 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

Not enough evaluations received (5 total) from Mentors to do useful summary.

5. 1996 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 113 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from 1 (below average) to 5 (above average)

1. Your influence on selection of topic/type of work.
2. Working relationship with mentor, other lab scientists.
3. Enhancement of your academic qualifications.
4. Technically challenging work.
5. Lab readiness for you: mentor, task, work plan, equipment.
6. Influence on your career.
7. Increased interest in math/science.
8. Lab research & administrative support.
9. Adequacy of RDL's Apprentice Handbook and administrative materials.
10. Responsiveness of RDL communications.
11. Overall payment procedures.
12. Overall assessment of SRP value to you.
13. Would you apply again next year? Yes (92 %)
14. Will you pursue future studies related to this research? Yes (68 %)
15. Was Tour length satisfactory? Yes (82 %)

	Arnold	Brooks	Edwards	Eglin	Griffiss	Hanscom	Kirtland	Tyndall	WPAFB	Totals
# resp	5	19	7	15	13	2	7	5	40	113
1	2.8	3.3	3.4	3.5	3.4	4.0	3.2	3.6	3.6	3.4
2	4.4	4.6	4.5	4.8	4.6	4.0	4.4	4.0	4.6	4.6
3	4.0	4.2	4.1	4.3	4.5	5.0	4.3	4.6	4.4	4.4
4	3.6	3.9	4.0	4.5	4.2	5.0	4.6	3.8	4.3	4.2
5	4.4	4.1	3.7	4.5	4.1	3.0	3.9	3.6	3.9	4.0
6	3.2	3.6	3.6	4.1	3.8	5.0	3.3	3.8	3.6	3.7
7	2.8	4.1	4.0	3.9	3.9	5.0	3.6	4.0	4.0	3.9
8	3.8	4.1	4.0	4.3	4.0	4.0	4.3	3.8	4.3	4.2
9	4.4	3.6	4.1	4.1	3.5	4.0	3.9	4.0	3.7	3.8
10	4.0	3.8	4.1	3.7	4.1	4.0	3.9	2.4	3.8	3.8
11	4.2	4.2	3.7	3.9	3.8	3.0	3.7	2.6	3.7	3.8
12	4.0	4.5	4.9	4.6	4.6	5.0	4.6	4.2	4.3	4.5
Numbers below are percentages										
13	60%	95%	100%	100%	85%	100%	100%	100%	90%	92%
14	20%	80%	71%	80%	54%	100%	71%	80%	65%	68%
15	100%	70%	71%	100%	100%	50%	86%	60%	80%	82%

INLET DISTORTION TEST CONSIDERATIONS FOR HIGH CYCLE FATIGUE IN GAS TURBINE ENGINES

and

A COMPARATIVE STUDY OF NUMERICAL SCHEMES AND TURBULENCE MODELS IN PREDICTING TRANSVERSE JET INTERACTIONS WITH A SUPERSONIC STREAM

Awatef Hamed
Professor
Department of Aerospace Engineering and Engineering Mechanics
University of Cincinnati
Cincinnati, OH 45221-0070

Final Report for:
Summer Faculty Research Program

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

and

Flight Dynamics Laboratory
Wright Patterson Air Force Base

August 1996

INLET DISTORTION TEST CONSIDERATIONS FOR HIGH CYCLE FATIGUE IN GAS TURBINE ENGINES

Awatef Hamed
Professor
Department of Aerospace Engineering and Engineering Mechanics
University of Cincinnati

Abstract

Requirements for lower life cycle cost for the next generation of USAF aircraft will dictate more compact inlet designs because of the lack of length for inlet flow distortion decay. It is imperative that tests reproduce the actual flow conditions at the engine interface plane. This report discusses inlet distortion test considerations for high cycle fatigue in gas turbine engines, and outlines the limitations of current screen tests to simulate the proper levels of turbulence and swirl in the induction system.

INLET DISTORTION TEST CONSIDERATIONS FOR HIGH CYCLE FATIGUE IN GAS TURBINE ENGINES

Awatef Hamed

Introduction

The problem of Aeroelasticity in turbomachines has been the subject of many investigations over the last 20 years. A major part of the research efforts dealt with flutter, or oscillating cascades in a uniform flow. More recently however, as both experimental and computational techniques became more sophisticated, they have been applied to study rotor-stator interactions and stall (1). Only one study was found on the unsteady aerodynamic rotor blade loading due to steady circumferential inlet flow distortion in the inlet flow total pressure, and flow swirl angles (2). To this authors knowledge, no study has been specifically performed to investigate the effects of inlet distortions on fan blade aeroelasticity and high cycle fatigue.

It is well known that the quality of air delivered by the inlet affects the stability as well as the aerodynamic and aeromechanical performance of turbojet and turbofan engines. Guidelines for assessing the effects of inlet total pressure distortion on engine stability and aerodynamic performance were established in the 70's by the SAE Aerospace Council Division (3) The recommended experimental and analytical methodology was based on the fundamental concept that inlet flow quality can be characterized by descriptors that characterize the intensity and extent of the total pressure distortion pattern(4), and that engine stability can be demonstrated by tests using equivalent levels of steady state distortion. The document did not address the effects of other forms of distortion on the propulsion system stability and performance, nor the effects of any type of distortion on aeroelastic stability.

Based on these guidelines current tests of engine response to distortion are conducted with screens to generate steady state total pressure distortion patterns. However, differences have been reported recently between in-flight engine response compared to screen tests. This, and the concerns over the fan rotor blades' high cycle fatigue, raise questions about the adequacy of current evaluation techniques in representing the types and levels of distortion in future inlet designs. This report discusses some of the distortion characteristics which can affect the fan blades' aeromechanical stability, but which can not be simulated properly by current screen tests, whose limitations are briefly outlined.

Screen Inlet Distortion Tests

According to AIR-1419 guidelines, screen sections are placed upstream of the inlet/engine Aerodynamic Interface Plane (AIP to produce specified levels of distortion for a given vehicle/engine combination. The resulting steady state total pressure pattern is measured using 40/48 probe matrix consisting of eight equiangularly spaced rakes with five/six probes in each. In addition to the distortion pattern, the data is used to obtain nondimensional numerical distortion descriptors that characterize the

intensity and extent of the total pressure in the circumferential and radial directions. The use of these descriptors in forecasting and assessing engine stability vary depending upon the stage of engine development (3).

It is necessary to be able to resolve the higher order harmonic contents of the inlet distortion, if the screen tests are to be useful in evaluating the aeromechanical stability of the fan blades. Currently this is limited by the number of rakes, except if the probes are placed at the leading edges of the inlet guide vanes (IGV). In addition, higher levels of multiple per revolution distortion patterns will be required, since the forcing frequencies for blade vibrations are much higher than those for engine dynamic response.

Many of these issues will be addressed in the planned ADLARF rotor resonance tests (5), in which the unsteady blade surface pressure measurements will be obtained in the first stage fan rotor (6) with 2, 3, and 8 per revolution sinusoidal total pressure distortion patterns. The dynamic blade loading will be obtained using two sets of seven pressure transducers imbedded in two adjacent blades. The inlet total pressure distortion pattern will be measured using 40 total pressure probes arranged to cover equal annulus areas on the eight equiangularly spaced rakes. The measured distortion pattern resolution will be enhanced by rotating the distortion screens through 10 positions within one period of the distortion pattern.

Turbulence

The high unfavorable pressure gradients in short subsonic diffusers or across shock boundary layer interactions can cause flow separation in induction systems. High levels of pressure fluctuations have been measured at the engine face, in the separated flow regions. RMS values as high as 7.35% have been reported in conventional short diffusers with uniform inlet conditions and thin boundary layers (7). Power spectral densities obtained from the high frequency response probes in short diffuser indicates large circumferential variation in the energy distribution with spikes at 160 and 300 HZ (8).

Shock-boundary layer interactions on the other hand can cause the turbulent kinetic energy to increase by an order of magnitude even prior to separation (9). Figure 2 from reference (10) presents the streamwise variation in the maximum kinetic energy and shear stress resulting from shock wave boundary layer interactions in a two dimensional transonic channel. One can see from this figure that when the flow is separated the turbulence kinetic energy reaches a maximum level which is ten to nine times the initial values. The situation is further complicated by the different recovery and decay rates of the turbulence intensities in the inner and outer parts of the boundary layers (11), and the strongly distorted turbulence structure in the core region due to shock oscillations (12). Since screens cannot duplicate these levels of turbulence, the inflight dynamic total pressure distortions in supersonic inlets can be much higher than the simulated steady state distortions in the tests.

Few investigations (13, 14) were conducted to study the propagation and decay of pressure fluctuations in inlets. Oates et al. (13) used a converging-diverging duct with plug centerbody upstream of the constant pipe feeding the engine as a turbulence generator. The turbulence was generated in the high shear regions behind the normal shock-boundary layer interactions. They measured the steady state and fluctuating stagnation pressures, at three stations upstream of the compressor face, to determine the turbulence decay characteristics. Based on their experimental data, they concluded that the turbulence decay is primarily limited by dispersion rather than dissipation. Martin et al. (14) conducted an experimental investigation to study the propagation of unsteady disturbances in a Mach 2.6 and 3.0 axisymmetric mixed compression inlet. They used an oscillating vane upstream of the inlet to generate external disturbance, and sinusoidal variations in the exit area to generate air flow demand disturbances. The fluctuating pressure measurements at the engine face were obtained using high frequency response probes (up to 2000 HZ). The measurements indicated that most of the turbulence energy was at frequencies at least an order of magnitude higher than the sinusoidal disturbance frequencies, which mainly produced quasi-steady operating conditions. The Power Spectral Density (PDS) characteristics of the turbulence at the engine face indicated that the total pressure fluctuations in the outer and inner zones associated with the turbulence generated by the shock-boundary layer interactions on the cowl and center body were 180° out of phase. Finally, the amplitude of the power density had a small positive skew, indicating that the incidents of very high values are more common than in a normal distribution.

Swirl

High levels of flow swirl are produced in supersonic combat aircraft inlets under subsonic high incidence, and supersonic low incidence flight conditions. According to Fig. 2 from Ref. 15, the trends in inlet swirl mean angle variations with flight conditions are similar to the total pressure distortion (15). Lecht and Meyer (16) reported that preswirl distortions tended to cause more intense blade force fluctuations than the strongest total pressure distortion. Inlet swirl was also implicated in the high cycle fatigue problem encountered in the APU of the Airbus A300 (17).

High offset diffusers will be utilized in the highly integrated air induction system designs of the next generation tactical aircrafts. High offset can produce severe secondary flow or swirl since the centrifugal forces produce a transverse pressure distribution that moves the low energy flow towards the convex side and the high energy flow towards the concave side. In general, inlet swirl patterns are the product of the superposition of two types of swirl, twin and bulk where both increase with incidence at subsonic flight (15). Bulk or the mean swirl bulk (circumferential mean value of swirl at each radius) is produced when a region of low energy (total pressure) is located at one position of the duct perimeter. It is very sensitive to flight conditions and inlet installation, since they both affect the circumferential location of flow separation (low energy) region. Twin swirl on the other hand consist of two counter rotating components in the cross section, and causes a reversal in the flow swirl angle from hub to tip. It

is typically produced in all curved pipes due to the action of the centrifugal forces on the high energy core flow towards the concave side, which in turn forces the boundary layer flow around the circumference to move inward. Generally twin swirl is less sensitive to external conditions, including flight Mach number, incidence and engine installation.

Alford (18) presented estimates for the reduction in maximum swirl angle at the compressor face, based on a simple two dimensional streamline curvature analysis in the IGVS. According to these results 46% and 74% reductions in the maximum swirl angle were estimated for 10 and 20 IGVS. These estimates do not apply however to twin swirl, since the maximum swirl angles reverse sign around the circumference and between the hub and tip.

Soeder & Botenlal (19) showed that 180° screen patterns can produce a sinusoidal variation in swirl of about ± 15. Lecht and Weyer (16) also produced similar circumferential variations of ± 25° using guide vanes. However, neither screen nor the guide vane tests reproduced the out of phase hub to tip variation of twin swirl (Fig. 3).

Recommendations

Current guidelines for inlet distortion testing address compression system aerodynamic stability further, but not the aeromechanical stability of fan blades. Distortion screen tests cannot duplicate the turbulence of levels, nor the flow swirl patterns in the highly offset short diffusers of future tactical aircrafts. Even though inlet guide vanes reduce the circumferential variation in the flow swirl at the rotor, its higher order harmonic content at the rotor is not known, and warrants further investigation. It is also recommended that future development of supersonic compact inlets should be accompanied by measurements of their turbulence characteristics, because of the high levels of fluctuations associated with flow separation due to the high diffusion rates, and the lack of length for their decay ahead of the engine face.

Acknowledgments

I would like to acknowledge the following people for sharing their experience and discussing with me various aspects of this complex problem. Dr. Joseph Shang, Mr. Keith Numbers and Mr. Lou Surber of the Flight Dynamic Lab, Dr. Joe Hollkamp of the Structure Lab, Mr. Ted Fecke, Mr. John Lueke, and Dr. Doug Rabe of the Propulsion lab at WPAFB, Dr. Jim Mace of McDonnell Douglas, Dr. George Stefco of NASA LeRC, Dr. Yehia El-Ainy of P&W and Dr. Virgil Smith and Mr. Sam Weholfer of AEDC. I would also like to thank Mr. Dan Carrigan, Mr. Dan Sell, and Ms. Sandy Whitney in WPAFB library for helping to find specific references without which this work would not have been possible.

References

1 Tanida, Y. and Namba, M., "Unsteady Aerodynamics and Aeroelasticity of Turbomachines," The 7th ISUATT, Elsevier 1995.

2. Lecht, M., and Weyer, H.B., "On the Unsteady Aerodynamic Rotor Blade Loading in a Transonic Axial Flow Compressor with Unsteady State Inlet Distortion," IUTAM Symposium on Aeroelasticity in Turbomachines. Paris. 18-23 October. 1976. pp. 1-16.

3. SAE Aerospace Council Division Technical Committee S-16. "Inlet Total-Pressure-Distortion Considerations for Gas-Turbine Engines" SAE Aerospace Information Report, AIR-1419, May 1983.

4. SAE Aerospace Council Division Technical Committee S-16. "Gas Turbine Engine Inlet Flow Distortion Guidelines" SAE Aerospace Information Report, ARP-1420, March, 1978.

5. Rabe, D.C., Personal Communication, Aug. 1996.

6. Manwaring, S.R., Rabe, D.C., Lovence, C.B. & Wadia, A.R., "Inlet Distortion Generated Forced Response of a Low Aspect Ratio Transonic Fan" ASME 96-GT-376, June 1996.

7. "Subsonic Diffusers for Highly Survivable Aircraft," McDonnell Aircraft Company, Final Report, AFWAL-TR-86-3025, May 1986.

8. MacMiller, C.J., "Investigation of Subsonic Duct Distortion," Air Force Flight Dynamics Laboratory, Technical Report AFFDL-TR-69-21, March 1969.

9. Hamed, A. and Shang, J.S. "Survey of Validation Data Base for Shockwave Boundary-Layer Interactions in Supersonic Inlets". AIAA Journal of Propulsion and Power, Vol. 7, No. 4, July-August 1991, pp 617-625.

10. Delery, J. M., "Experimental Investigation of Turbulence Properties in Transonic Shock Boundary Layer Interactions," AIAA Journal, Vol. 21, No. 2, 1983, pp. 180-185.

11. Hayakawa, K., Much, K. C., Smits, A. J., and Bogdonoff, S. M., "The Evolution of Turbulence in Shock in Shock-Wave/Boundary Layer Interactions," Presented at 8th Australian Fluid Mechanics Conf., New South Wales, Nov. 28-Dec. 2, 1983.

12. Sajben, M., Morris, T. J., and Bogar, T. J., "Confined Normal Shock/Turbulent Boundary-Layer Interaction Followed by and Adverse Pressure Gradient," AIAA Paper No. 89-0354, Jan. 1989.

13. Oates, G.C., Sherman, D.A., Motycke, D.L. "Experimental Study of Inlet-Generated Pressure Fluctuations", 96AFAPL-TR-69-103, Proceedings of the Airforce Airframe-Propulsion Compatibility Symp., June 1996.

14. Martin, A. W., Bealieu, W.D., Kostinl, L. C. "Analysis and Correlation of Inlet Unsteady Flow Data" NA-71-1146, North American Rockwell Corporation/Los Angeles Division, Prepared under Contract No. NAS-2-5916

15. Kimzey, W.F., "An Investigation and Calibration of a Device for the Generation of Turbulent Flow at the Inlet of a Turbojet Engine." A.E.D.C., Technical Report 65-195, Oct. 1965.

15. Aulehla, F., "Intake Swirl- A Major Disturbance Parameter in Engine/Intake Compatibility" ICAS-82-4.8.1 Proceedings of the 13th Congress of ICAS/AIAA, Seattle, WA August 22-27, 1982. Also MBB/FE/123/S/PUB/74.

16. Lecht, M. and Weyer, H. B., "Unsteady Rotor Blade Loading in an Axial Compressor with Steady-State Inlet Distortions," AGARD CP 248--30, Cleveland, Oct. 1978.

17. Lotter, K. W. and Jorg, J., "The Effect of Intake Flow Disturbances on APU Compressor Blade High Cycle Fatigue in the Airbus A300," ICAS-82-4.6.2, Seattle, Aug. 1982.

18. Alford, J.S.. "Design Criteria for Turbomachinery Periodic Structures to Improve Tolerance to Inflow Distortion and Resonant Oscillatory Flows", Society of Automotive Engineers Paper No. 690388, National Air Transportation Meeting, New York, NY April 21-24, 1969.

19. Soeder, R. H. and Bobula, G. A., "Effect of Steady-State Pressure Distortion on Flow Characteristics Entering a Turbofan Engine," NASA TM 79134, Apr. 1979.

20. Yates, B. and Eagles, C., "Engine 3/23B Gauze and Swirl Tests", Rolls Royce Ltd., Bristol, 1979.

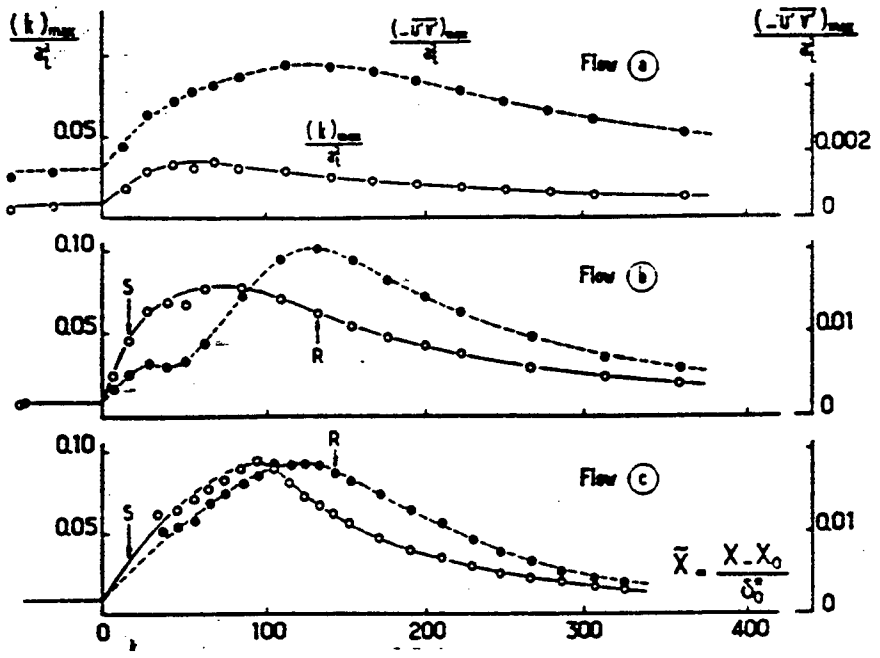


Fig. 1. Maximum kinetic energy and shear stress variation. (10)
(a) Incipient separation (b) Separation (c) High separation

- Engine bellmouth tests - swirl simulation by gauzes and vanes (16), (19), (20).

Twin swirl as measured in intake models

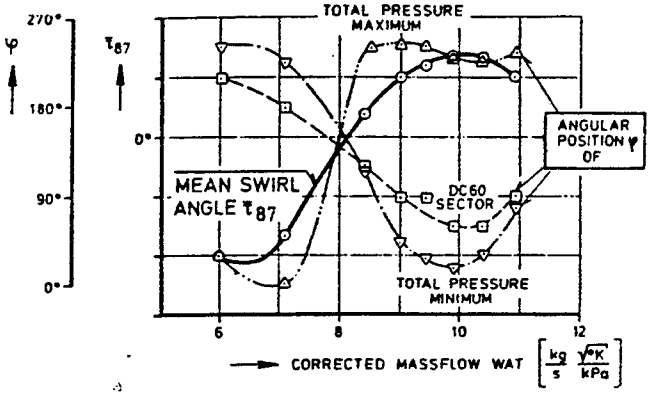


Fig. 2 Rotation of Total Pressure Maximum and Minimum Corresponds to Change of Circumferential Mean Swirl Angle τ_{87}, Basic Intake, $M_\infty = 0{,}5$, $\alpha = 15$. (15)

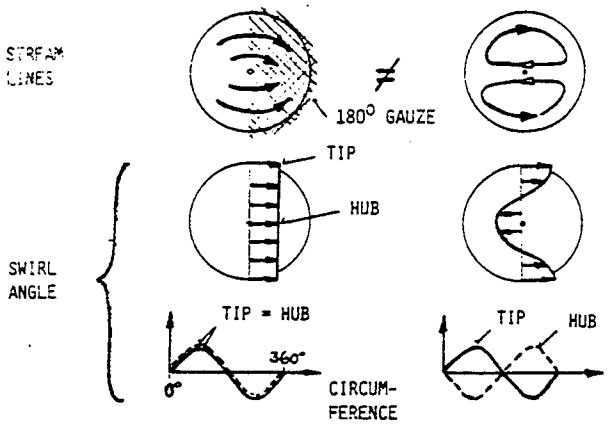


Fig 3. Discrepancy Between Required and Achieved Swirl Simulation.

A COMPARATIVE STUDY OF NUMERICAL SCHEMES AND TURBULENCE MODELS IN PREDICTING TRANSVERSE JET INTERACTIONS WITH A SUPERSONIC STREAM

Awatef Hamed
Professor
Department of Aerospace Engineering and Engineering Mechanics
University of Cincinnati

Abstract

A numerical simulation investigation was conducted to study the flow field near a normally injected secondary flow into a turbulent, supersonic freestream over a flat plate. Different numerical schemes and turbulence models were assessed in terms of their ability to predict the resulting flow field with its complex shear layers, multiple shocks and subsonic recirculation regions. Two implicit Navier-Stokes solvers, namely, COBALT and NPARC, and four turbulence models, were used in the numerical simulation. The computational results are presented and compared with existing experimental data for the separation point and surface pressure distribution. The results indicate that the best agreement with the experimental data in terms of separation point location were obtained with COBALT using the one equation Spalart-Allmaras turbulence model, while the closest agreement with the separation induced pressure rise upstream of the slot was obtained with NPARC with the two equation k-ω model.

A COMPARATIVE STUDY OF NUMERICAL SCHEMES AND TURBULENCE MODELS IN PREDICTING TRANSVERSE JET INTERACTIONS WITH A SUPERSONIC STREAM

Awatef Hamed

Introduction

New advanced, lighter weight, lower cost nozzle designs are required for future advanced fighter aircraft. One of the approaches considered to achieve the USAF nozzle technology development goals by the year 2000 is to use fluidic control, by injecting secondary air symmetrically for throttling or asymmetrically for thrust vectoring [1]. Numerical simulations can be used to investigate the effects of injection slot location, injection angle, and injected mass flow on the nozzle performance in order to reduce the number of experimentally tested configurations [2]. They can also provide more insight into the detailed nozzle flow characteristics under these conditions. The flow field in the vicinity of the jet injection is very challenging to simulate numerically due to the presence of a Mach surface, multiple shocks, shear layers, flow separation and low speed recirculating flow zones. A high degree of confidence in the predictions of the numerical computations is required, especially for the separation location, since it has a strong influence on the nozzle static thrust coefficient [3]. Furthermore, the onset of flow separation and the extent of the shock induced flow in supersonic flows has been found to be affected by turbulence models and numerical algorithms [4].

In this project, two dimensional numerical computations were conducted to study the flow field resulting from normal jet injection in supersonic flow over a flat plate. The simulations were conducted using two different computer codes for the implicit solution of the compressible Navier Stokes equations and four different turbulence models. The computational results are compared with existing experimental data for the surface pressure distribution and separation location. The one equation Spalart-Allmaras turbulence model in the COBALT code gave separation point location predictions in closest agreement with the experimental data, while the two-equation k-ω model in NPARC gave best overall agreement with the experimental pressure distribution.

Configuration and Flow Conditions

The simulated flowfield is shown schematically in Figure 1[5]. Injecting secondary flow normal to the freestream acts as blockage to the primary flow forming a bow shock also termed the jet induced shock. The static pressure increases due to the induced shock which causes the flow to separate resulting in two counter-rotating recirculating regions and a separation shock. Downstream of the slot, a recirculating zone is formed along with a reattachment shock which redirects the flow parallel to the plate surface. In addition, a characteristic Mach surface is formed where the high pressure jet adjusts to the free stream pressure.

Aso et al. [5] reported experimental results for secondary flow injection normal to the turbulent boundary layer over a flat plate in a supersonic stream. The geometry tested consisted of a flat plate with a sharp leading edge and a slot located 0.330 m downstream of the leading edge. The experimental data consisted of Schlieren photographs and surface pressure distributions obtained at free stream Mach numbers ranging from 3.70-3.81 and unit Reynold's number of 5.83×10^6 /m. Measurements were obtained for slot width's of 0.5 mm, 1.0 mm, and 2.0 mm, and injected flow total pressure ratios relative to free stream of 0.083, 0.180, 0.310, 0.440, and 0.510. Presented herein is a comparison of the numerical results with the experimental data for the case of 3.71 free stream Mach number, 1.0 mm slot width, and 0.310 pressure ratio.

Code Descriptions

Two computer codes, namely NPARC [6] and COBALT [7], were used to obtain the numerical solution to the compressible two dimensional Navier-Stokes equations. NPARC solves the Reynolds averaged Navier-Stokes equations in conservation law form and general curvilinear coordinates using a central difference, second order accurate algorithm based on the approximate factorization scheme of Beam-Warming. The algorithm forms an ADI type of scheme in which each sweep involves the inversion of a set of scalar pentadiagonal matrices. Second and forth order dissipation modeled after Jameson is included to increase the robustness of the code. Version 2.2β of the code, which was used in the present investigation, incorporates six different turbulence models, namely Baldwin-Lomax, Baldwin-Thomas, Baldwin-Barth [8], and RNG one equation algebraic models, and the two equation k-ε and k-ω turbulence models of Chien [9] and Wilcox [10] respectively.

COBALT uses a finite-volume, cell-centered, first-order accurate in space and time, exact Riemann solver of Godunov [11]. The approximate Riemann solution method due to Collela [12] in combination with the iterative method of Gottlieb and Groth [13] is used. Second-order accuracy in space is patterned after van Leer's [14] MUSCL scheme where the flow is assumed to vary linearly within each cell. The linear variations are constructed by a least squares method that is solved by QR factorization. Second-, third- and fourth-order temporal accuracy is achieved via the low-storage methods of Williamson [15]. Second-order accurate viscous terms patterned after McCormick [16] are added to yield a Navier-Stokes solver. The one equation Spalart-Allmaras [17] and Baldwin-Barth turbulence models are available in the code. The code was developed primarily for unstructured grids but is applied to a structured grid in this project.

Computational Details

The computational domain used in assessing the two codes is similar to that used by Rizetta [18]. The upper boundary was placed far enough from the plate surface such that all shocks passed through the exit boundary. The downstream boundary was placed far enough downstream of the reattachment point.

Figure 2 represents the 253x81 grid used in the numerical solutions which was generated using the software package GRIDGEN [19]. Ten equally spaced points were placed ahead of the plate leading edge to accommodate the use of the free-stream conditions at the in-flow boundary. A total of 120 points were distributed from the plate leading edge to the slot. Within the slot, 15 equally space points were placed corresponding to 16 segments of $\Delta x/l = 1.89 \times 10^{-4}$. The grid was clustered near the slot in the streamwise direction using algebraic stretching to ensure adequate resolution near the injection location. Normal to the plate surface the grid was stretched such that $\Delta y/l_{min} = 1.0 \times 10^{-5}$ corresponding to a y^+ value of about 1.0 which is necessary for producing good results with the k-ω model [20] and well below the recommended value of 3.0-5.0 for both the Baldwin-Barth [8] and Spalart-Allmaras models [17].

The free stream static pressure, temperature and Mach number were specified at the in-flow boundary. Symmetry condition was invoked at the lower boundary connecting the incoming flow and the leading edge of the flat plate. No-slip, adiabatic wall condition was specified over the plate surface. In the NPARC solution, first order extrapolation of all flow variables was applied at the upper and exit boundaries whereas for COBALT the Riemann invariant boundary condition was invoked at the same boundaries.

One of the fundamental differences between the two codes was in the specification of the boundary conditions across the injection slot opening. NPARC required specification of the stagnation pressure and temperature at each grid point whereas COBALT required specification of the static pressure, temperature and Mach number across each cell surface. The appropriate ratios between stagnation and static pressure and temperatures at choked conditions were used to specify uniform flow conditions in each case.

Turbulence quantities for the simulation were not readily available. Therefore, first order extrapolation was applied since it was the most straightforward approach for maintaining consistency between the two codes and different turbulence models.

Starting from uniform initial conditions corresponding to free stream conditions, the solution was advanced using local time stepping. For NPARC, the CFL number was varied from 0.2 initially to 0.6 at latter stages in the computations. For COBALT, the CFL number was determined within the explicit solution throughout the entire iterative process. A total of 30,000-40,000 iterations were necessary for NPARC while 70,000-80,000 iterations were performed using COBALT. The solution was considered converged when the maximum change in the friction coefficient was < 0.5% after 2000 consecutive iterations, which occurred directly downstream of the separation point. In the case of NPARC, this corresponded to 3 to 4 orders of magnitude reduction in the total L2 residual.

Results and Discussions

Figure 3 presents the Mach number contours from the solutions using COBALT and NPARC omitting the Baldwin-Thomas model due to its poor prediction of separation and surface pressure. The

figure clearly shows the jet induced and separation shock upstream, and the reattachment shock downstream, which is especially well defined in the NPARC k-ω case. The general shape of the plume is in good agreement between the two codes and different turbulence models.

Figures 4 and 5 compare the computed surface pressure distributions and surface friction using the two codes and four turbulence models to the experimental data. Rizzetta's computational results [18] which were obtained using McCormack explicit scheme and k-ε turbulence model are reproduced in the two figures. From Fig. 4, one can see that the predictions using NPARC with the k-ω turbulence model is closest to the experimental data in terms of the onset of the surface pressure rise associated with the flow separation, and pressure gradient upstream of the injection. COBALT with the Spalart-Allmaras model gave the second best predictions with the pressure rising earlier and at a higher initial rate than the experimental results. The Baldwin-Barth of NPARC predicted the onset of the pressure rise much earlier than the experimental data, however, it most accurately predicted, the static pressure level near the slot. COBALT predictions using the Baldwin-Barth model are close to Rizzetta's with the onset of the pressure rise shortly behind the experimental data, and a much higher shock strength than the experimental data and all other predictions except for NPARC with the algebraic Baldwin-Thomas model. All computations underpredicted the pressure recovery behind the slot, but COBALT predictions were the closest to the experimental results further downstream.

The computed skin friction coefficients, are presented in Figure 5, and the predicted separation locations are compared to the experimental results in Table 1. According to this data, the computed separation location using COBALT with the Spalart-Allmaras model is in closest agreement with experimental results (within 0.14%) followed by NPARC with the k-ω model (within 0.53%). The Baldwin-Barth model with NPARC and COBALT resulted in overpredicting and underpredicting the separation location by 1.84% and 1.81% respectively, with Rizzetta's results very close to the latter. Again, the Baldwin-Thomas algebraic model resulted in the greatest error in predicting the separation location, deviating by more than 3%. Further investigation of the effect of grid refinement study may be required to resolve the spread in the computed friction coefficients ahead of the separation region.

Immediately downstream of the slot there is a noticeable difference between the friction coefficient distributions predicted by COBALT and NPARC. Comparing the velocity vectors in the recirculation zone downstream of the slot for the two codes and Baldwin-Barth model as compared in Figures 6 and 7, one can see only one recirculation in the COBALT predictions, but while NPARC predicts a secondary smaller counter-rotating recirculation zone can be seen in NPARC predictions. It was determined in the present investigation that this second vorticity was formed irrespective of the turbulence model, pressure ratio, slot width, artificial viscosity, and CFL number. Yeneriz et al. reported similar findings concerning this type of vortical structure in their PARC code predictions of hypersonic flow over an axisymmetric cone with injection [22]. However, Rizzetta also reported only one recirculation region downstream of the slot as in the COBALT predictions.

The mass and momentum flux were numerically integrated across the slot opening and the results are presented in Table 1. The flow and thrust coefficients, which represent the ratio between the computed and ideal sonic jet, were found to be in agreement within 99.99% and 98.31% respectively. According to the computational results COBALT predicts higher pressures and densities and lower Mach numbers and momentum compared to NPARC. The difference between the predictions of the two codes is associated with how the boundary conditions at the slot were implemented in each code.

	COBALT Baldwin-Barth	COBALT Spalart-Allmaras	NPARC Baldwin-Barth	NPARC Baldwin-Thomas	NPARC k-ω	Experiment/ Ideal
Separation Point (x/l)	0.917	0.899	0.884	0.929	0.905	0.900
Percent Error	*+1.84*	*+0.14*	*-1.81*	*-3.11*	*+0.53*	-
Jet Mass Flux (kg/s/m)	0.8989	0.8990	0.8994	0.8993	0.8994	1.0278
Flow Coefficient	*0.8750*	*0.8750*	*0.8751*	*0.8750*	*0.8751*	-
Jet Momentum Flux (J)	*236.4804*	*236.5067*	*240.6067*	*240.5800*	*240.6067*	*275.4461*
Thrust Coefficient	*0.8585*	*0.8586*	*0.8735*	*0.8734*	*0.8735*	-

Table 1. Comparison of different turbulence models with experimental results.

Conclusions

Two dimensional numerical simulations were conducted to study the flow field resulting from normal jet injection in supersonic flow over a flat plate. The simulations were conducted using two different computer codes for the implicit solution of the compressible Navier-Stokes equations and four different turbulence models. The computational results were compared with existing experimental data for the surface pressure distribution and separation point location. The results indicated that the one equation Spalart-Allmaras model which was implemented in the COBALT code gave the closest agreement with the experimental data in terms of separation point location. On the other hand, the k-ω model which was implemented in the NPARC code gave the best agreement with the experimentally measured surface pressure rise upstream of the slot.

Acknowledgments

I would like to thank Dr. Joseph Shang, Dr. Don Rizzetta and Mr. Keith Numbers of the Flight Dynamics Lab, WPAFB for their help and support. I would also like to thank my graduate student, Greg Laskowski, for his dedication, thoroughness and hard work, without whom this work would not be complete.

References

1. Gridley, M. C., and Walker, S. H., "Propulsion Integration Issues for Advanced Fighter Aircraft," Air-Force Technical Report 142-96, Wright Laboratory, Wright-Patterson AFB, OH.

2. Wing, D. J., "Static Investigation of Two Fluidic Thrust-Vectoring Concepts on a Two-Dimensional Convergent-Divergent Nozzle," NASA Technical Memorandum 4574, Dec. 1994.

3. Hamed, A., Vogiatzis, C., and Yeuan, J. J., "Performance Prediction of Overexpanded 2DCD Nozzles," 12th International Symposium of Air Breathing Engines, September 1995.

4. Hamed, A., "An Investigation of Oblique Shock/Boundary Layer Interaction Control," AFOSR Technical Report Contract No. 91-0101, January 1992.

5. Aso, S., Okuyama, S., "Experimental Study on Mixing Phenomena in Supersonic Flows with Slot Injection," AIAA Paper No. 91-0016, January 1991.

6. Cooper, G. K., and Sirbough, J. R., "PARC Code: Theory and Usage," Arnold Engineering Development Center Report AEDC-TR-89-15, Arnold AFS, December 1989.

7. Strang, W. Z., "COBALT User's Manual," WL/FIMC WPAFB, OH, February 1996.

8. Baldwin B. S., and Barth T. J., "A One-Equation Turbulence Model for High Reynold Number Wall-Bounded Flows," AIAA Paper No. 91-0610, January 1991.

9. Chien, K-Y., "Prediction of Channel and Boundary-Layer Flows with a Low Reynolds Number Turbulence Model," AIAA Journal, Vol. 20, pp. 33-38, January 1982.

10. Wilcox, D. D., "The Remarkable Ability of Turbulence Model Equations to Describe Transition," 5th Symposium on Numerical and Physical aspect of Aerodynamic Flow, Long Beach, CA, Jan. 1992.

11. Gudunov, S. K., "A Finite Difference Method for the Numerical Calculation of Discontinuous Solutions of the Equations of Fluid Dynamics," Mat, Sb., Vol. 47, pp. 271-290, 1959.

12. Collela, P., "Glimm's Method for Gas Dynamics," SIAM Journal of Scientific Statistical Computing, Vol. 3, No. 1, pp. 76-110, 1982.

13. Gottlieb, J. J., and Groth, C. P. T., "Assessment of Riemann Solvers for Unsteady One-Dimensional Inviscid Flows of Perfect Gasses," Journal of Computational Physics, Vol 78, pp. 437-458, 1988.

14. van Leer, B., "Towards the Ultimate Conservative Difference Scheme. V. A Second-Order Sequel to Godunov's Method," Journal of Computational Physics, Vol 32, pp 101-136, 1979.

15. Williamson, J. H., "Low-Storage Runge-Kutta Schemes," Journal of Comp. Physics, Vol 35, pp 48-56.

16. MacCormack, R. W., "The Effect of Viscosity in Hypervelocity Impact Cratering," AIAA Paper No. 69-0354, 1969.

17. Spalart, P. R., and Allmaras S. R., "A One-Equation Turbulence Model for Aerodynamic Flows," AIAA Paper No. 92-0439, January 1992.

18. Rizzetta, D. P., "Numerical Simulation of Slot Injection into a Turbulent Supersonic Stream," AIAA Journal, Vol 30, No. 10, October 1992.

19. Steinbrenner, J. P., Chawner, J. R., and Fouts, C. L., "The GRIDGEN 3D Multiple Block Grid Generation System," WRDC-TR-90-3022 Vols. I and II, Wright-Patterson AFB, OH, 1990.

20. Hamed, A., and Vogiatzis, C., "Assessment of Turbulence Models in Overexpanded 2D-CD Nozzle Flow Simulations," AIAA Paper No. 95-2615, 1995.

21. Yeneriz, M. A., Davis, J. C., Cooper, G. K., and Harvey, D. W., "Comparison of Calculation and Experiment for a Lateral Jet From a Hypersonic Biconic Vehicle," AIAA Paper No. 89-2548, July 1989.

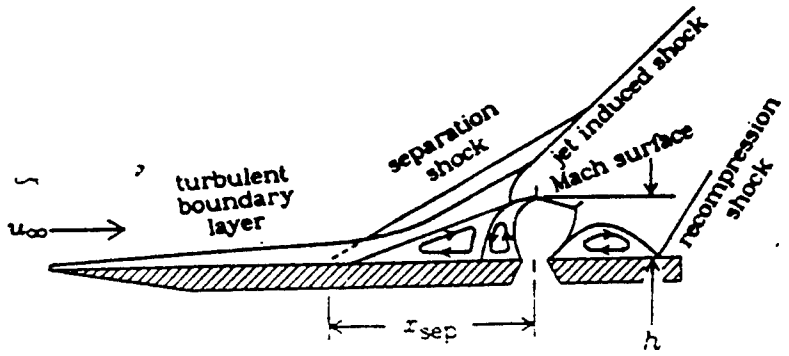


Figure 1. Schematic of flowfield interaction characteristics.

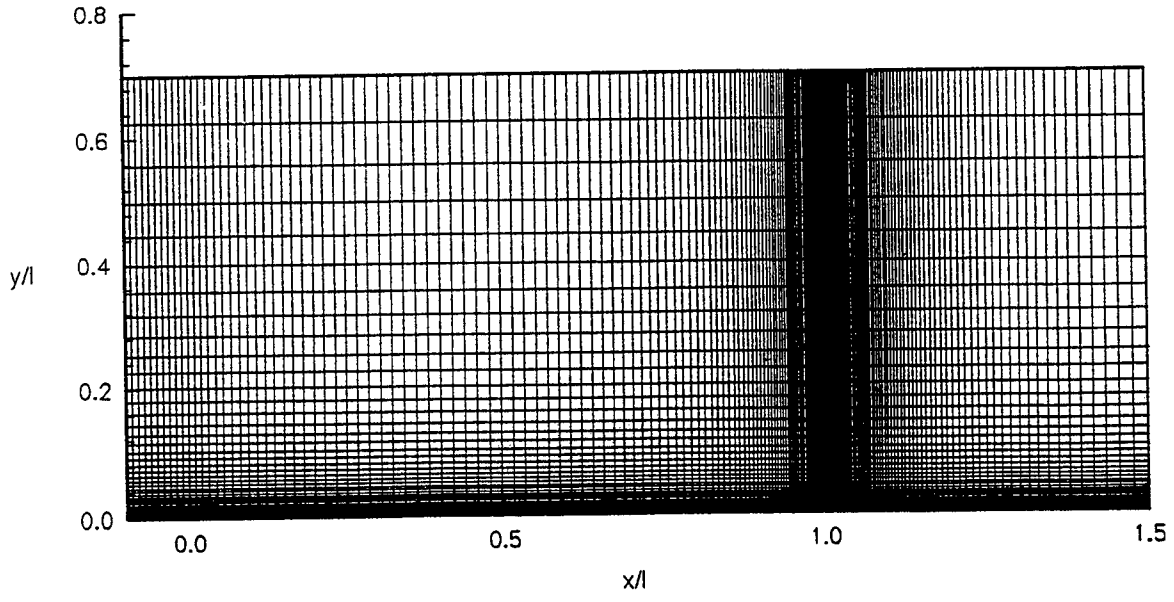


Figure 2. Computation grid for entire flow domain.

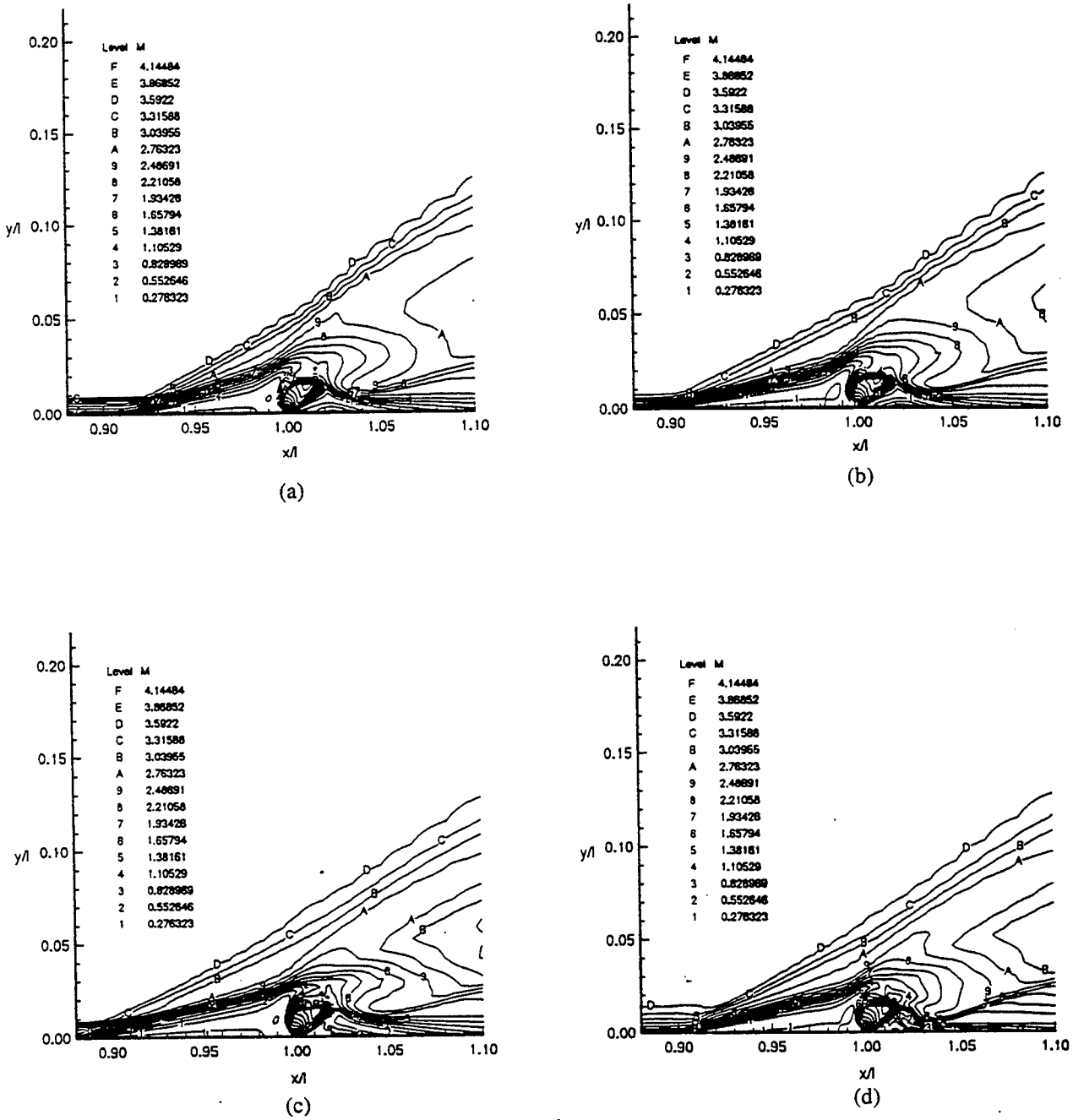


Figure 3. Mach number contours near slot opening. (a) COBALT, Baldwin-Barth (b) COBALT, Spalart-Allmaras (c) NPARC, Baldwin-Barth (d) NPARC, k-ω.

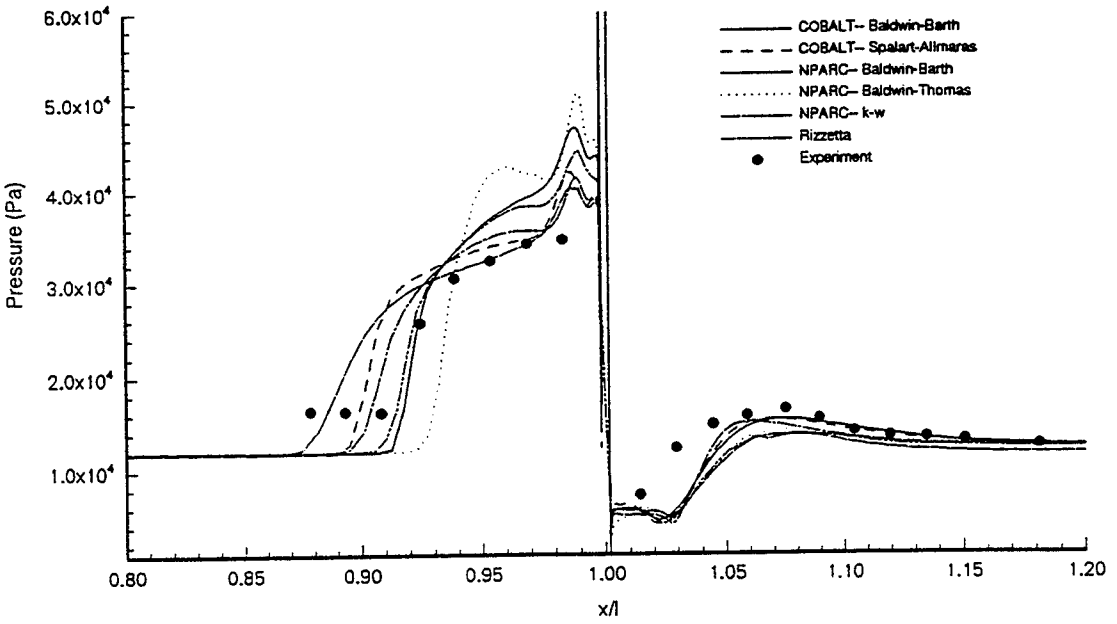


Figure 4. Surface pressure distribution near slot opening.

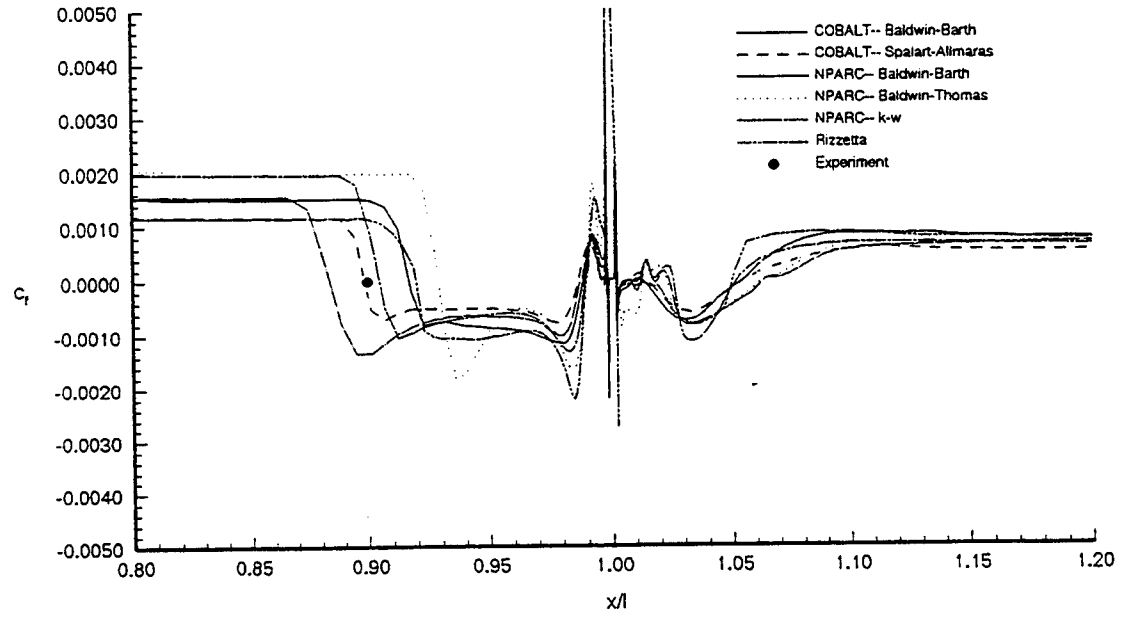


Figure 5. Friction coefficient near slot opening.

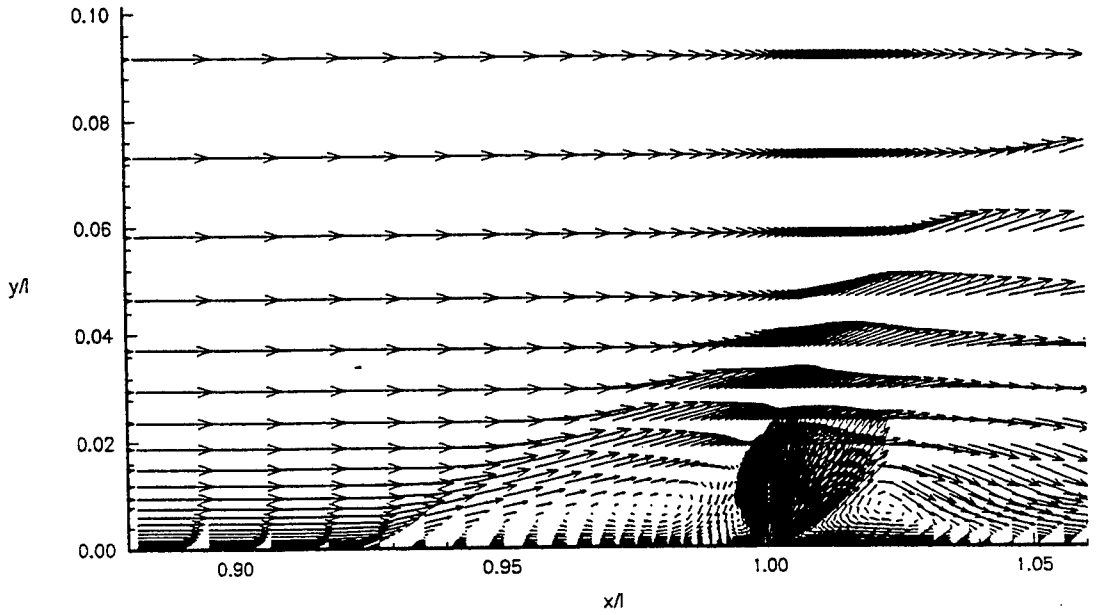


Figure 6. Velocity vectors near slot opening for COBALT using the Baldwin-Barth turbulence model.

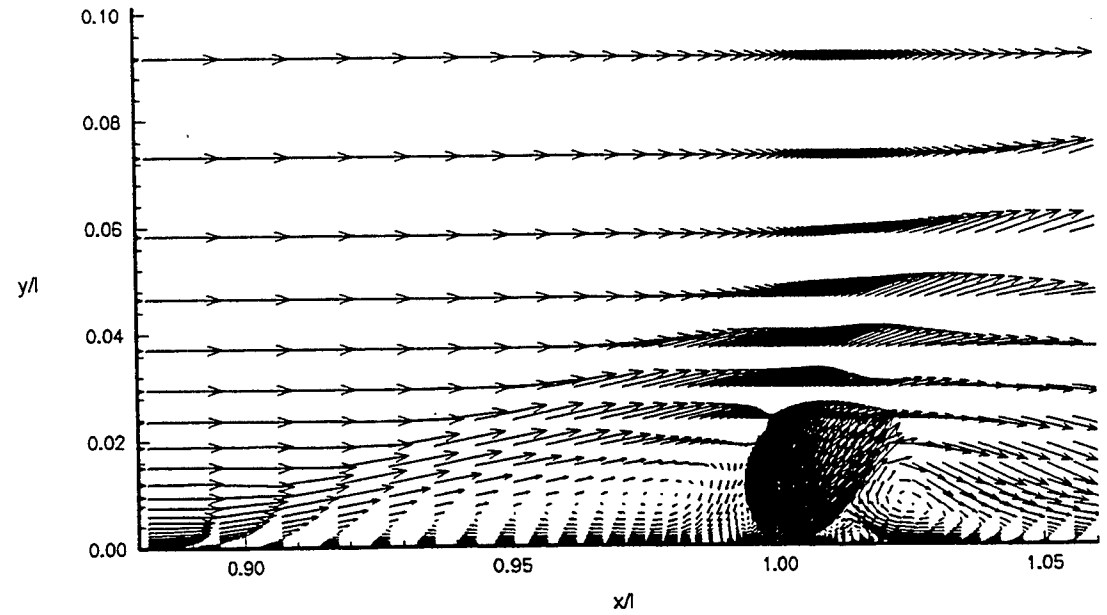


Figure 7. Velocity vectors near slot opening for NPARC using the Baldwin-Barth turbulence model.

COMPOSITIONAL MODULATION DURING EPITAXIAL GROWTH OF SOME III-V HETEROSTRUCTURES

Stewart Harris
Professor
Department of Mechanical Engineering

SUNY at Stony Brook
Stony Brook, NY 11794

Final Report for:
Summer Faculty Research Program
Wright Laboratory

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

and

Wright Laboratory

August 1996

COMPOSITIONAL MODULATION DURING EPITAXIAL GROWTH OF SOME III-V HETEROSTRUCTURES

Stewart Harris

Professor

Department of Mechanical Engineering

SUNY at Stony Brook

Abstract

Analytical models describing the vertical transfer reaction that occurs during the epitaxial growth of certain III-V alloys were formulated and solutions for the mole fractions of the metal components as a function of the layer thickness obtained. These results support recent simulation studies which indicate the occurrence of compositional modulation in the high temperature MBE growth of (Al,Ga)As. Since this has not yet been experimentally verified, the most detailed model that we considered was for the growth of (In,Ga) As where our model provides a more accurate description of the transfer reaction and a better interpretation of the experimental results than the only previous treatment.

COMPOSITIONAL MODULATION DURING EPITAXIAL GROWTH OF SOME III-V HETEROSTRUCTURES

Stewart Harris

I. Introduction

Exchange reactions during growth of III-V alloys provide a mechanism that, if properly understood, could allow the controlled growth of superlattice heterostructures with novel properties that could provide the basis for a variety of new device applications. There is now experimental evidence [1] that under appropriate conditions (temperature, growth rate, flux-ratios) such reactions occur but as of yet there appears to be no established complimentary theoretical kinetic description of these processes. This is not suprising since, even in the case of well-studied ternary alloys such as (Al,Ga)As a kinetic description that includes deposition, desorbtion, and exchange is quite complicated, involving coupled non-linear differential equations. Kinetic monte-carlo (KMC) studies [2] carried out at the host laboratory over the past several years have uncovered an interesting phenomena, which we refer to as compositional modulation (CM), in (Al,Ga)As at high temperatures (approximately 930K) where both desorbtion and exchange kinetics are significant. Over a range of III-V flux ratios in this temperature regime it was found that the Al and Ga mole fractions oscillate between approximately fixed values with a period of one monolayer, i.e. $Al)_x$(layer n) = $Al)_x$(layer n+2) with different values for n even, n odd. This is quite different from the lateral peridicity found previously during growth on vicinal surfaces under conditions where desorbtion does not occur (3). The goal of the research described here is to provide a theoretical confirmation of CM that will both compliment the KMC results and also provide additional insights into the role

of the various system parameters. It is interesting to note that CM is an example of a class of phenomena that is of great topical interest, namely that of self-organization in complex systems [4], which is the focus of extensive activity in a wide variety of systems ranging from biology to tectonic dynamics.

Since experimental confirmation of CM is not expected until Fall 1996 at the earliest, we have focused our efforts on kinetic modeling on the Ga-In transfer mechanism for which there is experimental benchmarking [5] to compare with. A discussion of the (Ga,In)As system along with new results for the transfer coefficient is presented in Section II; this represents completed work that will be submitted for journal publication. Our consideration of the more complicated (Al,Ga)As system began with some very simple cellular automata to verify the occurance of CM prior to undertaking the more difficult task of kinetic modeling. In Section III we discuss a number of such models that exhibit CM and provide us with the motivation for further, more detailed study. We then proceed to develop and discuss a general model of this type in Section IV that allows us to describe the compositional evolution in terms of four parameters that depend on the system kinetics and some preliminary results are obtained for a specific choice of these parameters. It is important to emphasize that this analysis is based on the consideration of the monolayer level as the relevant variable and not the time. This is analogous to examining a grown sample and inferring the kinetics from the resulting structure rather than following the growth evolution directly in time.

In Section V we describe the current state of our study modeling (Al,Ga)As

kinetics at the temperatures where CM is expected to occur. We indicate the problems encountered and how these are being resolved. The major difficulty is a consequence of the three-dimensional growth that examination of the KMC data indicates is present in these simulations. We will very shortly have data for two-dimensional (layer-by-layer) growth which will provide a basis for our first modeling effort. Additional work to be undertaken for (Al,Ga)As is also discussed in Section V. We conclude this report in Section VI with a summary of the results obtained, the work now in progress, and our suggestions for further work together with a restatement of the potential significance of these results relative to applications to Air Force programs.

II. Exchange Kinetics for a InAs/GaAs Structure

When a III-V compound semiconductor consists of two metals having markedly differing bond energies with the group V element there is a possibility of the weaker bonding metal being replaced in the layer below the growing surface by the stronger bonding metal. The occurence and extent of the vertical reaction depends on a number of factors including the group III component flux ratio, the group III-V flux ratio, overall growth rate, and temperature. Because the compositions on both the two uppermost layers influence the reaction the kinetic description is not first order and the resulting kinetic equations are non-linear. This poses a potential problem for any analytical analysis, and it is particularly useful therefore to be able to test analytical results against experimental data. For this reason we initiated our study by considering a system that had been considered previously both experimentally and theoretically and in

which exchange kinetics occurs. As we describe below, the more detailed theoretical description considered here provides a number of insights not apparent in the original theoretical treatment which should prove useful in considering CM in (Al,Ga)As alloys.

A major simplification in the study of exchange kinetics is made possible by employing the deposition technique referred to as migration enhanced epitaxy (MEE) [5] in which the group III components are deposited separately between depositions of the group V component. The system considered is (Ga,In)As in which the Ga replaces the In in the underlayer because of the bonding differences mentioned above. This system was analyzed using secondary ion mass spectroscopy (SIMS) following deposition of 1 ML of InAs and an overlayer of GaAs at 500C; it was found that the SIMS profile was asymmetric which indicates migration of the In into the GaAs overlayer consistent with the replacement of In by Ga. Corresponding SIMs profiles at 400C and 560C were symmetric, in the former case because replacement was not activated at the lower temperature and in the latter case due to the onset of In evaporation. To reconcile these experimental facts with a theoretical model it was assumed that deposition was infinitely rapid; in the actual experiment the schedule following the deposition of the InAs monolayer was to deposit 1ML of Ga in two seconds followed by a three second stage for As deposition and annealing and then repeating this. It was also implicitly assumed that growth was layer-by-layer and that replacement only occurs between the two uppermost layers. With these restrictions the In mole fraction in the last completed layer, $I_2(t)$, and the layer immediately below, $I_1(t)$, are given by two non-linear kinetic equations

$$\frac{dI_2}{dt} \equiv I_2' = kI_1I_2 \tag{1}$$

$$I_2' = -I_1' \tag{2}$$

which have the solution

$$I_2 = I_o[\exp(1-I_o)kt - 1]/[\exp(1-I_o)kt - I_o] \tag{3}$$

$$I_1 = I_o[1-I_o]/[\exp(1-I_o)kt - I_o] \tag{4}$$

where $I_o = I_1(0)$, k is the kinetic coefficient regulating the transfer process, and $0 \quad t \quad T$ with T the time for a single GaAs deposition cycle (five seconds). Note that in general $I_1 + I_2 = I_o$, reflecting conservation of In, and that in applying the solutions to the initial In monolayer and first Ga overlayer that L'Hospital's rule must be used since $I_o = 1$. These equations can be used to determine the In concentration profile after the deposition of the GaAs overlayer by repeated application with the value of I_o calculated for each new GaAs monolayer from the previous cycle.

The material property that appears in eqs(1),(2) is the coefficient k. This can be directly related to the probability P that an In atom will be replaced which has a more direct physical meaning and is related to the parameters we will use to characterize (Al,Ga)As in the following two sections. Specifically, we have

$$P = \lim_{I_o \to 0} I_2(T)/I_o = 1 - \exp(-kT) \tag{5}$$

so that unlike the material property k, P depends on the cycle schedule as well as the material properties (through k).

Composition profiles calculated using eqs(3),(4) can be converted to SIMS profiles by inverting the standard convolution algorithm; this has been done in [5] assuming a background intensity of $2\text{x}10^{19}\text{cm}^{-2}$ and

instrument resolution of 12 nm; in what follows we therefore only need to refer to the corresponding concentration profile, which is shown in figure 1. The values of k and P are 0.78 sec^{-1} and 0.98, respectively. Since the focus of interest in the work reviewed above was to show that the In migration as observed by SIMS analysis could be explained in terms of a simple model of the transfer kinetics, no attempt was made to consider aspects of the problem that are relevant to an understanding of CM in the MBE deposition of (Al,Ga)As. In particular, the solution of the basic kinetic equations becomes more complicated when these include the deposition rates as must be done with MBE. In the remaining part of this section we consider this generalization as well as the conclusions this leads to. As might be expected, we find that regardless of the cycle details and specific (but finite) value of k, the quantity P when properly determined fully charar-acterizes the In migration; however, P will depend on both the value of k and the cycle details.

As noted in [5], eqs (1) and (2) are "easily solved". This is a result of the fact that when the conservation relationship $I_o = I_1 + I_2$ is used in eq(1) that equation is separable and the solution follows directly by quadrature. Despite the non-linearity, the separability puts this equation in the simplest class of first order ODE's. When we relax the restriction that the deposition time is infinitely rapid and include the details of the deposition schedule eqs(1) and (2) are no longer valid and instead we must consider two pairs of equations per cycle, one for $0 < t < 2$ (Ga deposition) and one for $2 < t < 5$ (As deposition and anneal). We then have

$$G_2' = R - G_2 I_1 k \tag{6}$$

$$I_1' = - G_2 I_1 k \tag{7}$$

in the interval $0 < t < 2$ where G_2 is the Ga mole fraction in the overlayer in which deposition is occurring. For $2 < t < 5$ we have the same equations except that R is no longer included since no Ga deposition occurs in this interval. Eq(6) is not separable, but using the mass balance relationship $G_2 = Rt - (I_o - I_1)$ we can write eq(7) as

$$\begin{aligned} I_1' &= -[Rt - (I_o - I_1)]I_1 k \\ &= -kI_1^2 - kI_1(Rt - I_o) \end{aligned} \tag{8}$$

which is solvable despite the non-linearity and non-constant coefficient. The solution follows from the substitution $U = I_1^{-1}$, which transforms eq(8) into a solvable equation for U from which I_1 is then determined as

$$I_1 = I_o / \exp k(\tfrac{1}{2}Rt^2 - I_o t) \times [1 + kI_1(0)\int_o^t dt' \exp -k(\tfrac{1}{2}Rt'^2 - I_o t')] \tag{9}$$

The integral in eq(9) can be written as an error function by a simple change of variable, and I_1 can then be directly calculated.

At the conclusion of Ga deposition we have $I_1 = I_1(t=2) = I_o^*$. Re-setting the clock so that we now consider the interval $0 < t < 3$ with $I_1(0) = I_o^*$, and making use of the mass balance relationship $G_2 + I_2 = 1$, $I_2 + I_1 = I_o$ we find that I_1 now satisfies

$$I_1' = k[I_o - (1+I_1)]I_1 \tag{10}$$

which is separable and can be solved either by quadrature or by again using the substitution $I_1 = 1/U$ to remove the non-linear term. In any event we find

$$I_1 = I_o^*(1-I_o) / [\exp k(1-I_o)t \ (1 - I_o + I_o^*) - I_o^*] \tag{11}$$

where we must again use L'Hospital's rule for the initial In and first Ga layers since $I_o = 1$. From eq(11) together with eq(9) it follows that P now depends on R as well as k and T

$$P = 1 - \exp-k(T-\tfrac{1}{2}R) \qquad (12)$$

and for the cycle which is the basis of the SIMS analysis in [5], with T = 5sec, R = $\frac{1}{2}$sec^{-1} we find that P = 0.978 corresponds to a transfer coefficient of k = 0.98, roughly 25% higher than that estimated previously.

As might be expected, the parameter of significance relative to determining the In distribution in the Ga overlayer is P, and any realizable combination of k (temperature) and cycle schedule T, R that results in a common value of P also produces an identical concentration profile, and simulated SIMS profile. Thus, for the conditions of reference [5], the correct value of k is 0.98. An identical result would be obtained according to our results if, e.g. the temperature is slightly decreased so that k = 0.9, the cycle time is increased to T = 5.1 sec, and the deposition rate increased so that R = 2/3 sec^{-1}. Finally we note that if only R is varied, then for k = 0.98 (500C) and T = 5 sec, increased R to 1 sec^{-1} (so that P = 0.988) results in a longer tail with a concentration at each Ga monolayer reduced by approximately 10% from the values shown in figure 1; conversely, if deposition is continuous, so that R = 1/5 sec^{-1}, then P is reduced to 0.913 and the concentration profile is considerably enhanced but reduced in breadth.

In the following two sections we consider the (Al,Ga)As system using an approach motivated by the above results that emphasize the significance of P.

III. CM in Some Simple Cellular Automata

In the previous section we described the analysis of a kinetic model

that has many of the features required for a study of CM in (Al,Ga)As. This provides a useful first step for considering the more complicated transfer process that occurs when the two group III components are deposited together as in MBE. In parallel with that work we also carried out an analysis based on some simple cellular automata for the purpose of verifying that CM could occur and was not just an artifact of the KMC simulation model. This verification was of particular importance since there is no current experimental evidence for CM [6]. As is often the case, these two parallel lines of inquiry led to some similar conclusions which we were able to more formally generalize as will be discussed in Section IV..

The main feature of the approach considered in this and the following section is that we look at the composition evolution in terms of the changes in successively deposited monolayers, i.e. the monolayer level is the variable of change, not time. All the kinetics are implicit in the set of parameters (model rules) specifying that the probability that a Ga or Al adatom in layer n will have a Ga or Al adatom above it in layer n+1, (As a distinct problem for subsequent study) These probabilities are related to the kinetics which involve a number of processes depending on the deposition rates of the group III and V components. Here, as we discussed above, our primary goal is to establish that CM can occur in a framework independent of the KMC model.

Because the interaction between the substrate and the initial layer is much different from that between successive deposited layers we begin with the first layer. Also, we consider at most six monolayers (KMC results only extend to seven) since the calculations become unwieldly beyond that level. A number of models have been considered which lead to CM and we present specific results for two of these that differ in the treatment of

of the transfer reaction, here deposited Al replacing Ga in the underlayer. As discussed above, the models are distinguished by specifying the probabilities P_{ij} that a j (=Ga,Al) adatom occupies a site in the nth monolayer and an i adatom (=Ga,Al) occupies the corresponding site in monolayer n+1.

The first model emphasizes the greater bonding energy of mixed pairs

$$P_{ii} = 0.85,\ P_{ij} = 0.15;\quad i,j = Ga, Al,\ i \neq j \tag{13}$$

We also introduce the rule that a deposited Al adatom will replace a Ga adatom if that adatom is bonded below by anothe Ga adatom. The initial deposited monolayer composition is taken to be .45/.55 (Al/Ga). The above rules then lead to a CM structure as shown in figure 2.

In the model described above exchange occurred because of the relatively weak bonding between Ga atoms. We can implicitly include the exchange effect more directly by simply allowing $P_{GG} = P_{AG} = 1$ (we use A=Al, G=Ga for simplification), and in table 1 we show the results for a numbe of models of this type, all with one exception showing some degree of CM through six monolayers. Note that the CM here is not so pronounced and the variation is not between fixed values. The single exception found indicates that the window for CM may be fairly narrow.

These models have provided us with a means to quickly implement a verification that some form of CM occurs, at least in the early stages of growth. However, they are <u>ad hoc</u>, with the values of P_{ij} selected intuitively rather than by direct connection to the underlying kinetics. Further, there is no general result, even if a formal kind, that would provide the basis for a more detailed analysis - what you see is what you get! The importance of these results is that they provide a measure of assurance (at a cheap price) that CM can occur and consequently the motivation to formalize this approach

in a systematic manner that provides the basis for making contact with the relevant kinetic processes taking place.

IV. Probabilistic Composition Evolution Model

The most general features of the above models are in the use of the monolayer level rather than the time as the variable of change together with an implicit representation of the underlying kinetics through probabilities that specify the occupation of identical lattice sites in adjacent monolayers. This suggests that a master equation [7] approach, for which a variety of solution techniques have been developed, will provide a basis for a systematic description along these lines. An advantage of this approach over that of kinetic equations is that the resulting equations are linear; the tradeoff is that the kinetics are only implicit. We have pursued the former at the present time because this can be carried out independently of additional KMC studies that are required for the latter (we discuss this in the next section).

The most general model is based on the four probabilities P_{ij}, i,j=Al,Ga where

$$P_{GA} + P_{AA} = 1 = P_{AG} + P_{GG} \tag{14}$$

We have again used A and G to refer to Al and Ga and will denote their respective mole fractions in layer n as An, Gn. As before, P_{GA} is the probability or corresponding lattice sites in layers n, n+1 as having G in layer n+1, A in layer n, etc.. It is again convenient to disregard the substrate and call the first deposited layer, for which we prescribe the G and A compositions, as layer zero, and then for n = 1,2,...

$$Gn = P_{GG}Gn\text{-}1 + P_{GA}An\text{-}1, \quad An = P_{AG}Gn\text{-}1 + P_{AA}An\text{-}1 \tag{15}$$

The solution to these equations can be found through the use of generating

functions; define

$$G(s) = \sum_{n=o} Gn\ s^n\ ,\quad A(s) = \sum_{n=o} An\ s^n \tag{16}$$

where the variable s is taken small enough so the series converges but otherwise has no significance. If G(s) is determined then Gn is given either by expressing the former as a polynomial in s or by the value of the nth derivative at s=0. Multiplying eqs(15) by s^n and summing over n leads to

$$G(s)[1 - sP_{GG}] = Go + P_{GA}sA(s) \tag{17-a}$$

$$A(s)[1 - sP_{AA}] = Ao + P_{AG}sG(s) \tag{17-b}$$

so that eliminating A9s) and solving for G(s)

$$G(s) = [Go + s(P_{GA}Ao - P_{AA}Go]/[1 - s(P_{GG} + P_{AA}) + s^2(P_{GG}P_{AA} - P_{AG}P_{GA})] \tag{18}$$

Since $P_{ij} < 1$ we can always find a value of s so that the rhs of eq(18) can be expanded and written as a polynomial in s and the Gn identified as the coefficients of s^n. The first two coefficients are

$$G1 = P_{GA}Ao + P_{GG}Go\ ,\quad G2 = G1 + (1 - P_{AG} - P_{GA})(AoP_{GA} - GoP_{AG}) \tag{19}$$

and only simple algebra is required to determine additional terms.

Because of time limitations this model has not yet been fully analyzed. In particular, the connection between the P_{ij} and the kinetics has not yet been made. We conclude this section with a general observation followed by a specific example. First, it is apparent that for this model CM with fixed values only occurs for the special case $P_{AA} = P_{GG} = 0$, i.e. Al is always found over Ga and vice versa. While this is a possible consequence of the kinetics it is unlikely that other possibilities are excluded, e.g. we expect that four-dimensional P_{ij} phase space contains a (hyper-)volume corresponding to a range of system parameters for which CM occurs. As an

example that supports this conjecture we consider the simple case where $P_{GA} = 1-b$, $P_{AG} = 1-a$, $a,b \ll 1$, Go=0.55, G1=0.44, G2=0.54, then determine a,b for these values and use these to find G3 and G4. Neglecting higher order terms in a,b makes this relatively simple and we find

$$G1 = Ao + b(Go-Ao),\ G2 = Go + (a+b)(Ao-Go),\ G3 = Go + (a+2b) + Ao(1-a-2b),$$
$$G4 = Go -2(a+b)(Ao-Go) \qquad (20)$$

leading to a = 0.2, b = -0.1 (justifying the neglect of the higher order terms), and G3 = 0.45, G4 = 0.57 consistent with CM.

The preliminary results we have obtained using this model provide the motivation for further work along these lines. They also add support to the KMC results.

V. Kinetic Models of (Al,Ga)As Growth

The results reported in sections III and IV have provided us with the confidence that CM in the MBE growth ao (Al,Ga)As can be understood in terms of the kinetic processes taking place. The situation is more complex than that considered in section II since the two group III metals are deposited continuously and together. The models studied by KMC included Ga desoption, as has been experimentally found in previous work at the host laboratory [8]. In addition, analysis of the simulation data indicated that growth was not layer-by-layer so that at a given time a growing or filled monolayer might have a number of of less-filled layers growing above it. The implications of this are that the resulting transfer process would involve many layers and the kinetic equations become even more complicated because coverages as well as mole-fractions must be considered in the same equations. It would not be prudent to initiate an analytic study by looking at this most

general case, and therefore we have requested additional KMC simulation data at layer-by-layer growth conditions. These will determine whether a kinetic model along the lines of section II, generalized to include simultaneous deposition and desorption, can be used to study CM. In this regard we emphasize that the growth mechanism is not an explicit issue in the models formulated in the previous two sections and only has an influence on the connection between the kinetics and the P's. From the perspective of the present section this feature can be appreciated as a significant asset. In the event we find that CM occurs when the growth is two-dimensional we can disregard simultaneous exchange as a primary factor, and directly assess the relative effects of desorption and exchange between the two top-most layers. Should the KMC indicate that CM does not occur when growth is two-dimensional then we will have a more complicated scenario to consider. We have previously used mean-field models [9] to model collective effects, as would occur in three-dimensional growth, and would formulate our description for this situation this way. A growth layer would be described as interacting with two fields, and underlayer field and an overlayer field, which collectively incorporate the effects of all the other layers interacting with the designated layer. The analytical treatment of both two and three dimensional growth offers the prospect of stimulating and topical research of a very applied nature. The phenomena of CM provides an example of self-organization that has considerable potential for application to the design of novel heterostructures and devices based on these. First though we must have a better understanding of the relevant mechanisms through both experimental (including KMC) and theoretical study.

VI. Summary

The specific accomplishments of our ten week tour are:

* Improved theory for exchange mechanism in (In,Ga)As. These results are being put in a form for journal publication.
* Formulation and development of cellular automata and probability models to describe CM in (Al,Ga)As. The latter will be developed more fully and the results then prepared for publication, either jointly, or as a companion paper with KMC studies carried out by Drs. Dorsey and Mahalingam.
* Identification of the issues related to formulating a full kinetic model for the MBE deposition of (Al,Ga)As and the new KMC results required to resolve these prior to initiating this program.

The last two items indicate where we believe our future efforts can be most productively focused. The potential for utilizing CM in new graded superlattice structures is considerable once their formation is understood. The development of such new materials having applications ranging from LEDs, lasers, and detectors to the full alphabet of microelectronic devices is directly in line with the goals of Air Force materials reasearch programs.

REFERENCES

1. J. Moison et.al., J. Cryst. Growth 111, 141 (1991).
2. M. Krishnamurthy et.al., J. Vac. Sci. Technol. B11, 1384 (1993).
3. K. Mahalingam and D. Dorsey, to be published.
4. For a comprehensible review see S. Kaufman, At Home in the Universe (Oxford University Press, Oxford, U.K., 1995).
5. H. Yamaguchi and Y. Houkoshi, J. Appl. Phys. 68, 1610 (1990).
6. Analysis of samples grown at the required conditions is scheduled for Fall 1996; if additional samples and analysis are required the time-line

will be extended even further

7. see,e.g. N. Van Kampen,Stochastic Processes in Physics and Chemistry, (North-Holland, Amsterdam, 1981)VI.2.

8. K. Evans et.al., J. Vac. Sci. Technol. B7, 2428 (1991).

9. L.Borovinski and L. Tzindergozen, A. Dokl. Akad. Nauk. USSR 183, 1308 (1968).

P_{GA}	P_{AA}	ML1	2	3	4	5	6
0.6	0.4	0.55	0.67	0.60	0.64	0.61	0.67
0.7	0.3	0.55	0.61	0.57	0.60	0.58	0.63
0.5	0.5	0.55	0.72	0.64	0.68	0.66	0.73
0.6	0.4	0.60	0.64	0.62	0.63	0.62	0.67
0.7	0.3	0.60	0.58	0.60	0.58	0.59	0.62
0.5	0.5	0.60	0.70	0.65	0.68	0.66	0.71

TABLE 1: Al mole fraction by monolayer for simple cellular automata with $P_{GG} = 0$, $P_{AG} = 1$.

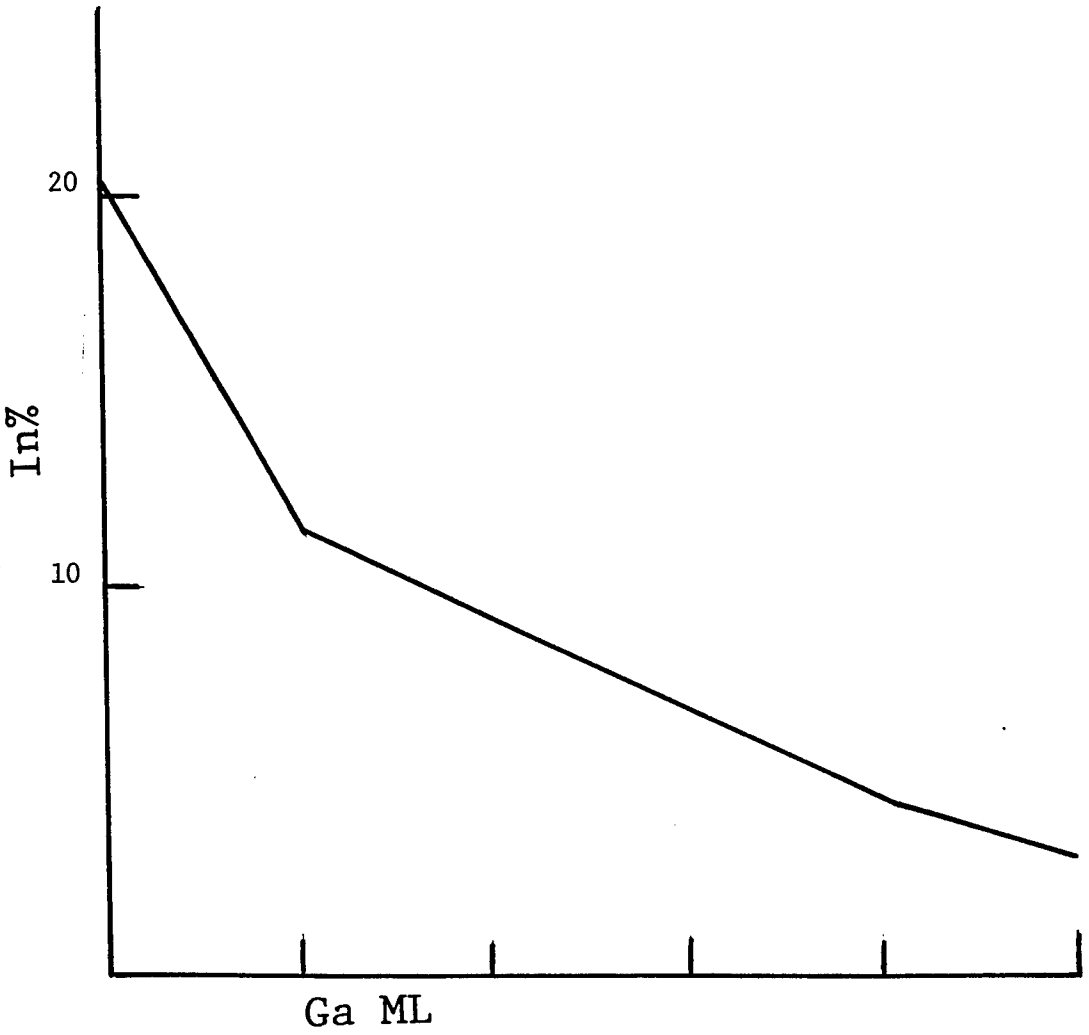
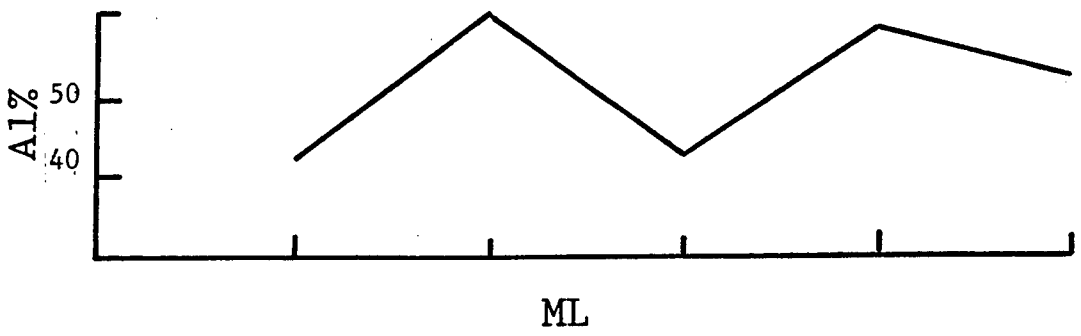


Figure 1. In composition x Ga monolayer for P = 0.98. The surface is towards the right; the initial In monolayer is located at zero on the Ga Ml axis. Five Ga ML are shown. The same "curve" is obtained from eqs (3), (4) with k = 0.78 and eqs (9), (11) with the mass balance relationships and k, R, T chosen so that P = 0.98.



		GAA(.15)	(.55)(.85)(.15)
	GA(.85)	GAG(.85)	(.55)(.85)(.85)
G(.55)			
	GG(.15)	GGG(.15)	
		GGA(.10)	
		GAG(.75)	(.55)(.15)(.75)
		AGA(.85)	
	AG(.85)	AGG(.15)	
A(.45)			
	AA(.15)	AAA(.15)	(.45)(..5)(.15)
		AAG(.85)	(.45)(.15)(.85)

Figure 2. Top: Al composition through five ML. Bottom: Schematic for determining the composition of ML 2. Here Ga=G, Al=A, and the probability of adding the rightmost adatom is shown. ML 2 is finalized when ML 3 is filled as shown in the rigtmost column. The probability of each outcome that results in an Al adatom in ML 2 is also shown far right. The Al % is the sum of these.

EFFECT OF HUMIDITY ON WEAR OF M-50 STEEL WITH A KRYTOX LUBRICANT

Larry S. Helmick
Professor of Chemistry
Department of Science and Mathematics

Cedarville College
Box 601
Cedarville, OH 45314

Final Report for:
Summer Faculty Research Program
Wright Laboratory

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

and

Materials Directorate
Wright Laboratory

August 1996

EFFECT OF HUMIDITY ON WEAR OF M-50 STEEL WITH A KRYTOX LUBRICANT

Larry S. Helmick
Professor of Chemistry
Department of Science and Mathematics
Cedarville College

Abstract

Using a Cameron-Plint tribometer under controlled environmental conditions, wear of M-50 steel with a Krytox (a branched perfluoropolyalkyl ether) lubricant was studied under boundary lubrication conditions at 150°C in air with relative humidity ranging from 1% to 95%. Both wear and friction decrease sharply as humidity is increased from 1 to 20%, then are constant as humidity increases to 95%. Thus, wear is highly dependent on humidity when relative humidity is less than 20%. The similar effect of humidity on wear previously observed for Fomblin Z, a linear perfluoropolyalkyl ether containing difluoroacetal groups, and Demnum S, a linear perfluoropolyalkylether which does not contain difluoroacetal groups, indicates that this may be a general property of all commercially available perfluoropolyalkylethers.

EFFECT OF HUMIDITY ON WEAR OF M-50 STEEL WITH A KRYTOX LUBRICANT

Larry S. Helmick

Introduction

Perfluoropolyalkyl ethers (PFPAE) are presently being investigated as liquid lubricants for aerospace applications (1,2). PFPAE fluids have been tested with the four-ball tribometer (3,4) and the newer Cameron-Plint reciprocating tribometer under sliding boundary lubrication conditions with the test cell exposed to the environmental atmosphere. When testing Fomblin Z, a linear PFPAE fluid containing difluoroacetal groups (OCF_2), with steel specimens with either instrument, erratic results have sometimes been obtained. A recent investigation revealed that environmental relative humidity has a major effect on friction and wear for this fluid (5), and therefore may be a major cause of these erratic results. That study was then broadened to include Demnum S (6), a linear PFPAE fluid which does not possess difluoroacetal groups. A similar effect of humidity was observed. Therefore, the current study was initiated to broaden the investigation still further and determine whether humidity also affects friction and wear for a branched PFPAE fluid. If humidity affects friction and wear for

this fluid also, then it may be that this is a general property of all commercially available PFPAE fluids, and is not determined by a specific molecular structure. Furthermore, it will be necessary to measure and control humidity when evaluating potential antiwear additives (7-9) under sliding boundary lubrication conditions on steel specimens with all commercially available PFPAE fluids.

Experimental Procedure

The commercially available fluid used in this study was Krytox 143AC, a branched PFPAE with a C/O ratio of 3/1, which does not contain difluoroacetal groups. Data obtained with this fluid are compared to data previously obtained under similar conditions using Fomblin Z-04 (5), a linear PFPAE with a C/O ratio of 1.4/1, which does contain difluoroacetal groups, and Demnum S-65(6), a linear PFPAE with a C/O ratio of 3/1 which does not contain difluoroacetal groups.

Krytox was tested in air under sliding boundary lubrication conditions using a modified Cameron-Plint Tribometer with a controlled environmental chamber and automatic data acquisition system. A schematic diagram and a full description of this instrument are given in references 5 and 10.

The detailed test procedure is described in reference 5, but a brief review here may be helpful. An M-50 steel cylinder was

loaded and slid with a reciprocating motion against an M-50 flat specimen. Friction was monitored throughout the 5 hour runs and recorded at 30 second intervals. An average coefficient of friction was calculated for each run by dividing the average friction by the 250 N load.

Data were obtained with specimen temperatures of 150°C +/- 2°C and ambient atmospheric temperatures (29 +/- 3°C). Specimen temperatures and electrical resistance between the cylinder and disk were also recorded at 30 second intervals throughout the runs. Formation of films on the surface of the wear scar were indicated by an increase in contact point electrical resistance and confirmed by grazing angle microscope / Fourier transform infrared (GAM-FTIR) spectroscopy, as described in references 11 and 12.

Relative humidity (RH) was measured (+/- 0.1%) at ambient temperature and recorded at 30 second intervals as well. Humidities below ambient (approximately 50%) were obtained by controlling the flow rate of dry air into the environment chamber, and were constant to +/- 0.3% or less. Humidities above ambient were obtained by controlling the flow rate of air through a Tatung TUH-500H Ultrasonic Humidifier, and were constant to +/- 1.5% or less.

Areas of wear scars were determined by measuring (+/- 1 X 10^{-3}mm) the length and width of the wear scar on the cylinder at the end of the 5 hr runs.

All temperature, humidity, friction, and wear data for Krytox are recorded in Table 1 and plotted in Figures 1 and 2, along with data previously reported for Fomblin (5) and Demnum (6), for easy comparison.

Results

Figure 1 shows the effect of humidity on wear scar area at 150°C for all three fluids. There is a sharp decrease in wear scar area as humidity increases from 1 to 20% for Krytox as well as for Fomblin and Demnum. No additional decrease is obvious for any of the fluids as humidity increases from 20 to 95%.

Figure 2 shows the effect of humidity on the average coefficient of friction at 150°C for all three fluids. A sharp decrease in friction is observed as humidity increases from 1 to 20% for all three fluids, but no additional decrease occurs for any of the fluids as humidity increases to 100%. Fomblin exhibits a lower coefficient of friction than either Demnum or Krytox because of the presence of difluoroacetal units in its polymeric molecular structure.

Figure 3 shows the variation in electrical resistance at the contact point with time as well as with relative humidity for Krytox. Fomblin and Demnum exhibit similar curves. Although resistance is relatively erratic, generally it is high during the low load run-in period, due to poor metal-metal contact, but then drops rapidly when the 250 N load is applied and metal-metal contact improves. For low humidity runs, it then stays relatively low, demonstrating good metal-metal contact throughout the 5 hour run. For high humidity runs, however, it rapidly increases again, indicating that an electrically insulating film is being formed on the wear scar surface as the run continues. The formation of this film at high humidities has now been observed by GAM-FTIR spectroscopy for Krytox just as it was previously for Fomblin (11) and Demnum (6).

Discussion

The dependence of both friction and wear on humidity was previously observed for Fomblin and assumed to be related to the presence of difluoroacetal groups (5) in the linear molecular structure. A similar dependence was then observed for Demnum (6), which has a linear structure but does not possess difluoroacetal groups. Thus it is clear that this dependence is not due solely to the presence of difluoroacetal groups in Fomblin. This dependence has now been observed for Krytox, which has a branched molecular structure. Therefore, the effect of humidity on wear is a property of branched as well as linear

fluids, and may very well be a general property characteristic of all commercially available PFPAE fluids.

The cause of the high wear at low humidity is not yet clear. It may be corrosive wear, in which case it would be directly related to the rate of decomposition of the PFPAE fluid to form corrosive products. Or it may be abrasive wear, in which case it would be inversely related to the rate of formation and thickness of the film which is produced at higher humidities, which in turn may also be related to the rate of decomposition of the PFPAE fluid. More likely, it is a combination of both of these processes rather than a result of just one or the other. In any event, the reaction mechanisms for decomposition of PFPAE fluids under tribological conditions, and the chemistry of film formation are being investigated (11-16).

Conclusions

Under boundary lubrication conditions, wear of M-50 steel with linear and branched perfluoropolyalkylether lubricants with and without difluoroacetal groups depends strongly on relative humidity for humidities below 20%, but is nearly independent of humidity above 20%. Therefore, this dependence cannot be attributed solely to the presence of difluoroacetal groups in the PFPAE lubricant or to the linear or branched molecular structure. It seems to be a fundamental property of all commercially available PFPAE fluids. Therefore, for boundary lubrication wear

tests performed below 20% relative humidity, it is important that humidity be monitored and carefully controlled.

Disclaimer

The U. S. Air Force does not endorse for other purposes or criticize the materials used in this study.

References

1. Fusaro, R.L., and Khonsari, M.M., "Liquid Lubrication for Space Applications," NASA TM-105198, 1992.

2. Fusaro, R.L., "Tribology Needs for Future Space and Aeronautical Systems," NASA TM-104525, 1991.

3. Masuko, M., Jones, W.R., Jr., and Helmick, L.S., "Tribological Characteristics of Perfluoropolyether Liquid Lubricants Under Sliding Conditions in High Vacuum," NASA TM-106257, 1993.

4. Masuko, M., Jones, W.R., Jr., Jansen, R., Ebihara, B., Pepper, S.V., and Helmick, L.S., "A Vacuum Four-Ball Tribometer to Evaluate Liquid Lubricants for Space Applications," NASA TM-106264, 1994.

5. Helmick, L.S., and Sharma, S.K., "Effect of Humidity on Friction and Wear for a Linear Perfluoropolyalkyl Ether Fluid Under Boundary Lubrication Conditions," Lubr. Eng., 52, 6, 437-442 (1996).

6. Helmick, L.S., Gschwender, L.J., Sharma, S.K., Snyder, C.E., Jr., Liang, J.C., and Fultz, F.W., "The Effect of Humidity on the Wear Behavior of Bearing Steels with $R_fO(n\text{-}C_3F_6O)_xR_f$ Perfluoropolyalkylether Fluids and Formulations," Submitted for publication (1996).

7. Gschwender, L.J., Snyder, C.E. Jr., and Fultz, G.W., "Soluble Additives for Perfluoropolyalkylether Liquid Lubricants," Lubr. Eng., 49, 9, 702-708 (1993).

8. Srinivasan, P., Corti, C., Montagna, L. and Savelli, P., "Soluble Additives for Perfluorinated Lubricants," JSL, 10, 2, 143-164 (1993).

9. Sharma, S.K., Gschwender, L.J., and Snyder, C.E., Jr., "Development of a Soluble Lubricity Additive for Perfluoropolyalkylether Fluids," JSL, 7, 1, 15-23 (1990).

10. Cavdar, B., Sharma, S.K., and Gschwender, L.J., "Wear-Reducing Surface Films Formed by a Fluorinated Sulfonamide Additive in a Chlorotrifluoroethylene-based Fluid," Lubr. Eng., 50, 11, 895-902 (1994).

11. Liang, J., and Helmick, L.S., "Tribochemistry of a PFPAE Fluid on M-50 Surfaces by FTIR Spectroscopy," STLE Preprint No. 95-TC-5A-2,(1995).

12. Liang, J., Tung, Y.S., Henderson, D.O., and Helmick, L.S., "A Study of Boundary Lubricating Thin Films Produced from a

Perfluoropolyalkylether Fluid on M-50 Surfaces Part 2: Humidity Effect and Topography Studies," Submitted for Publication (1996).

13. Karis, T.E., Novotny, V.J., and Johnson, R.D., "Mechanical Scission of Perfluoropolyethers," J. Appl. Polym. Sci., 50, 1357-1368 (1993).

14. Zehe, M.J., and Fout, O.D., "Acidic attack of Perfluorinated Alkyl Ether Lubricant Molecules by Metal Oxide Surfaces," Trib. Trans., 33, 634-640 (1990).

15. Kasai, P.H. "Perfluoropolyethers: Intramolecular Disproportionation," Macromolecules, 25, 6791-6799 (1992).

16. Vurens, G., Zehringer, R., and Saperstein, D., "The Decomposition Mechanism of Perfluoropolyether Lubricants during Wear," Chapter 10 in Surface Science Investigations in Tribology, Chung, Y.W., Homola, A.M., Sheet, G.B., eds., ACS Symposium Series, American Chemical Society, 1992.

TABLE 1
Temperature, Humidity, Friction, and Wear Data

Run Number	Temperature (°C)	Rel.Humid. (%)	Avg. Coef. of Friction	Wear Area (mm^2)
1	150	1	0.123	4.120
2	150	1	0.127	4.108
3	150	5	0.139	4.150
4	150	5	0.130	3.493
5	150	10	0.139	3.284
6	150	10	0.135	3.505
7	150	15	0.112	1.156
8	150	20	0.107	0.675
9	150	40	0.114	0.677
10	150	60	0.114	0.624
11	150	95	0.117	0.852

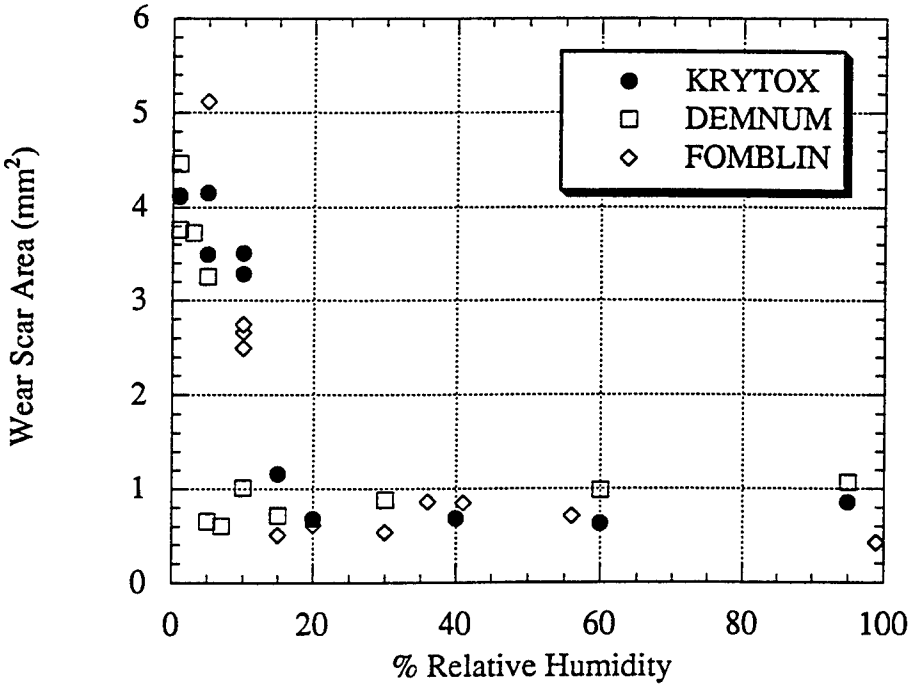


Fig. 1. Effect of Humidity on Wear

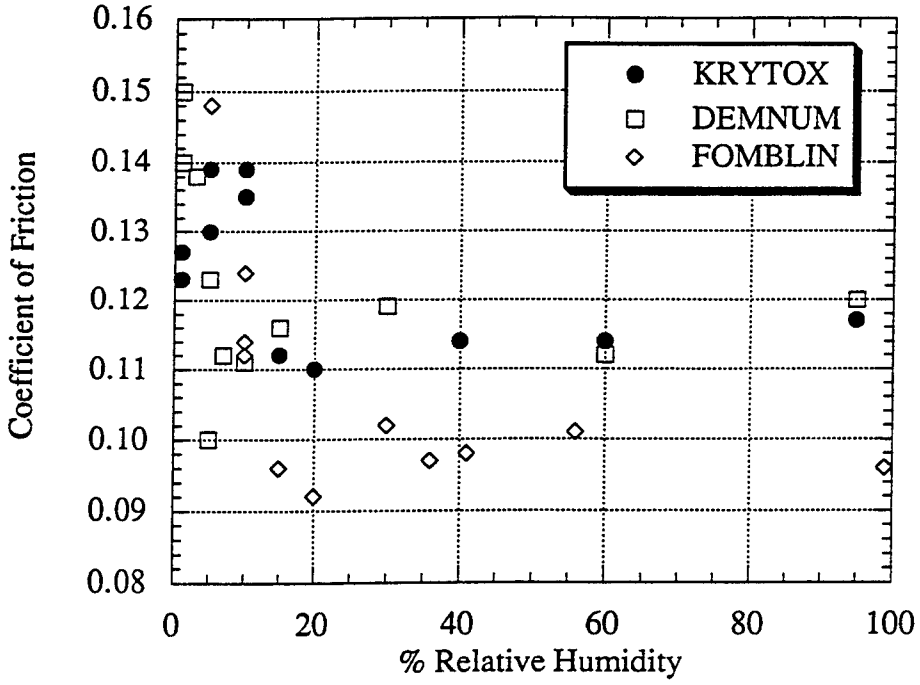


Fig. 2. Effect of Humidity on Friction

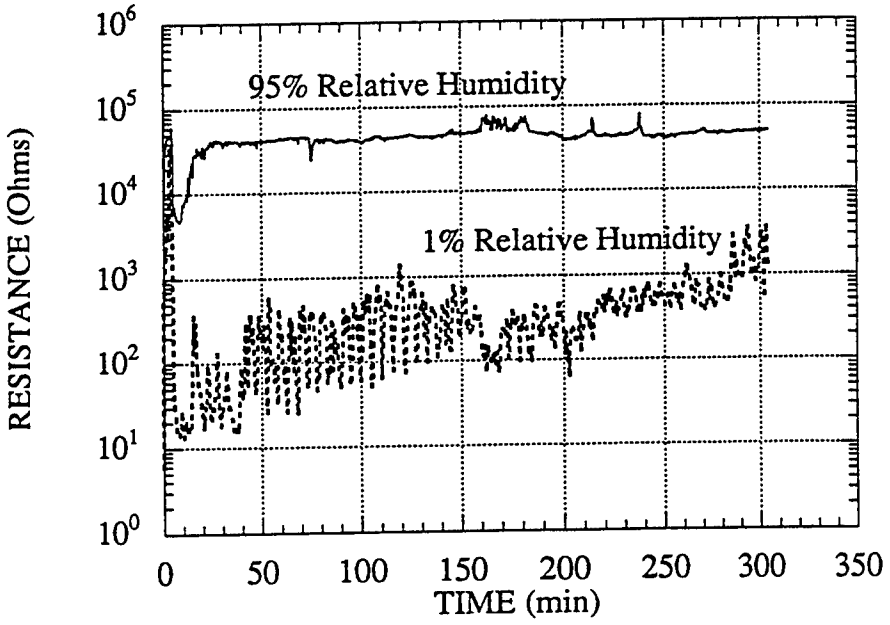


Fig. 3. Effect of Humidity on Resistance for Krytox

HYPERBARIC OXYGEN EFFECTS ON THE POSTISCHEMIC BRAIN

Kenneth Hensley

Research Associate, Oklahoma Medical Research Foundation
825 NE 13th St.
Oklahoma City, OK 73104

Adjunct Assistant Professor, Chemistry and Biochemistry
University of Oklahoma
Norman, OK

Final Report for:
Summer Faculty Research Program
Wright Patterson Air Force Base

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

and

Wright Patterson Air Force Base

September 1996

HYPERBARIC OXYGEN EFFECTS ON THE POSTISCHEMIC BRAIN

Kenneth Hensley

Associate Research Scientist, Oklahoma Medical Research Foundation
825 NE 13th St.
Oklahoma City, OK 73104

Adjunct Assistant Professor, Chemistry and Biochemistry
University of Oklahoma
Norman, OK

Abstract

Hyperbaric oxygen therapy has been previously studied as a method of accelerating wound healing, and is used medically in the treatment of carbon monoxide poisoning, nitrogen embolism, and burn trauma. Hyperbaric oxygen therapy has been suggested as a method for counteracting some aspects of ischemia / reperfusion pathophysiology, by augmenting tissue oxygenation or by other mechanisms. In this study, an animal model of ischemia / reperfusion was employed to assess the effects of hyperbaric oxygenation on biomarkers of tissue damage in the postischemic brain. Mongolian gerbils were surgically implanted with carotid loops, which could be drawn tight to induce a defined period of near-total cerebral ischemia. Animals were exposed to two atmospheres of 100% oxygen for one hour prior to, or following ischemia. Animals were killed at time points after the experimentally induced ischemia / reperfusion, and organs archived for latter analysis. Initial results indicate that hyperbaric oxygen treatment does not counteract ischemia-induced decrease in brain glutamine synthetase activity, which has been repeatedly shown to index brain pathology in animal models of ischemia / reperfusion. To the contrary, in the current study, hyperbaric treatment following ischemia / reperfusion significantly exacerbated ischemia-induced loss of glutamine synthetase activity in the gerbil hippocampus as assayed in tissue removed 24 hours following the ischemic event. Furthermore, animals receiving the experimental ischemic insult demonstrated significantly increased hippocampal creatine kinase activities 24 hours after reperfusion, whereas animals receiving hyperbaric oxygenation (either before or after ischemia / reperfusion) failed to demonstrate this CK response.

HYPERBARIC OXYGEN EFFECTS ON THE POSTISCHEMIC BRAIN

Kenneth Hensley

Introduction

Ischemia / reperfusion injury, whether suffered by the central nervous system or by peripheral tissue, is a matter of tremendous medical significance. Furthermore, head trauma produces a major element of military (and civilian) casualties, and a major portion of the mortality and morbidity they cause. Essentially all severe, but not immediately fatal, head or spinal trauma involves a critical element of ischemia due to some combination of concussion, vascular disruption, and edema within a confined space. Cerebral ischemia can also be induced by cardiovascular shock or toxic edema. In recent years, evidence has accumulated that damage to ischemic tissue largely occurs during the reperfusion phase or at times thereafter, as the tissue attempts to recover from the ischemic insult. Furthermore, reperfusion-associated injury seems to be strongly associated with the generation of toxic, bioreactive oxyradicals by physiologically compromised tissue, which exacerbates damage to the energy-compromised postischemic cells.

Tissue undergoing recovery from ischemia therefore faces a paradoxical challenge: the need to restore metabolism, necessitating effective reperfusion and oxygenation, and the need to counteract oxygen toxicity. Hyperbaric oxygen (HBO) treatment has demonstrated medical efficacy in promoting wound healing and enhancing perfusion of vasocompromised tissue. Zamboni and colleagues report that HBO treatment of postischemic gracilis muscle inhibits leukocyte adherence to endothelium, which might serve to mitigate a significant component of postischemic inflammatory stress {1}. Furthermore, brief exposure of cells to a mild oxidant stress elicits the expression of numerous genes including heat shock proteins, which may serve a protective function to cells challenged by energy deprivation or chemical sensitization. It has therefore been suggested that hyperbaric oxygen therapy may have medical relevance in promoting recovery from an ischemic state, if administered to subject immediately following the ischemia.

This study has attempted to begin addressing the role of postischemic tissue oxyg en tension in modulating recovery of brain tissue from an ischemic insult, and in the postischemic presentation of inflammatory sequelae. The goal of this study is to more closely define oxygen-

dependent mechanisms of ischemia / reperfusion injury, which may be amenable to therapeutic intervention. This goal is critical to improvement of medical management of head and spinal injuries. Successfully defining these mechanisms and interactions may also have important implications for understanding and management of ischemia/ reperfusion phenomena in other tissues involved in military casualties such as crush syndrome, thermal burns, and blast injury.

A rodent model of unanesthetized cerebral ischemia was employed {1}. Mongolian gerbils were subjected to 10 minutes of forebrain ischemia, which has been reported to induce alterations in trauma-sensitive enzyme activities [e.g.glutamine synthetase (GS, EC 6.3.1.2) and creatine kinase (CK, EC 2.7.3.2)] and cellular death. Animals were exposed to 2 atmospheres of 100% oxygen prior to or following the ischemic event. Brain tissue was removed 2 or 24 hours after ischemia / reperfusion for analysis of biomarkers (glutamine synthetase and creatine kinase activities {2}). In ongoing work, the tissue will be analyzed for quantifiable sequelae indicative of an inflammatory response, which growing evidence suggests may be involved in the protracted neurodegeneration that follows cerebral ischemia. Inflammatory sequelae to be measured include messenger RNA for inducible nitric oxide synthetase (iNOS), heat shock protein 70 (HSP70), as well as ornithine decarboxylase (ODC), polyubiquitin, and vascular cell adhesion molecule-1 (VCAM-1). These various gene products are to be detected by reverse-transcriptase polymerase chain reaction (RT-PCR) in hippocampus and cerebral cortex.

Methodology

Male Mongolian gerbils, 3-5 months of age (approximately 100 g, Tumblebrook Farms, W. Brookfield, MA) were implemented in an established carotid artery occlusion model of ischemia / reperfusion {2-3}.The protocol for this treatment is as follows:

Animals were anesthetized with pentobarbital (40 mg/kg intraperitoneal injection). The ventral portion of the neck was surgically opened and the carotid arteries exposed. A loop of surgical floss was placed loosely about each artery. The ends of the loop were threaded through a small double-lumen catheter made to extend through the dorsal portion of the animal's neck. The catheter was secured with a small drop of nontoxic acrylic adhesive, the neck stitched close, and animals allowed to recover for 48 hours before ischemia.

Ischemia was initiated by applying tension to the portion of surgical floss extending from the catheters. Following 10 minutes of cortical ischemia the tension was released from the floss. Animals were returned to the home cage, and killed by decapitation at a designated time following ischemia.

Hyperbaric oxygen treatment is used medically and experimentally at Wright Patterson Air Force Base. Experimentally, animals are placed in a pressure- and gas-regulated steel drum and slowly acclimated to hyperbaric oxygen (100 %, 2 atmospheres) for 1 hour at ambient temperature (approximately 23°C). Pressure was adjusted over the course of 3-5 minutes. Duration of hyperbaric oxygenation (60 minutes) denotes actual time at full 2 atm. No seizures or other gross effects were observed in animals as a function of hyperbaric treatment.

HBO treatment preceding ischemia. Animals treated with HBO received a 60 min. exposure to HBO beginning 90 min. and terminating 30 min. prior to the ischemic event. Control groups were included to account for the effects of HBO only in the absence of ischemia; ischemia without HBO; and a "blank" group received neither ischemia nor HBO.

HBO treatment following ischemia. Animals were subjected to 10 minutes of global forebrain ischemia, followed by 60 minutes of HBO (2 atmospheres as described above) and sacrificed 2 hours following ischemia (one hour following completion of HBO) or 24 hours following ischemia. Appropriate control groups were included (HBO only or ischemia only, or sham surgery groups which had carotid implants but no deliberate carotid occlusion).

Animals were subjected to 10 minutes of forebrain ischemia by occlusion of both carotid arteries. Animals were then allowed to recover for 2 or 24 hours. Brain tissue removed from the animals following sacrifice was apportioned into sections for reverse-transcription polymerase chain reaction (RT-PCR) and analysis of marker enzyme activities. Tissue was frozen intact by immersion in liquid nitrogen, and maintained at less than -70°C until analysis. Tissue from hippocampus, cortical mantle, and cerebellum was stored separately. In addition, tissue was collected from liver, lung, heart, kidney, gracilis muscle, and blood (collected in heparinized tubes), frozen in liquid nitrogen, and archived. For enzyme assays, tissue was thawed by immersion in isolation buffer (10 mM HEPES, 5 μg/mL pepstatin, μg/mL aprotinin, 100 μg/mL phenylmethylsulfonyl fluoride, 2 mM EGTA, pH 7.4). Tissue was homogenized thoroughly at 4°C and allowed to incubate at 4°C in the isolation buffer for 2-3 hours. Protein concentration was assayed by the Lowry method {4}. Glutamine synthetase assays were performed essentially as reported by Miller et al {5} using γ-glutamyl hydroxymate as a standard,

and corrected for nonspecific glutaminase activity. Creatine kinase activity was measured by means of the Voges-Proskaur reaction as reported by Hughes {6} using a commercially available kit (Sigma Chemical, St. Louis, MO).

Results

The experimental objectives of this study, i.e. the execution of carotid occlusion and hyperbaric treatment, were accomplished. The major problem encountered was an unexpectedly high mortality of animals due to ischemia. Approximately 40 % of animals died prematurely (i.e. before deliberate sacrifice) during ischemia, in the 24 hours following ischemia, or in several instances, during recovery from surgery. This is much higher than the projected mortality of 15 % prior to the initiation of the project. As a result, some experimental groups were abbreviated from the desired number of 5 animals / group anticipated at the inception of the study. Hence, the group of animals which received HBO + 2 hours recovery (no ischemia) contained 3 animals rather the projected 5. The group designated to receive HBO prior to ischemia, followed by 24 hours recovery, contained 4 animals rather than the desired 5. The group designated to receive HBO prior to ischemia, followed by 2 hours recovery, contained a final of 3 animals. Nonetheless, the most experimentally relevant groups (control animals, HBO treatments, ischemic groups receiving no HBO or HBO following ischemia) contained a final total of 5-6 animals each, thereby allowing an accurate assessment of experimental endpoints.

Animals receiving ischemia, regardless of HBO treatment, were lethargic and in some cases demonstrated seizure activity for several hours following ischemia / reperfusion. In many cases the lethargy was easily recognizable up to 24 hours following the ischemic event. There were no signs of seizures in animals receiving only HBO treatment. Otherwise, no gross behavioral differences were noted among animal groups.

Hippocampal glutamine synthetase activity measurements indicate that animals receiving experimental ischemia / reperfusion did encounter a significant degree of central nervous system trauma (Figure 1). As reported by Oliver and colleagues, brain GS activity decreases at 2 hours following ischemia / reperfusion, then rebounds to near baseline levels at 24 hours, in the absence of hyperbaric treatment. On the contrary, animals receiving HBO treatment after ischemia / reperfusion exhibit a loss of GS activity which fails to rebound over the 24 hour postischemic period. HBO treatment in the absence of ischemia / reperfusion, resulted in a marginal but nonsignificant decrease in GS activities at

1 and 24 hours following HBO. These data are consistent with previous observations {7} that chronic exposure to 100 % isobaric oxygen elicits a loss in brain GS activity, which young animals can recover, but from which old animals cannot. These data may complement recently reported findings {8} that HBO treatment, if taken to the point of producing seizures, significantly elevates brain lipid peroxides and protein carbonyls levels, indicating enhanced oxidative damage to sensitive biomolecules. These preliminary data suggest that tissue oxygenation in the period immediately following ischemia / reperfusion affects the ultimate fate of the traumatized tissue. Furthermore, events occurring after ischemia (in this case HBO treatment) can significantly affect delayed tissue response to injury, hence, there may be a therapeutic window after ischemia / reperfusion when protective measures can be taken to mitigate tissue injury.

Creatine kinase activity has been suggested as a readily measured index of the extent of traumatic damage to the central nervous system. Shiavone and Kaldor {9}and Dubo and colleagues {10} report elevated serum CK levels in patients having experienced acute ischemia / reperfusion, head injury, or aflicted with meningitis or schitzophrenia. Creatine kinase responds to a variety of cardiovascular and pulmonary stresses including myocardial infarction, pneumonia, and cardiac massage; moreover, serum and central nervous system levels of CK do not necessarily correlate {11} so that interpretation of this parameter must be cautioned.

As shown in Figure 2, hippocampal CK activity increased significantly within 24 hours following ischemia / reperfusion injury in animals which did not receive HBO treatment. CK activity also increased significantly within 24 hours following the hyperbaric oxygen treatment, in animals which did not receive ischemia / reperfusion. Paradoxically, animals receiving HBO following ischemia / reperfusion injury did not demonstrate significantly increased hippocampal CK activity. Likewise, in three animals which received HBO treatment prior to ischemia, no elevated CK activity was observed in the hippocampus at 24 hours post reperfusion.

Effects observed 24 hours following reperfusion likely reflect, among other phenomena, gene inductive events. A consensus is emerging that oxyradical-mediated stresses can profoundly influence transcription factor activation and gene induction {reviewed in 13}. The effect of HBO in counteracting ischemia / reperfusion-induced increase in hippocampal CK may represent differential gene inductive processes occurring among the experimental groups, with CK activity being upregulated more slowly in animals receiving HBO treatment. Since CK catalyzes the reversible

phosphorylation of adenosine-5'-diphosphate by phosphocreatine to form adenosine triphosphate (ATP), upregulation of CK may represent an attempt by the postischemic brain to repair ischemia / reperfusion-induced damage. If HBO suppresses postischemic gene induction in a widespread manner, then the failure of GS activity to rebound in HBO-treated, postischemic animals can be interpreted by differential degradation and biosynthesis of this metabolic enzyme.

Clearly, more detailed analysis of the effects of oxygen-modulating paradigms in influencing tramatic injury, is justified; the tissue collected in this study is part of this ongoing research. Further analysis of gene inductive events following ischemia / reperfusion may clarify the mechanism of oxygen in modulating postischemic trauma.

References

1 Zamboni W.A., Roth A.C., Russell R.C., Graham B., Suchy H., and Kucan, J.O. *Plast. Recon. Sur.* **91**: 1110-1123, 1993.

2 Chandler M.J., Deleo J., and Carney J.M.. *J. Pharmacol. Method.* **14**: 137-146, 1985.

3 Oliver C.N., Starke-Reed P.E., Stadtman E.R., Liu G.J., Carney J.M., and Floyd R.A. *Proc. Natl. Acad. Sci. USA* **87**: 5144-5147, 1990.

4 Lowry O.H., Rosebrough N.J., Farr A.L, and Randall, R.J. *J. Biol. Chem.* **151**: 265-275, 1951.

5 Miller R.E., Hadenberg R., and Gersham H. *Proc. Natl. Acad. Sci. USA* **75**: 1418-1422, 1978.

6 Hughes P.P. *Clin. Chim. Acta* 7: 579, 1962.

7 Hensley K., Howard B., Carney J.M., and Butterfield D.A. *Biochim. Biophys. Acta* **1270**: 203-206, 1996.

8 Chavko M. and Harabin A.L. *Free Rad. Biol. Med.* **20**: 973-978, 1996.

9 Schiavone D.J. and Kaldor J. *Med. J. Aust.* **2**: 790, 1965.

10 Dubo H., Park D.C., Pennington R.J.T., Kalbag R.M., and Walton J.N. *Lancet* **2**: 743, 1967.

11 Smith A.F. *Lancet* **2**: 178, 1967.

12 Remacle J., Raes M., Toussaint O., Renard P., and Rao G. *Mutation Res.* **316**: 103-122, 1995.

Figure 1 Hippocampal glutamine synthetase activity following ischemia / reperfusion and hyperbaric oxygen treatment. Groups shown in this figure are (lower bar to upper bar):

CONTROL : Control animals, sham surgery performed, no ischemia, no HBO. N = 5.

HBO / 1 H : Animals received 60 minutes HBO, 2 atm, 100 % oxygen, and were sacrificed 1 hour thereafter. N = 4.

ISC / 2 H : Animals received 10 minutes ischemia followed by 2 hours recovery (reperfusion) prior to sacrifice. N = 6.

ISC / HBO / 1 H : Animals received 10 minutes ischemia followed by 1 hour reperfusion under HBO, 2 atm, 100 % oxygen, followed by one additional hour reperfusion under ambient atmospheric conditions. N = 5.

ISC / HBO / 24 H: Animals received 10 minutes ischemia followed by 1 hour reperfusion under HBO, 2 atm, 100 % oxygen, followed by 24 hours reperfusion under ambient atmospheric conditions. N = 5.

HBO/ ISC / 24 H: Animals received 1 hour HBO, 2 atm, 100 % oxygen, followed 30 minutes later by 10 min. ischemia and a subsequent 24 hours reperfusion under ambient atmospheric conditions. N = 3.

ISC / 24 H: Animals received 10 minutes ischemia followed 24 hours reperfusion under ambient atmospheric conditions. N = 6.

HBO / 24 H: Animals received 60 minutes HBO, 2 atm, 100 % oxygen, and were sacrificed 24 hours thereafter. N = 5.

Error bars represent standard error about the mean; asterisk denotes significance decrease relative to controls, $p < 0.05$ (HBO / 1 H) or $p < 0.02$ (ISC / HBO / 24 H) by Student's t-test.

Figure 2. Hippocampal creatine kinase activity following ischemia / reperfusion and hyperbaric oxygen treatment. Groups shown in this figure are identical to those in figure 1. Error bars represent standard error about the mean. **X** = Significance relative to control (sham surgery) group, $p < 0.02$; * = significant effect of HBO on ischemia-induced increase in CK activity, $p < 0.02$, Student's t-test.

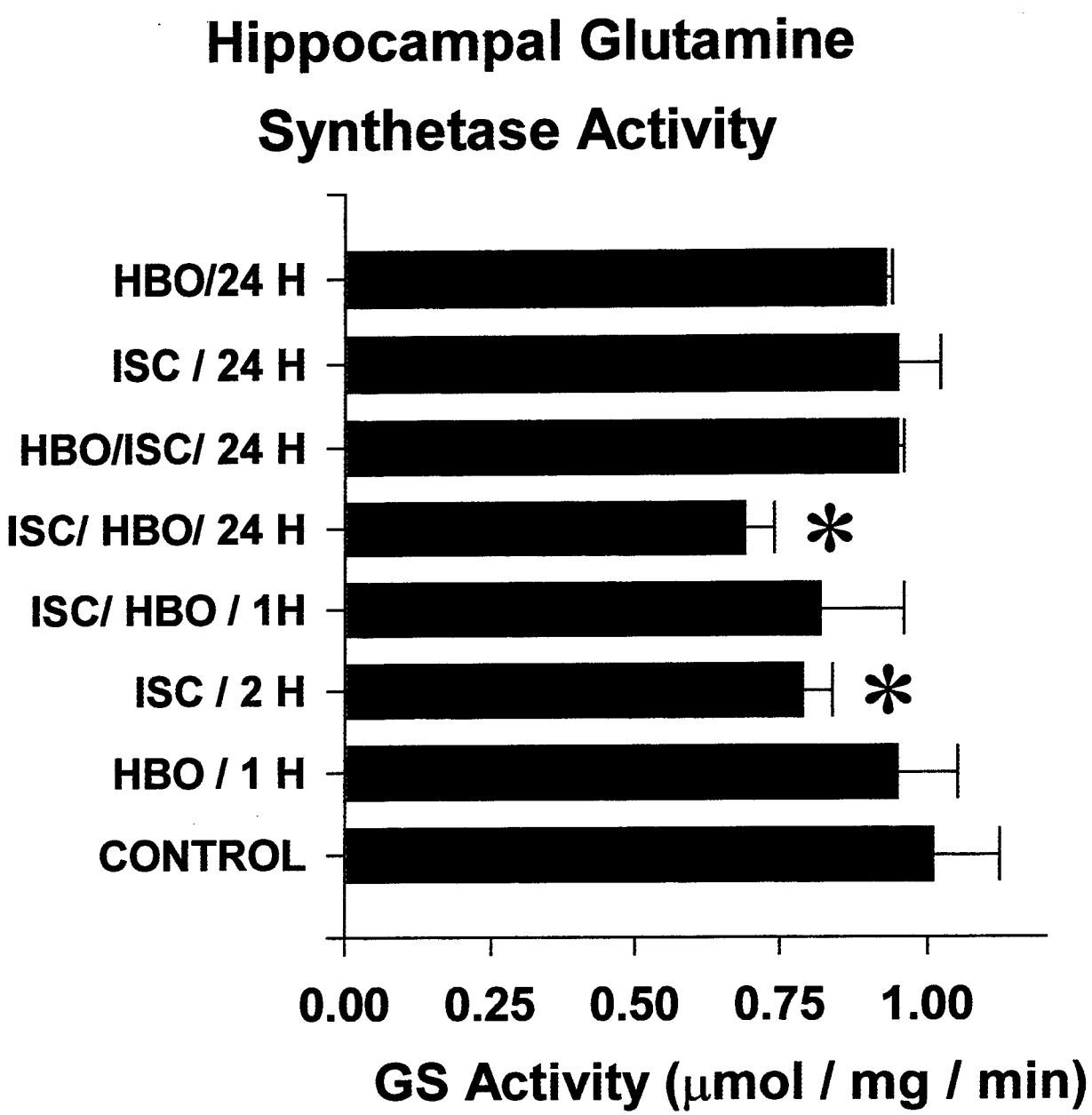

Hippocampal Glutamine Synthetase Activity
HBO/24 H
ISC / 24 H
HBO/ISC/ 24 H
ISC/ HBO/ 24 H
ISC/ HBO / 1H
ISC / 2 H
HBO / 1 H
CONTROL
*
*
0.00
0.25
0.50
0.75
1.00
GS Activity (μmol / mg / min)

Figure 1.

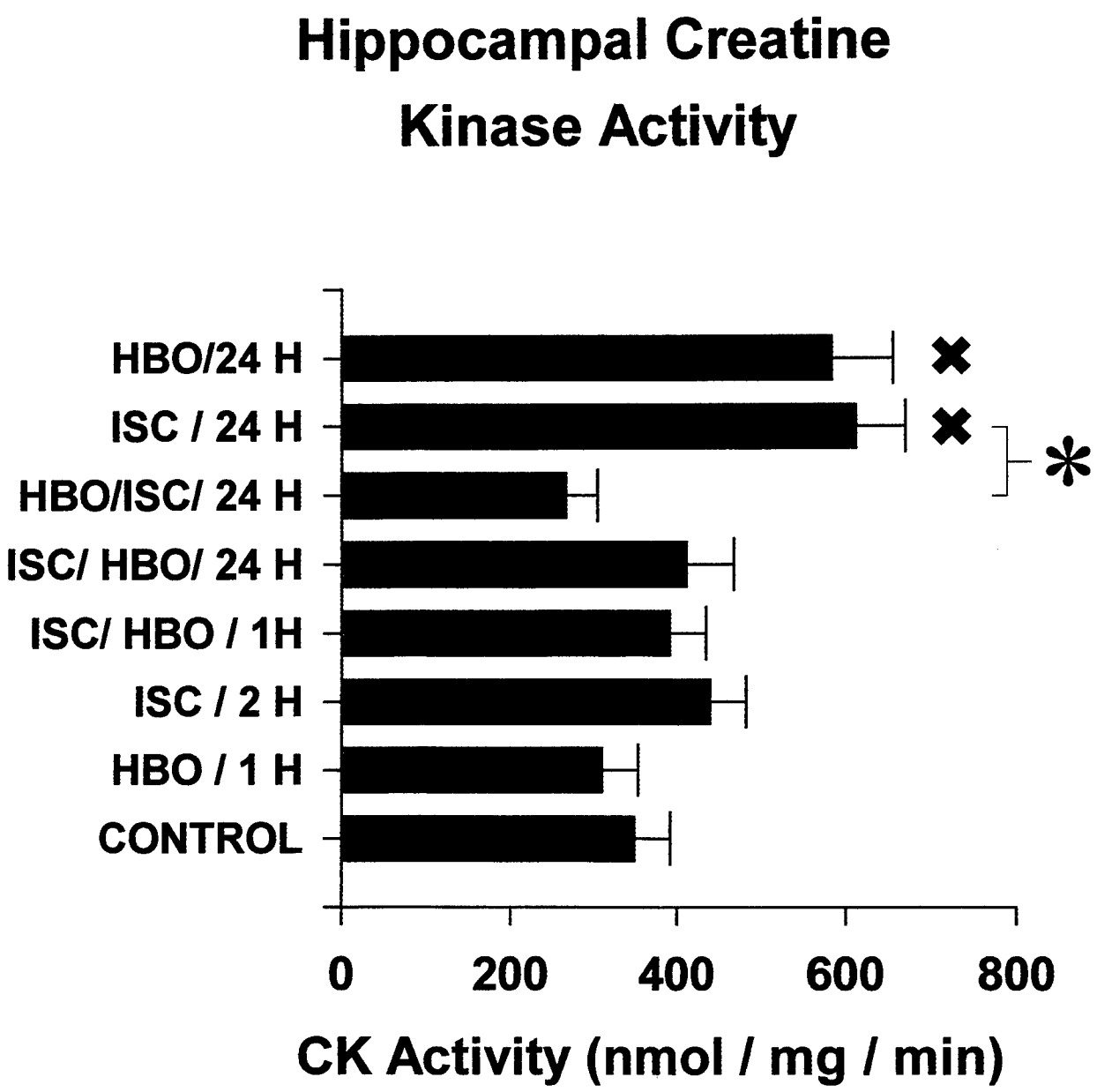

Hippocampal Creatine Kinase Activity
HBO/24 H
ISC / 24 H
HBO/ISC/ 24 H
ISC/ HBO/ 24 H
ISC/ HBO / 1H
ISC / 2 H
HBO / 1 H
CONTROL
0
200
400
600
800
CK Activity (nmol / mg / min)

Figure 2.

FAULT ANALYSIS AND EXCITATION REQUIREMENTS FOR SWITCHED RELUCTANCE STARTER-GENERATORS

Iqbal Husain
Assistant Professor
Department of Electrical Engineering

The University of Akron
356 Auburn Science and Engineering Center
Akron, OH 44325-3904

Final Report for:
Summer Faculty Research Program
Aero Propulsion and Power Laboratory
Wright Patterson Air Force Base
Dayton, Ohio

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

and
Wright Patterson Laboratory (WL-POOC)

August 1996

FAULT ANALYSIS AND EXCITATION REQUIREMENTS FOR SWITCHED RELUCTANCE STARTER-GENERATORS

Iqbal Husain
Assistant Professor
Department of Electrical Engineering
The University of Akron

Abstract

The excitation requirements with and without faults for switched reluctance starter/generators were studied in this research. The size of the excitation source required during the initial voltage build-up stage of generating mode and under fault conditions is a critical issue in aircraft starter/generator engines. In this research, simulation models were modified and further developed from existing models to simulate various switched reluctance starter/generator configurations. The models were used to investigate the fault modes and determine the excitation requirements of four different machine configurations.

The various fault modes in the External Integral Starter-Generator (EIS/G) were first analyzed, and the excitation requirements under these faulted conditions were simulated and studied. The load fault case presented the worst case scenario as far as the excitation requirements are concerned.

Three-phase machines with three-repetitions, and four-phase machines with two- and three-repetitions were designed using the same specifications as that of the EIS/G (three-phase, two-repetitions). Simulation showed that the excitation requirements for the four-phase machines were lower for initial voltage build-up, but more under similar load fault conditions compared to that of the three-phase designs.

The system level simulations including the faults were performed in Matlab-Simulink platform. Mathcad programs were used to design the machines.

FAULT ANALYSIS AND EXCITATION REQUIREMENTS FOR SWITCHED RELUCTANCE STARTER-GENERATORS

Iqbal Husain

I. Introduction

The switched reluctance machine is a prime candidate for the more electric aircraft (MEA) applications primarily because of its fault tolerance capabilities. The fault tolerance capabilities have thus far been discussed in an abstract sense for starter/generators and very few results have been published in the literature. Furthermore, the published literature [1,2] discusses the fault tolerance and detection of SRMs operated as a motor and not as a generator.

If the generator is supplying a load heavier than its regular load, such as under faulted conditions, the bus voltage would collapse to zero. In order to avoid such a situation, a separate excitation source has been suggested [3] to supply the excitation during voltage build-up and under fault condition. The question that arises next is how big an excitation source would be required for the starter-generator system. Simulations and analysis have been done during the past three months under the summer faculty research program to determine the excitation requirements under normal and various faulted conditions.

This report presents the results of simulations and analysis for excitation requirements of switched reluctance integral starter/generator system. The external integral starter generator (EIS/G) has been used initially for the simulations. Later, four-phase machines with 2 and 3 repetitions, and a 3-phase machine with 3-repetitions, have been designed. Repetition refers to the independent output channels that the generator is delivering power to. The simulation models for these systems were developed and then used for analysis. The excitation requirements for these four systems were compared. The results and analysis are reported after presenting a review on switched reluctance generators in section 2.

II. Switched Reluctance Generator

A switched reluctance machine (SRM) is operated in the generating mode by positioning the phase current pulses in the falling inductance region immediately after the poles have passed alignment. However, in this generator mode, the machine obtains its excitation from the same bus that it generates into. The phase is turned on before pole alignment, and energy is supplied by the DC bus to excite the winding. Once pole overlap begins and current is established in the winding, the phase excitation is removed. The work done by the mechanical system to pull the rotor poles away from the stator poles is returned to the DC bus, together with most of the excitation energy. The control key is to position the phase pulses precisely in order to maximize the efficiency and reduce stresses on the power electronic converter. The phase current pulses during generating are, in fact, the mirror images of the phase current pulses during motoring around the aligned position [3]. Typical phase current waveforms in the generating mode will appear later in the simulation section.

With SRMs, multiple channels can be built by repeating the basic stator and rotor pole arrangements. For example, a 2-channel, 3-phase machine can be built using a 12/8 pole configuration, which is a 2-repetition of the basic 6/4, 3-phase machine. The machine is wound in such a way that with the special independent nature of SRM, it behaves like two separate 3-phase machines. It is the independent nature of the SRMs that allows the single machine to supply power to two (or more) separate, independent, and presumably differently loaded output buses.

The switched reluctance generator is typically operated in the constant power region of the torque-speed characteristics of the machine. In this region, the machine operation is limited to the single pulse mode, where the phase switches turn on (at θ_{on}) and turn off (at θ_{off}) only once per cycle. The higher machine speeds in this mode also cause the back-emf (BEMF) to be significant and is often higher than the DC bus voltage. The SRM can generate at lower speeds, but it would then be limited to a constant torque capability rather than the constant power capability required of a generator.

A switched reluctance generator including the power electronic converter, the EMI filter and a "diode OR" excitation source is shown in Fig. 1. The power electronics is the two switch, two diode per phase classic converter topology. A resistive load is assumed at the output of the generator.

The terminal phase voltage of an SRM is given by

$$V_{dc} = iR_{ph} + \frac{d\lambda}{dt} = iR_{ph} + \frac{\partial\lambda}{\partial i}\frac{di}{dt} + \frac{\partial\lambda}{\partial\theta}\omega_m \quad \Rightarrow \quad \frac{di}{dt} = \frac{V_{dc} - iR_{ph} - \frac{\partial\lambda}{\partial\theta}\omega_m}{\frac{\partial\lambda}{\partial i}}.$$

R_{ph} is the phase winding resistance, while $\frac{\partial\lambda}{\partial\theta}\omega_m$ is the back-EMF (BEMF) term.

The average generated current for a given set of θ_{on} and θ_{off} angles is proportional to the bus capacitor voltage and can be written as

$$i_{av} = k(\theta_{on},\theta_{off})v_c = \frac{v_c}{R_k},$$

where $k(\theta_{on},\theta_{off}) = \frac{1}{R_k}$ is a proportionality constant which depends on θ_{on}, θ_{off}, machine speed ω_m and the bus capacitor voltage v_c. The bus capacitor voltage is governed by the following relation:

$$\frac{dv_c}{dt} = \left(\frac{1}{R_k C} - \frac{1}{R_L C}\right)$$

where R_L is the load resistance.

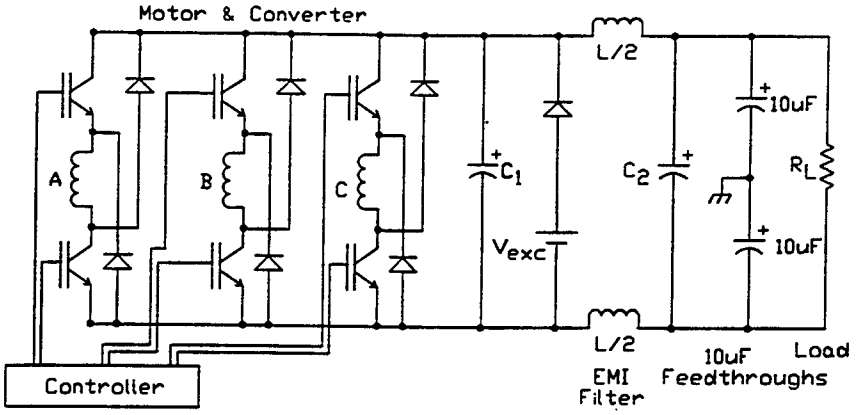


Fig. 1 Switched Reluctance Generator System.

For fixed θ_{on} and θ_{off} angles, if R_k is less than R_L, which would be the case for lighter loads, the bus capacitor voltage would gradually increase and cause instability. On the contrary, if R_k is greater than R_L, which would be the case for heavier loads, the bus voltage would collapse to zero. Therefore, θ_{on} and θ_{off} angles are continuously adjusted in the feedback control loop, regulating the generator output voltage v_c, for regular and lighter loads so that R_k equals R_L on the average. However, during a load fault, when R_L is too small, the angles θ_{on} and θ_{off} would saturate to the maximum possible difference among them due to the control action of the feedback loop. θ_{on} and θ_{off} would then become fixed and the generator output voltage v_c would collapse to zero. The generator would not be able to recover unless it is supported by a separate excitation source.

The "diode OR" excitation source V_{exc}, shown in Fig. 1, supplies the initial excitation current and the excitation current during and after a load fault. The diode blocks the generated current from going into the excitation source. Once the bus voltage builds up above the source excitation voltage, the diode isolates the initial excitation and fault source from the system and the machine draws its excitation current from the DC bus into which it generates. If a fault occurs and the bus voltage collapses below the excitation source voltage , the source supplies all the power during the fault and also the excitation energy required for output voltage build-up after the fault is cleared.

III. Modeling of SR Starter/Generator

3.1 EIS/G System Description

The EIS/G has a 4-pole magnetic configuration with 12 stator and 8 rotor poles. The system has two independent channels generating into two isolated 270V DC buses. Two separate power inverters and two VME rack-based controllers are used to drive the 4-pole, 12/8, 3-phase system. Diagonally opposite poles are tied together to form one phase leg of one three phase group in order to avoid unbalanced force conditions [3].

The EIS/G system which is currently at Sundstrand does not have a separate excitation source. However, EIS/G system can be modified so that the power and excitation buses are separated and the power bus is diode or-ed to the excitation bus.

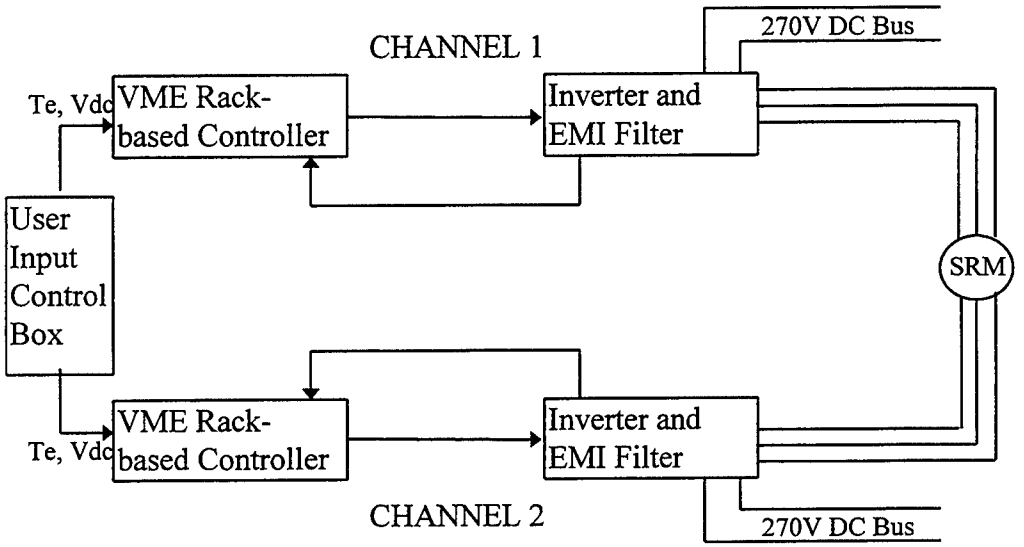


Fig. 2 Block diagram of the 250kW EIS/G system.

The power inverter buses are passed through an EMI filter to meet the EMI requirements. The EMI filter is designed to not only eliminate the differential ripple from the switched reluctance generator system's output, but also the common mode currents flowing in the chassis produced by the switching of the semiconductor devices.

A resistive load of .442Ω has been used to simulate the load for two channel SR generators, and a load of .663Ω for three channel SR generators. These resistive loads would result in a total output power of 330kW, which is the maximum overload power for the EIS/G.

3.2 Simulation Model

A Matlab-Simulink model of the three-phase, 12/8 EIS/G system developed by Dr. Arthur Radun of the University of Kentucky was the starting point of all the simulation and analysis reported here. The original model is capable of simulating the system with a step type load fault. The model has been modified to simulate various other step and ramp type load fault conditions as well as inverter and internal machine fault modes. Blocks were also added to determine the excitation requirements under various faulted conditions. A model was developed to simulate 4-phase SR generators which is shown in Fig. 3. The machine model itself was also re-written to simulate four-phase, 4-pole and 6-pole, and three-phase 6-pole systems.

The model is built in a modular fashion so that changes can be incorporated easily in the different blocks of the system. The machine model is embedded in the Matlab function block 'didt4ph'. The machine model is written in 'C language' and compiled using Matlab 'cmex' tools in order to reduce simulation time. All the other blocks are developed using Simulink tools. The various blocks of the model are described in the following.

Machine model: The machine model contains the detailed geometry and characteristics of the SR machine. The input to the block are the previous step phase currents, phase voltages and properly shifted rotor angles for the four phases. The outputs of the block are the phase current derivatives which are then passed through an integrator to obtain the phase currents.

Phase Currents: The phase current data is stored in this block to generate simulation plots.

Make Ilink: This block is a summing junction where all the phase currents are added to generate the dc link current. The identities of excitation current and generating current are maintained and stored separately as 'iexite' and 'ipower' respectively.

EMI Filter: The EMI block models filter components C1, C2 and L. The total resistance of the wires and the phase windings are also included in this block. The block outputs the excitation voltage and generated voltage separately as 'vexite' and 'vpower' as well as the load current. The power required by the 'diode-or' excitation source is also calculated and stored within this block.

Diode or: This block simulates a diode.

Load: The block has the model for the load. The block simulates a load fault after a certain period of generation and then clears the fault after 25ms.

Controller: The feedback control algorithms are modeled within this block. The inputs to the controller are the output voltage and the machine speed. The outputs are the command current 'icom' and the turn-on and turn-off angles 'qon' and 'qoff' respectively.

Current Regulator: The current regulator receives the control outputs, the current feedbacks and the shifted rotor angles, and generates the gate turn-on and turn-off angles for each individual phase of the machine.

Make Vph: The block generates the phase voltages based on the existing power bus and/or excitation voltage available. These phase voltages are then fed to the SR machine block 'didt4ph'.

Phase Angles: The block converts the speed in rpm to rads/sec and also calculates the appropriately shifted rotor angles for the four phases so that a model of only one phase can be used for simulation.

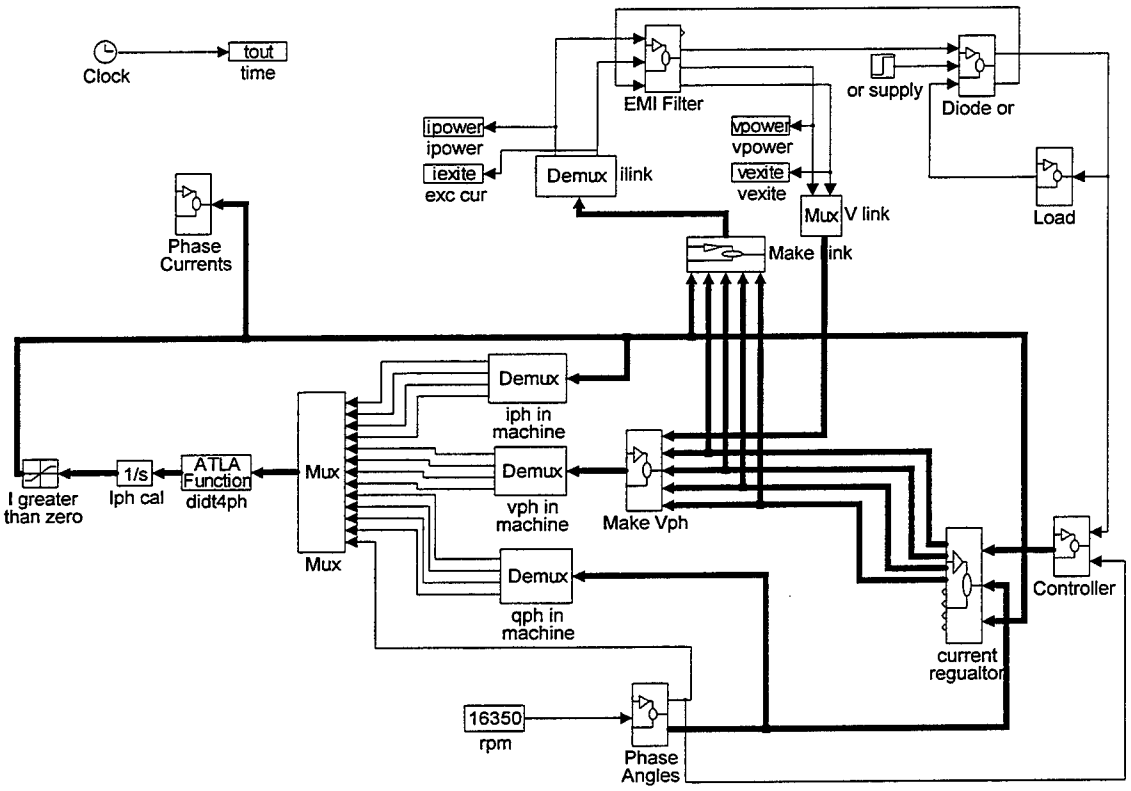


Fig. 3 Matlab-Simulink model to simulate 4-phase SR generators including various fault modes.

IV. Fault Analysis

The SR generator has been simulated for faults both internal and external to the machine. The excitation requirements under these faulted conditions were obtained from simulation.

4.1 Machine and Converter Faults

The various fault modes considered and simulated are:
i) Armature Phase Open: One phase of the machine was simulated to be open.
ii) Armature Phase Short: One phase winding were simulated to have shorted after 12ms.
iii) Switch Failure: Phase switches of one phase was open-circuited and short-circuited to simulate switch failure modes and determine the excitation requirements.
iv) Load Fault: Various levels of load fault were simulated and the excitation requirements under these fault conditions were determined. Load faults at start-up was also simulated.

The excitation source power requirements were determined for the fault conditions (i)-(iv) using modifications of the basic Simulink model. Load faults were superimposed on top of the phase or switch faults to find the excitation requirements under these severe fault modes. The diode failure modes were not simulated, since the failure leads to short circuit conditions or catastrophic failures. The excitation source power and current requirements under these various fault conditions are summarized below.

Table I
Excitation Requirements Under Various Failure Modes

	Phase Open		Phase Short		Switch Open		Switch Short		Load Fault
	Before Ld. Flt.	During Ld. Flt.	Before Ld. Flt.	During Ld. Flt.	Before Ld. Flt.	During Ld. Flt.	Before Ld. Flt.	During Ld. Flt.	
Source Power (kW)	0	50	0	70	0	50	500	500	50
Source Current (Amps.)	0	1000	0	1400	0	1000	10000	10000	1000

Simulation showed that the SR generator is able to maintain its output voltage even if there is phase winding open or short fault, or a phase switch opens, without the load fault. When a load fault is superimposed on a phase fault, the excitation power is drawn from the diode-or excitation source as expected.

The output voltages for phase winding short and converter switch short faults are shown in Fig. 4. The step input of Vout=270V is commanded at 5ms., and the machine and converter faults occur at 12ms. In the case of phase winding short fault, the voltage ripple increases significantly although the voltage is maintained well above 50V. Energy is not drawn from the excitation source unless there is a load fault, which occurs at 23ms. On the other hand, the phase-switch short results in abnormally high value of source currents which would lead to catastrophic failures.

4.2 Load Faults

The load faults are the most critical faults in the case of SR generators and need to be investigated further. The EIS/G system will be investigated first under various load fault conditions, and later load faults for three-phase, three- repetition, and four-phase, two- and three-repetition machines will be simulated.

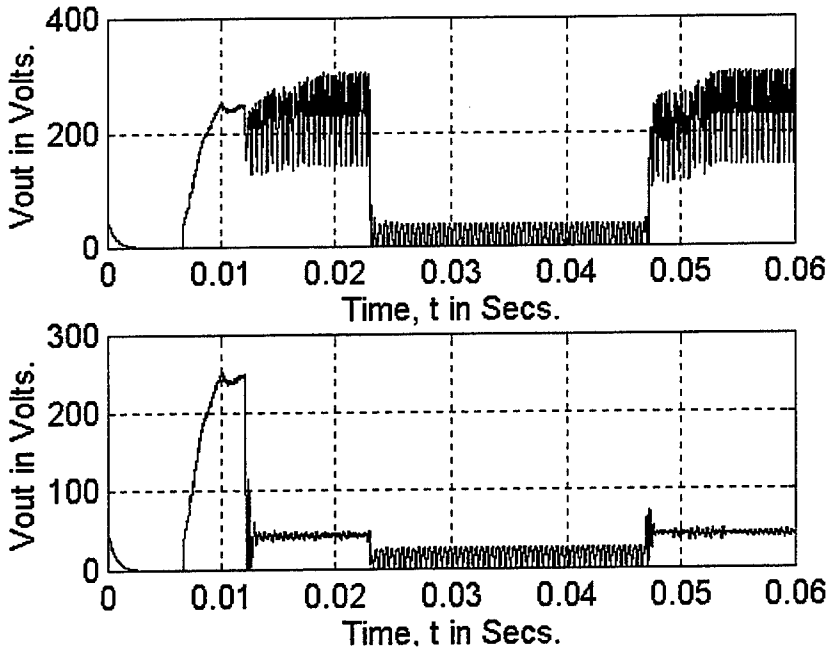


Fig. 4 Output voltage waveforms for phase winding short fault (top trace) and converter switch short fault (bottom trace).

EIS/G Load Fault Simulation

Plots of output voltage, phase voltage and currents, excitation and generator currents, load current, source current and excitation source power for the EIS/G system are shown in Figs. 5 through 9. A step command of Vout=270V is set at 5ms and the output voltage approaches the command voltage as seen in Fig. 5. A load fault is simulated at 23ms when the output voltage collapses to about 35V. The fault is cleared at 47ms after which the generator is seen to recover because of the 'diode-or' excitation source. Phase voltage and phase currents are shown in Fig. 6. The generated current 'ipower', and the excitation current 'iexite' are shown in Fig. 7. The excitation current is negative since this is drawn by the machine and is coming from either the source or dc power bus. The load current is about 1500A during fault, although it shows a transient current spike of 14kA (Fig. 8), which could have detrimental effects on the system. The phase voltages and currents, however, don't show any transient overshoots. The excitation current drawn from diode-or source is 400A initially and 1000A during fault, while the power drawn from the source is 20kW for the initial output voltage build-up and 50kW during faults which are about 12% and 30% respectively of the generated power.

Although the simulation predicts results with high magnitudes of currents, in real time the devices will fail, if the voltage and current stresses exceed the rating during the fault condition and in that case the generator will not recover from fault. Furthermore, after a load fault, over-current protection circuits should be activated automatically and the power drainage from the excitation source will not continue as shown in the simulations. After the fault is cleared, the generator will again take the required power from the excitation source and build-up the output voltage.

The fault condition existing at the beginning of generating mode is simulated and shown in Fig. 10. The generator is still able to pick up after the fault is cleared at 23ms. The power requirement from the excitation source during fault is around 50kW.

A linear increase in fault current has been simulated and are shown in Figs. 11 through 14. Figs. 11 and 12 show the excitation source power, load current and phase current when the load resistance drops gradually from 0.1Ω to 0.01Ω in 50ms and then recovers from the fault. The excitation power requirement surprisingly has a minimum of about 35kW around .05Ω of load resistance as seen in Fig. 11. The load current waveform in Fig. 12 shows a transient overshoot of about 2600A, although transients are absent in

the phase current waveform. A similar linear increase in load are shown in Figs. 13 and 14, where the load resistance drops from 1mΩ to 0.1mΩ. The load current shows a transient overshoot of 250kA, although the steady state current is about 2700Amps. The maximum excitation power requirement is about 75kW at R_L= 0.1mΩ. However, such a low load resistance is not realistic, since the sum of the contact resistances would be greater than this value. Also, the phase winding and power device on-state resistances would start dominating at this level to limit the load current.
The excitation source power requirement is presented in tabular form below in Table II.

Table II
Excitation Currents and Power Requirements under Load Faults

Load Resistance, RL	.1Ω	.05Ω	.02Ω	.01Ω	.001Ω
Load Current (steady state peak)	870A	860A	1490A	1950A	2750A
Load Current (average)	805A	830A	1450A	1860A	2650A
Excitation Source Current (peak)	990A	740A	1000A	1050A	1120A
Excitation Source Power (peak)	49.5kW	37kW	50kW	52.5kW	56kW

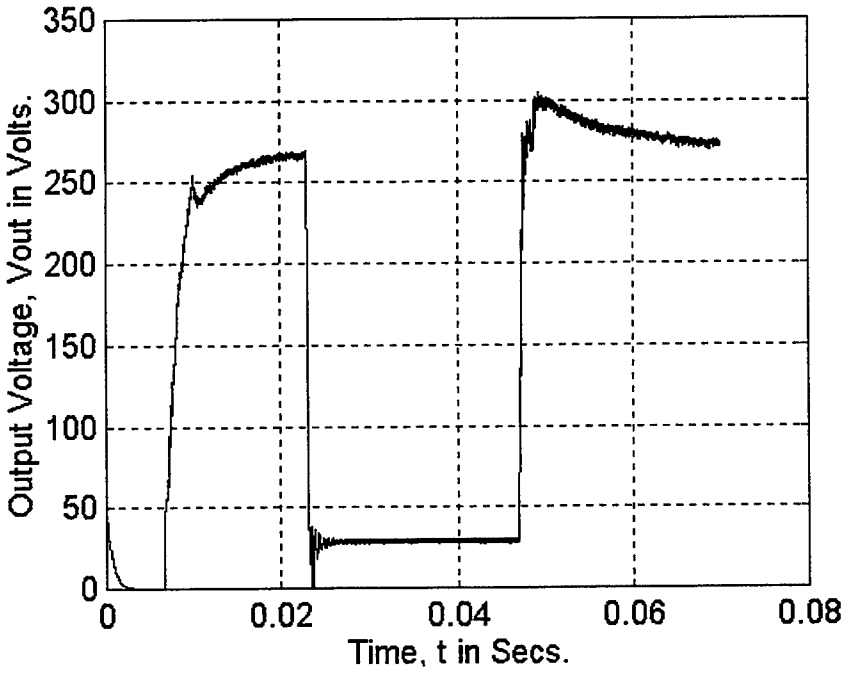


Fig. 5 Output voltage of one channel of the EIS/G.

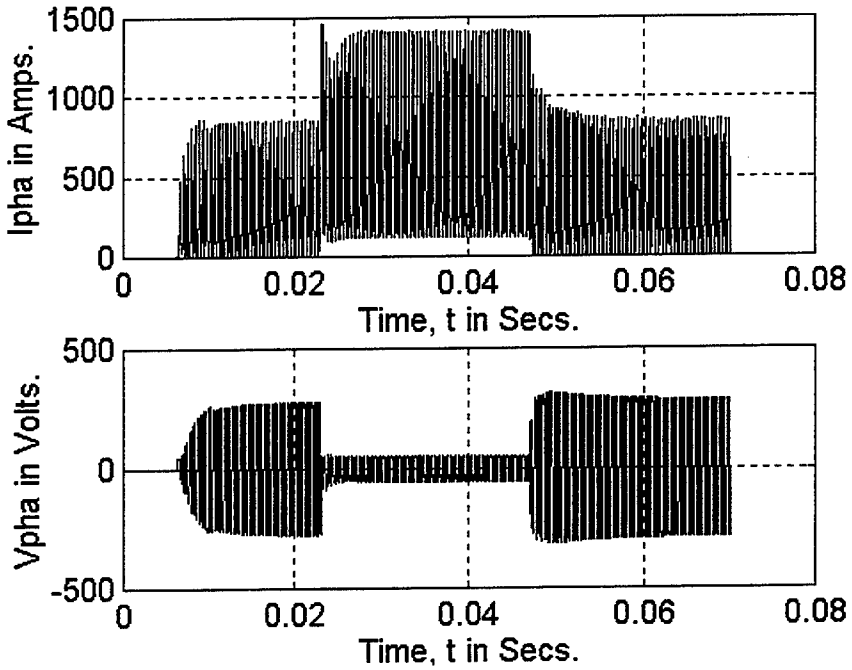


Fig. 6 Phase current and phase voltage.

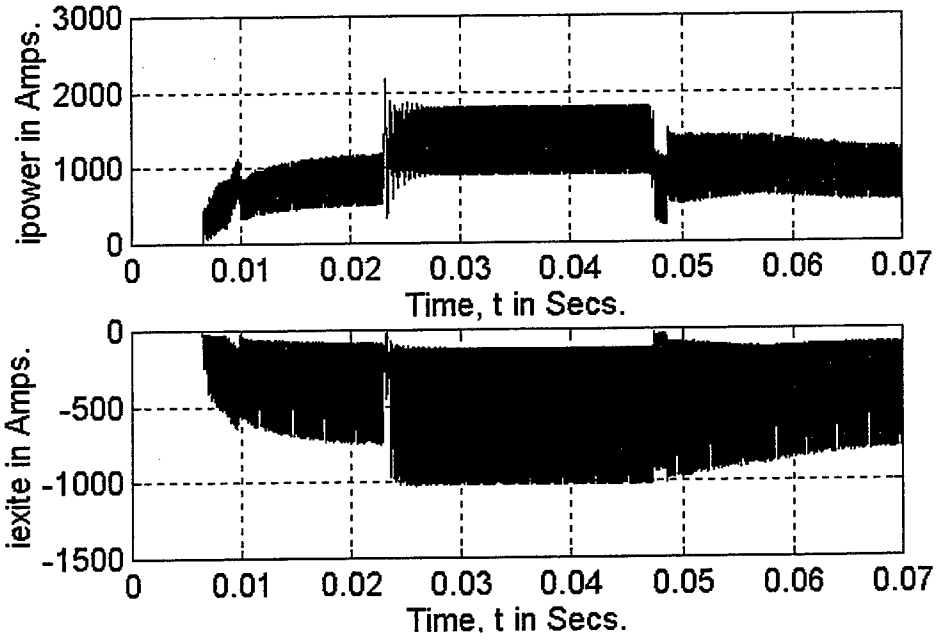


Fig. 7 Generated current 'ipower' and excitation current 'iexite'.

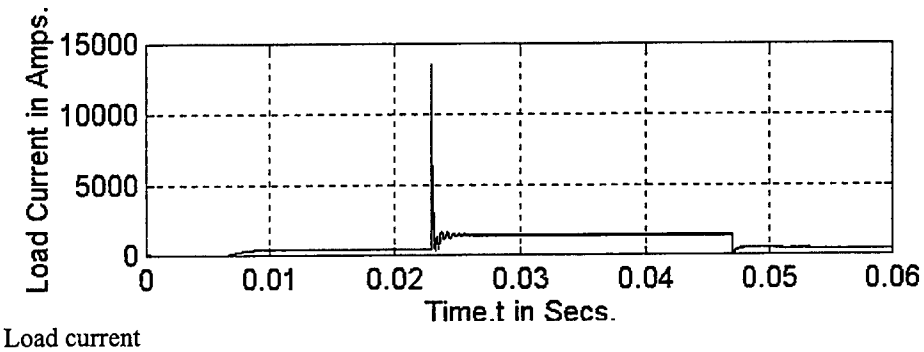


Fig. 8 Load current

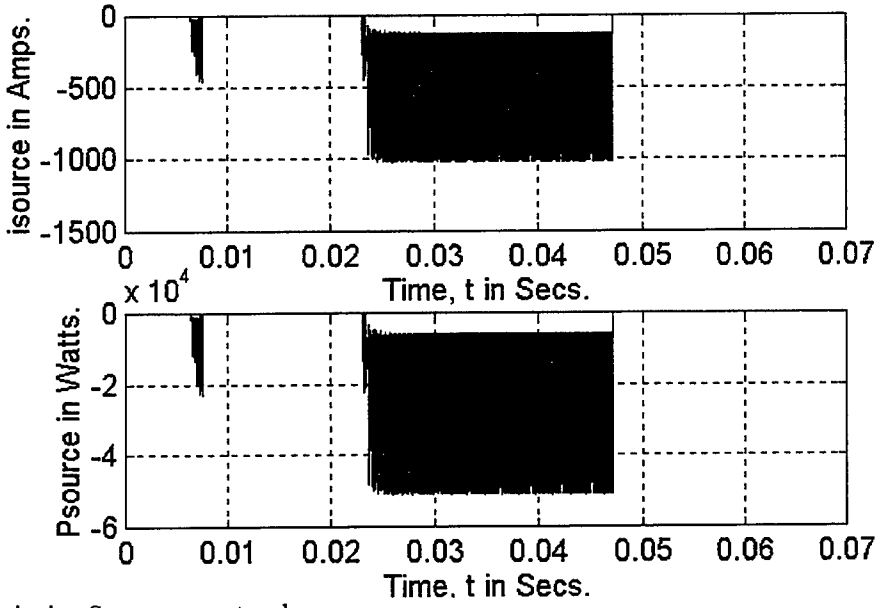


Fig. 9 Excitation Source current and power.

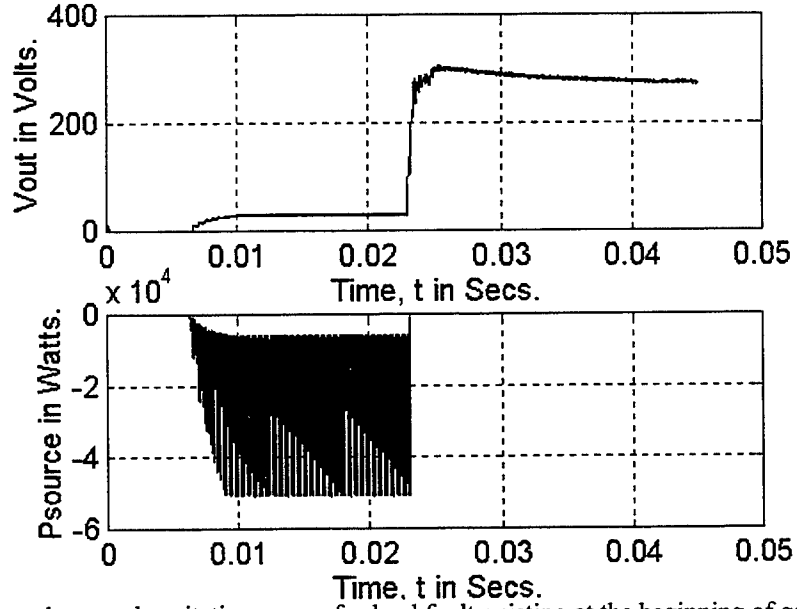


Fig. 10 Output voltage and excitation power for load fault existing at the beginning of generating mode.

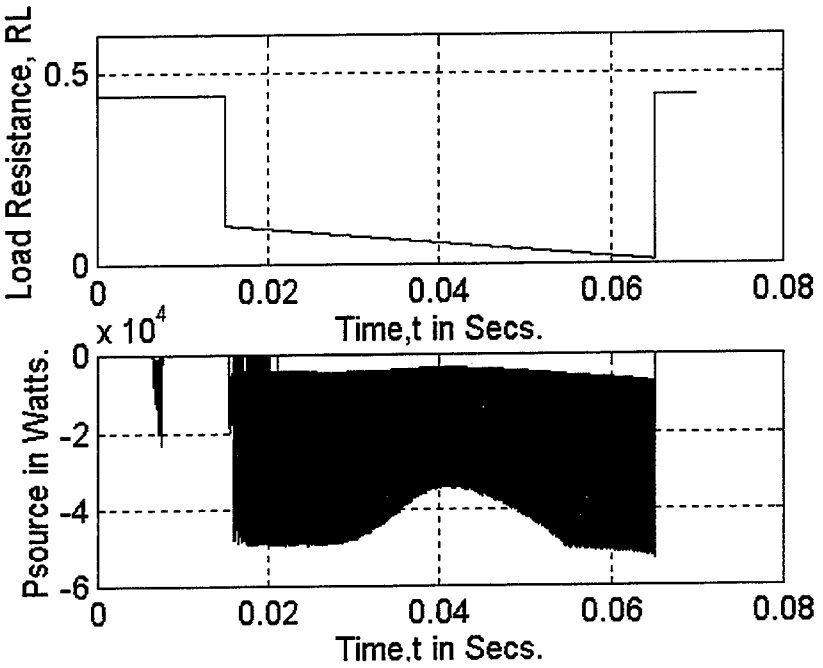


Fig. 11 Change in load resistance from .1Ω to .01Ω during fault and excitation source power.

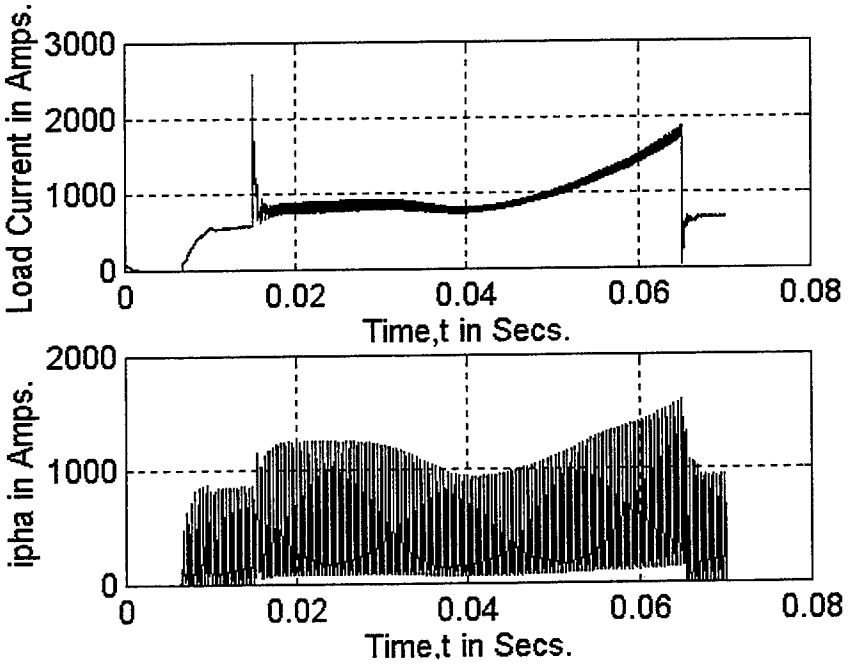


Fig. 12 Load current and phase current.

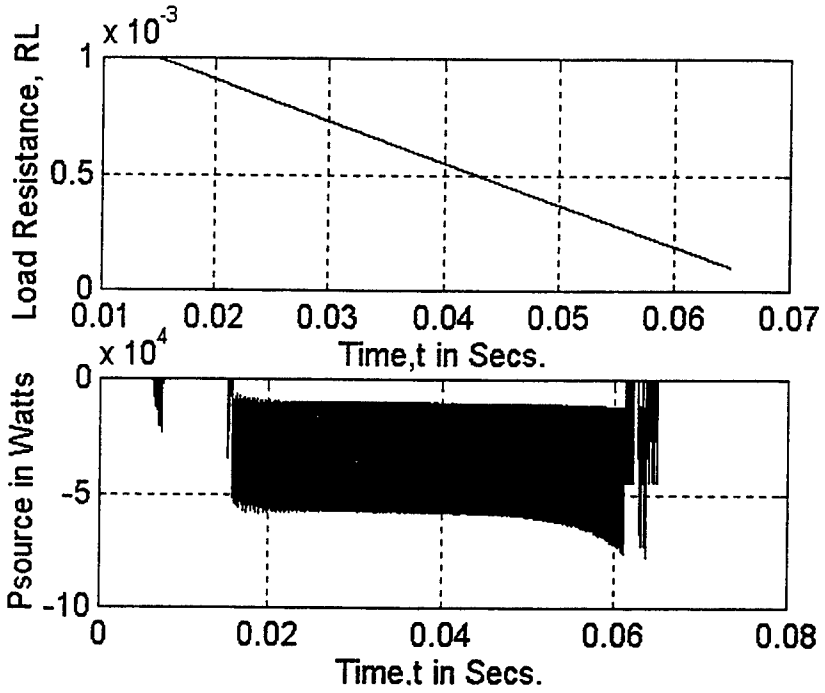


Fig. 13 Change in load resistance from 1mΩ to .1mΩ and the source excitation power.

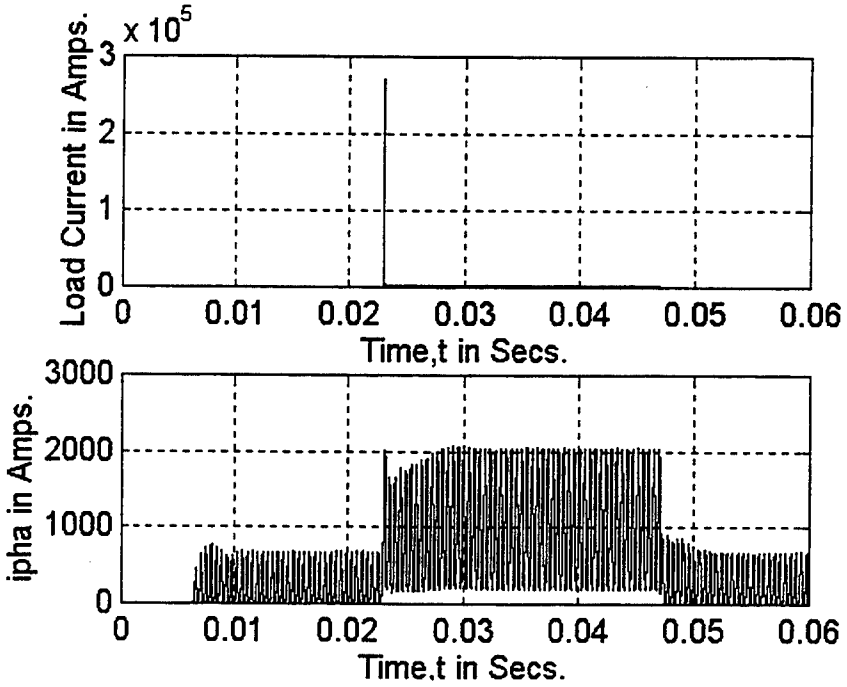


Fig. 14 Load current (steady state current during fault is about 2700A) and phase current.

V. Alternate Machine Configurations

5.1 Machine Designs

Mathcad programs were used here to design three-phase machines with 3 repetitions, and four-phase machines with 2 and 3 repetitions. The design programs were originally developed by Prof. Arthur Radun of the University of Kentucky. The analytical model used in simulations and the design equations used in this section appear in [6].

The design constraints and limits, as well as the magnetic materials for the alternate configurations, were kept the same as that used for 3-phase, 2 repetition EIS/G system. The following design guidelines were used:

Design Constraints
Rated speed= 22,224 rpm, Corner speed = 16,352 rpm,
Maximum Power =330kW, Rated power =250kW,
DC Link Voltage = 270 Vdc.

Design Limits
Maximum rotor hoop stress=40,400 psi
Rms current density =1254 A/cm^2
Maximum average rotor torque density=70,000 N-m/m^3
Power Factor=0.45
Efficiency=0.95

The Mathcad design program 'predesi.mcd' was run to design the three machine configurations. The machine geometry obtained are given in Table III. The geometry of EIS/G is also given for comparison.

Table III
SR Generator Geometry

Parameters	4-phase, 2-Rep. (16/12)	4-phase, 3-Rep. (24/18)	3-phase, 3-Rep. (18/12)	3-phase, 2-Rep. (12/8) - **EIS/G**
Airgap, g in.	.02	.02	.02	.02
Stack length, in.	2.133	5.426	13.169	7.17
Stacking factor	0.9	0.9	0.9	0.9
Rg,in.	3.616	3.16	2.929	3.125
Rry,in.	2.423	2.465	2.528	2.575
Rsy, in.	5.121	4.036	3.435	3.84
Np	15	9	4	7
Nrep	2	3	3	2
θpd	22.5	22.5	31	31
nser	1	1	1	1
npar	2	2	2	2
θu	30	30	45	45
θa	60	60	90	90
θsep	45	45	60	60

Rg = radius to the airgap at the rotor
Rry= radius to the outside rotor yoke
Rsy=radius to the inside stator yoke

Np = number of turns per pole
Nrep = number of machine repetitions
θpd= rotor pole width times Nrep in degrees
θu= unaligned rotor position times Nrep in degrees
θa= aligned rotor position times Nrep in degrees
θsep= stator pole separation times Nrep in degrees

The geometry is shown in Fig. 15. One repetition of a 6/4, three-phase machine is shown for simplicity in the figure.

The analytical model of SRMs developed in [6] were used to predict the magnetization curves and the phase current for one cycle of the four machines. These are plotted in Figs. 16 through 19. It must be mentioned here that the controller for the four-phase generators were not optimized, and the same control constants as those used for the 3-phase machines were used for the four-phase machine. However, this approach was sufficient to predict excitation requirements of the four-phase machines.

The EMI filter components used for three-phase and four-phase machines are given in Table IV:

Table IV
EMI Filter Components

	4-phase SRG	3-Phase SRG
C1	506 μF	842μF
C2	375 μF	624μF
L	2.4 μH	4μH

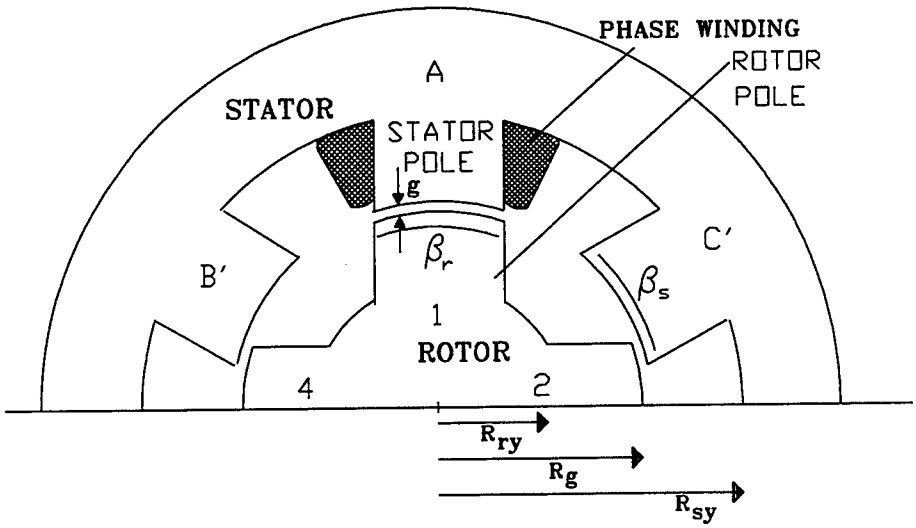


Fig. 15 Half cross-section of a switched reluctance machine.

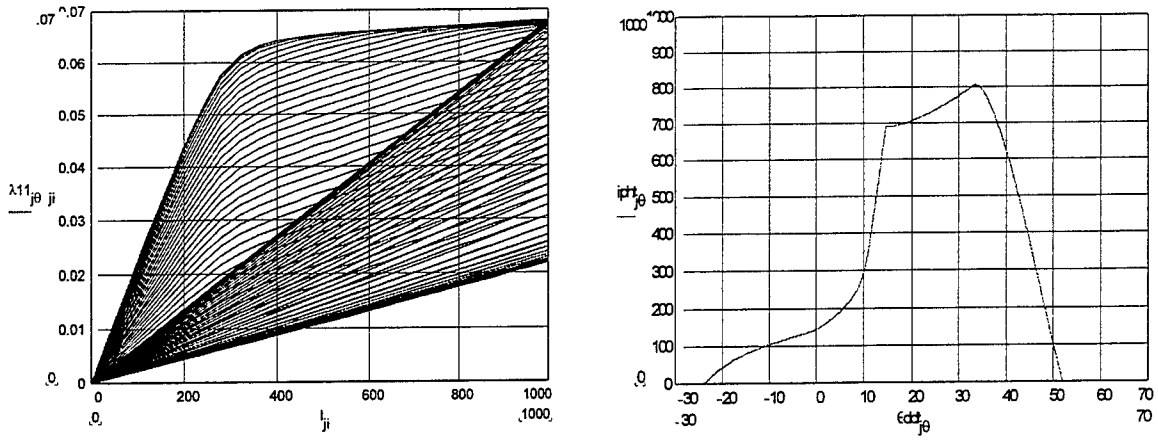


Fig. 16 Magnetization curves and phase current of the 3-phase 2 -repetition EIS/G machine.

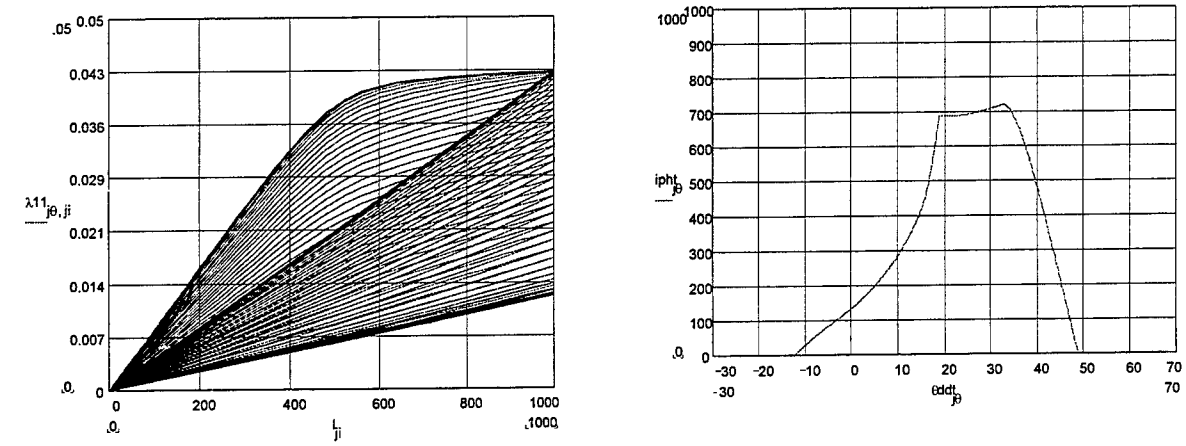
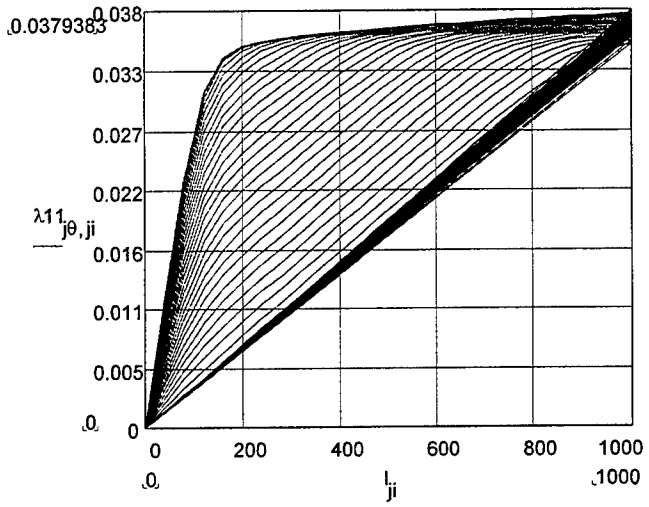


Fig. 17 Magnetization curves and phase current of the 3-phase 3 -repetition switched reluctance generator.



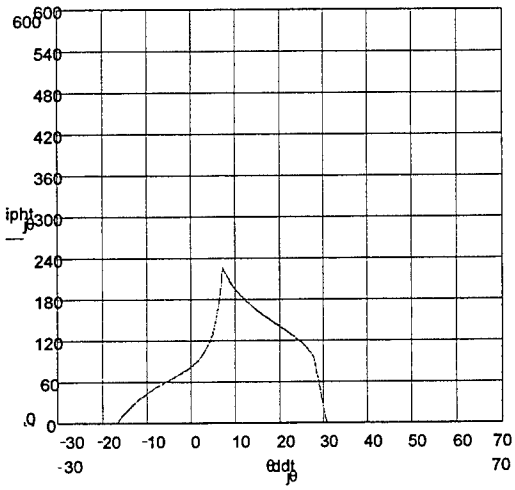
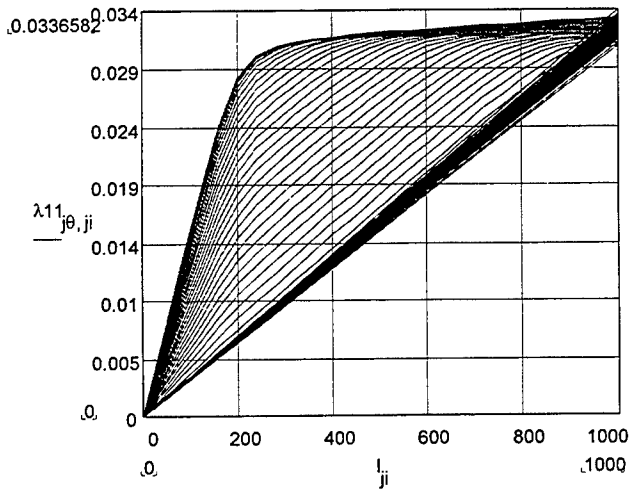


Fig. 18 Magnetization curves and phase current of the 4-phase 2 -repetition switched reluctance generator.



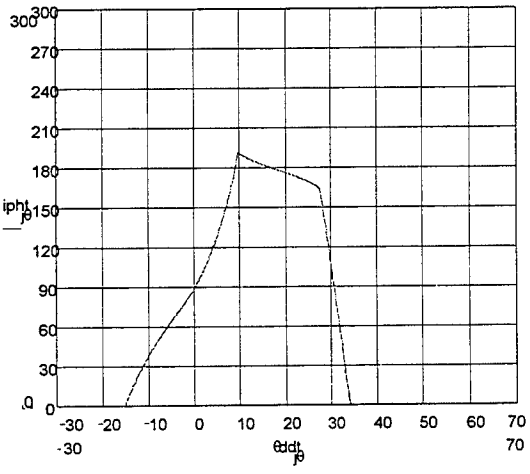


Fig. 19 Magnetization curves and phase current of the 4-phase 3 -repetition switched reluctance generator.

5.2 Fault Simulations

The output voltage and excitation source power without and with load faults for the three-phase and four-phase machines are shown in Figs. 20 through 22. The output voltages for all three machines reach the set point of 270V, which is the step input command at 5ms. The response for the four-phase machines is also surprisingly good considering that the control constants were still set for three-phase machines. The fault occurred at 23ms and cleared at 47ms and all three machines were able to recover from the fault with the energy coming from the excitation source during fault and at start-up. The excitation requirements at start-up and during fault are given in Table V.

Table V
Excitation Requirements for Alternate Configurations

Machine Configuration	3-Phase, 2-Repetition		3-Phase, 3-Repetition		4-Phase, 2-Repetition		4-Phase, 3-Repetition	
	Total Power	% of Channel Output Power	Total Power	% of Channel Output Power	Total Power	% of Channel Output Power	Total Power	% of Channel Output Power
Source Power for voltage build-up	20kW	12%	20kW	18%	12kW	7%	10kW	9%
Source Power during fault	50kW	30%	52KW	47%	65kW	39%	65kW	59%

The 4-phase designs require less excitation power from the source for initial voltage build-up, but more during load faults, as can be seen from Table V. The fault load was the same for 3-phase and 4-phase, as well as for 2-repetition and 3-repetition designs. The simulations were performed on a per channel basis and the percent values given in Table V are also on a per channel basis. The percentage excitation requirements for 3-repetition designs are more, although the power requirements are close. The airgap for the 3-repetition designs should be proportionately reduced to minimize the excitation requirements. However, airgap was kept at 20 mils for all the designs considering manufacturing limitations.

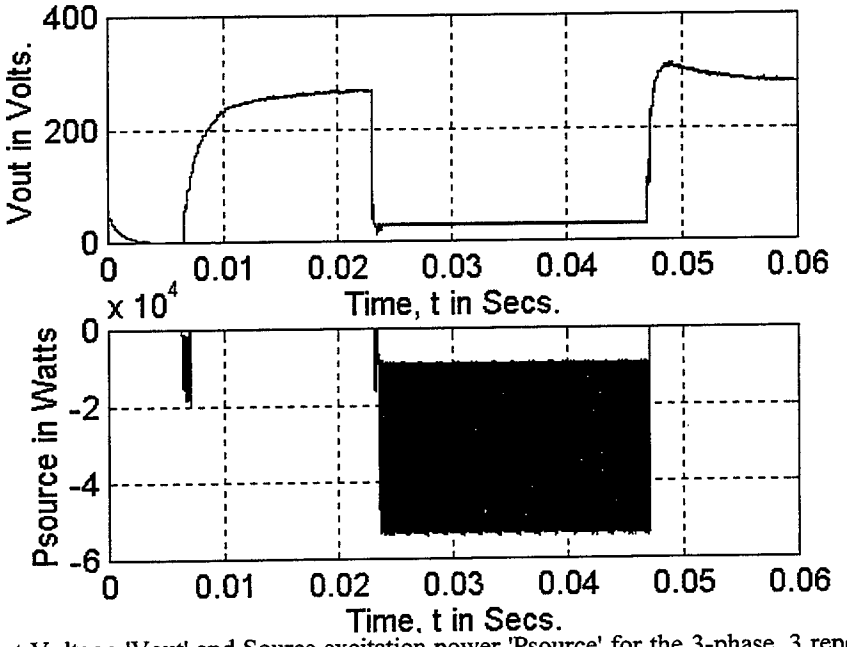


Fig. 20 Output Voltage 'Vout' and Source excitation power 'Psource' for the 3-phase, 3 repetition switched reluctance generator.

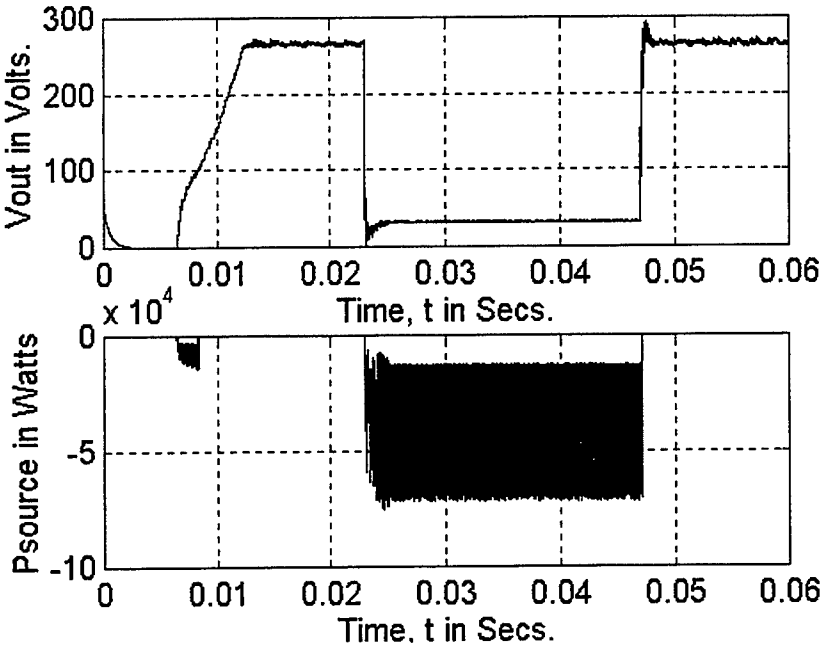


Fig. 21 Output Voltage 'Vout' and Source excitation power 'Psource' for the 4-phase, 2 repetition switched reluctance generator.

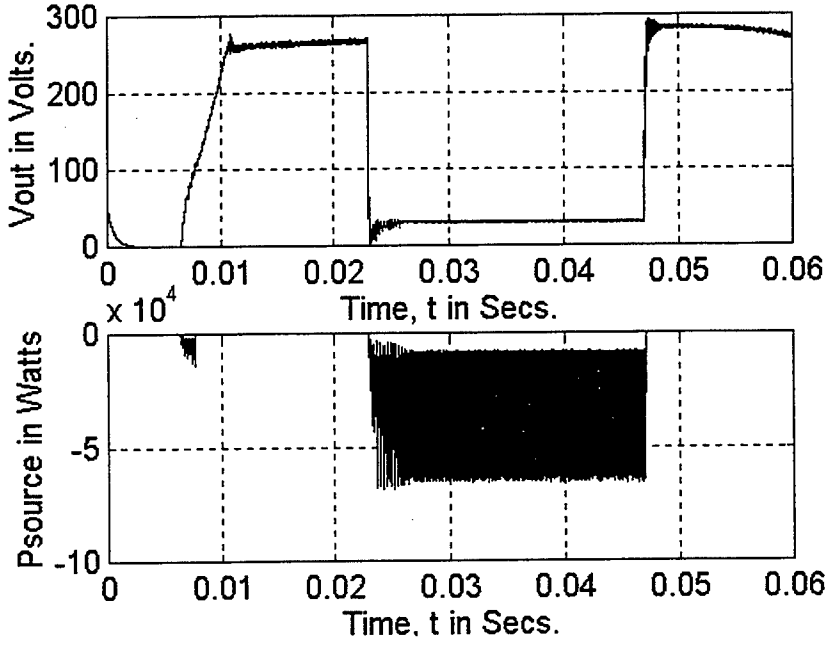


Fig. 22 Output Voltage 'Vout' and Source excitation power 'Psource' for the 4-phase, 3 repetition switched reluctance generator.

VI Conclusions

The excitation requirements for switched reluctance starter-generators for three-phase and four-phase machines were studied through simulation. The excitation requirement for initial voltage build-up in the EIS/G is 12% of the delivered power, while for a load fault condition the power requirement is 30% in one channel. The excitation requirement for initial voltage build-up increases as the number of repetitions increase, but decreases with four-phase machines. The excitation requirements for voltage build-up are 18% for three-phase, 3-repetition, 7% for four-phase, 2-repetition and 9% for four-phase, 3-repetition SR starter/generators. However, four-phase machines have higher excitation requirements under load fault conditions compared to three-phase machines.

The generator is able to maintain its output under a single phase short and a single phase open fault conditions, and also under converter switch open fault condition in one phase leg. The generator does not draw any power from the excitation source under these fault conditions provided there is no load fault. However, for a switch-short fault condition in one converter phase leg, the output voltage collapses and the currents reach exceedingly high magnitudes, which would result in device failures.

The simulation models developed and used in this research can be used for quick preliminary designs of switched reluctance machines and system simulation including fault conditions.

The simulation results obtained in this study need to be experimentally verified. However, experimental verification on the actual EIS/G system may be complicated and expensive. Experiments are recommended to be performed on scaled down versions of the systems in order to evaluate the validity of the simulation tools.

References

[1] C.M. Stephens, "Fault Detection and Management System for Fault Tolerant Switched Reluctance Motor Drives," IEEE-Industry Applications Society Conf. Rec. '89, pp. 574-578.

[2] A.A. Arkadan and B.W. Kielgas, " Switched Reluctance Motor Drive Systems Dynamic Performance Prediction under Internal and External Fault Conditions," IEEE Trans. on Energy Conversion, Vol. 9, No. 1, March 1994.

[3] A.V. Radun, "Generating with Switched Reluctance Motor," IEEE-Applied Power Electronics Conf. Rec. '94, pp. 41-47.

[4] C.A. Ferreira and E. Richter, "Detailed Design of a 250kW Switched Reluctance Starter/Generator for an Aircraft Engine," 1993 SAE Transactions, Vol. 102, Journal of Aerospace, Section 1, pp. 289-300.

[5] E. Richter, J.P. Lyons, C.A. Ferreira, A.V. Radun and E. Ruckstdter, "Initial Testing of a 250 kW Starter/Generator for Aircraft Applications," 1994 SAE Aerospace Atlantic Conference and Exposition, Dayton, OH, April 18-22, 1994.

[6] A.V. Radun,"Design Considerations for the Switched Reluctance Motor," IEEE Trans. on Industry Applications, Vol. 31, No. 5, pp. 1079-1087, Sept./Oct. 1995.

IN SITU FORMATION OF STANDARDS FOR THE DETERMINATION OF WEAR METALS IN PERFLUOROPOLYALKYLETHER LUBRICATING OILS

David W. Johnson
Associate Professor
Department of Chemistry

University of Dayton
300 College Park
Dayton, OH 45469-2357

Final Report for:
Summer Faculty Research Program
Wright Laboratory

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

September, 1996

IN SITU FORMATION OF STANDARDS FOR THE DETERMINATION OF WEAR METALS IN PERFLUOROPOLYALKYLETHER LUBRICATING OILS

David W. Johnson
Associate Professor
Department of Chemistry
University of Dayton

Abstract

Methods have been developed for the in situ formation of metal standards for use in perfluoropolyalkylether based lubricants. The standards are prepared by dissolving a known mass of metal; either as the pure metal or as a metal salt in either the lubricant, containing 5% of a chelating ligand, or in a co-solvent which is used to decrease the viscosity of the lubricant before analysis. Appropriate ligands for perfluoropolyalkyl ether based lubricants include 2,2-dimethyl-6,6,7,7,8,8,8-heptafluoro-3,5-octanedione and 1,1,1-trifluoro-2,4-pentanedione. For non-fluorinated lubricants various other β diketones are useable.

IN SITU FORMATION OF STANDARDS FOR THE DETERMINATION OF WEAR METALS IN PERFLUOROPOLYALKYLETHER LUBRICATING OILS

David W. Johnson

INTRODUCTION

Perfluoropolyalkylethers (PFPAE) are being developed by the Air Force as high temperature lubricants for the next generation of jet engines. While several PFPAE based lubricants are commercially available, there are no good analytical methods for the analysis of metals, due to wear of bearing surfaces in these fluids. Analytical methods based on techniques such as inductively coupled plasma- atomic emission spectroscopy (ICP-AES) which are commonly used for conventional lubricants[11] fail due to the insolubility of the metal standards in PFPAE fluids. The goal of this study was to develop metal standards which are both soluble and stable in PFPAE lubricants and to develop ICP-AES methods for the determination of wear metals in the commercial PFPAE lubricants.

Metal complexes for use as standards in wear metal analysis must be both soluble in PFPAE's to at least 100 ppm metal and stable over a long period of time (months) in the presence of air. The presence of a long fluorinated chain as part of the ligand was expected to provide solubility in the fluorinated oil while improving the hydrolytic stability of the complex[2]. The ligand of choice was 2,2-dimethyl-6,6,7,7,8,8,8-heptafluoro-3,5-octanedione (HFOD). In our previous work, metal FOD complexes were shown to make useable standards for wear metal analysis in Krytox 143AC and Fomblin Z.[3] The stability and utility of these complexes has been demonstrated in fluorinated fluids.[4]

The primary difficulty with this procedure is the need to prepare and isolate the pure β-diketonate complexes and then dissolve them in fluorinated liquids such as Demnum. In this report we describe the in situ preparation of metal standards by the direct reaction of metal powders with the β-diketonates in the fluid of interest. This method is applicable to both conventional lubricants and perfluoroalkyl ether based lubricants and does not require the isolation and purification of specific compounds..

Methods have also been examined for the determination of trace metals in Demnum S-65, a linear perfluoropolyalkylether which has superior properties at high temperatures and has been selected as fluid of choice for initial PFPAE formulations. Demnum, however has a significantly lower ability to dissolve species such as the metal complexes, previously developed as metal standards.

Experimental Section

Materials: The perfluoropolyalkylether based fluids Krytox 143AC (MLO 71-6), Fomblin Z (MLO 78-80), and Demnum S-65 (95-103) were supplied by WL/MLBT and used as received. Freon E-6 (MLO 69-286), a highly fluorinated alkylether with a single remaining hydrogen atom was supplied by WL/MLBT. A sample of the conventional ester based lubricant MIL-L-7808-J was supplied by WL/POSL. The β-diketone ligands 2,2-dimethyl-6,6,7,7,8,8,8-heptafluoro-3,5-octanedione(HFOD), 1,1,1-trifluoro-2,4-pentanedione (TFAcAc), 1,1,1,5,5,5-hexafluoro-2,4-pentanedione (HFAcAc) and 2,4-pentanedione (AcAc) were purchased commercially and used without further purification. Submicron sized metal powders (300 Å) were purchased from Vacuum Metallurgical, Inc. Other metal powders of various sizes were purchased from commercial sources and were of reagent grade or better.

Preparation of Standard Solutions: Solutions were prepared by dissolving 5% of the β-diketone in the appropriate fluid. A known mass of the metal adequate to yield a 1000ppm solution was added to 10 g of the solution of the ligand in oil. The test tube was placed in an ultrasonic cleaner for 30-45 minutes. The resultant mixture was allowed to stand overnight to allow the solubility equilibrium to be established.

Inductively Coupled Plasma-Atomic Emission Spectroscopy

Metal concentrations in PFPAE fluids were measured using a Jarrell Ash Polyscan 61-E spectrometer. Standards were prepared by dissolving known masses of each metal FOD complex in Freon E6 to give solutions of 25 ppm of each of the metals. Samples of the PFPAE fluid were mixed with appropriate amount of the standard solution and Freon E6 in a 1:4 ratio of oil to Freon in order to obtain the various concentrations of standard solutions. All oil samples were diluted 1:4 with Freon E6 before analysis. For analysis, an RF power of 1750 watts, a nebulizer pressure of 22 PSI and a pump rate of 80 RPM was used. The samples were aspirated for a minimum of 45 seconds to insure equilibration within the plasma.

Results

In Situ Preparation of Metal Standards: The preparation of metal β-diketonate complexes by direct reaction of the metal with a β-diketone in an inert solvent is well known[5]. It is not always the most convenient method when isolation of the pure complex is desired[6] because the complexes often

will not come out of solution easily. Even when the complex precipitates, solvent molecules often adhere to the complex resulting in non-stoichiometric compounds yielding poor elemental analyses.

The preparation of standards for metal analysis does not require isolation of a pure compound, just the preparation of a solution of known metal content. Initially, it was thought that submicron sized metal powders could be suspended to give solutions of a known metal concentration. Although the perfluoropolyalkylether based lubricants are quite viscous, the metal powders were found to settle out of solution on standing aftera day or two. The addition of a surfactant (perfluorooctanoic acid) was also attempted in order to stabilize suspensions of the metal powder. The addition of perfluorooctanoic acid increased the viscosity of the solution but did not significantly improve the stability of the suspension. The surfactant did, however increase the amount of time needed to prepare the suspension of the metal powders.

The use of β-diketonate ligands to prepare metal standard solutions began as an attempt to stabilize the submicron metal powders in solution. In the initial attempts, 0.1%, 1% and 5% solutions of the FOD ligand were prepared and added to 0.10g of iron powder (300 Å average particle size). After mixing with the ultrasonic cleaner, the solutions were allowed to settle overnight. After settling, the samples which contained 1% and 5 % FOD had a definite color characteristic of the complex $Fe(FOD)_3$. The solution was also observed to contain a small amount of unreacted iron powder.

In order to demonstrate the utility of this method for the preparation of metal standards, use of several β-diketones, metals and fluids were investigated. The metals included iron, chromium, nickel, magnesium, zinc, and molybdenum. In addition to the submicron sized iron, a standard iron powder (80 mesh) was also used. Ligands investigated included HFOD, 1,1,1-trifluoro-2,4-

pentanedione(TFAcAc), 1,1,1,5,5,5-hexafluoro-2,4-pentanedione (HFAcAc) and 2,4-pentanedione (AcAc). The fluids investigated were Demnum S-65, Krytox 143AC, Freon E-6 and MIL-L- 7808 (an ester based lubricant in current use). The results of these studies are summarized in Table 1.

The more reactive metals (Fe, Mg and Zn) were observed to react completely to yield a saturated solution of the metal complex in each of the fluids. If the quantity of metal added was more than was necessary to prepare a saturated solution, solid metal(β-diketonate)$_n$ was observed to be floating on top of the fluid. The degree of reaction was independent of the particle size of the metal powder. Less reactive metals(Cr, Ni and Mo) remained unchanged at the bottom of the test tube even after several days.

Several different β-diketone ligands were investigated to see if changes in structure would enhance reactivity, increase solubility or decrease cost. Data in Table 1 shows that the FOD ligand reacts to form complexes with the highest solubility in the perfluorinated fluids. TFACAC and HFACAC, while they react readily to form the metal complexes with Fe, Mg and Zn, do not appear to be any more reactive towards Cr, Ni and Mo. The decrease in the solubility of the metal complexes mitigates against the use of these β-diketonates as metal standards in perfluorinated liquids.

Variation of the reaction medium through variation of the fluorinated fluid showed only variations due to changes in the solubility of the complex in the reaction medium. The results show the expected trend where the complex was most soluble in the low molecular weight , monohydrofluoroether Freon E-6. The complexes showed intermediate solubility in the branched

Table I. Reactivity Data For Metals In Perfluoropolyalkylether Based Fluids

Metal	Size	Ligand	Fluid	[M]ppm
Fe	300	HFOD	Demnum S65	600
Fe	300	HFOD	Krytox	750
Fe	300	HFOD	Freon E-6	800
Fe	100 mesh	HFOD	Demnum S65	610
Fe	100 mesh	HFOD	Krytox	740
Fe	100 mesh	HFOD	Freon E-6	810
Fe	100 mesh	TFACAC	Demnum S65	380
Fe	100 mesh	TFACAC	Krytox	430
Fe	100 mesh	TFACAC	Freon E-6	510
Fe	100 mesh	HFACAC	Demnum S65	430
Fe	100 mesh	HFACAC	Krytox	490
Fe	100 mesh	HFACAC	Freon E-6	510
Mg	80 mesh	HFOD	Demnum S65	300
Mg	80 mesh	HFOD	Krytox	320
Mg	80 mesh	HFOD	Freon E-6	350
Mg	80 mesh	TFACAC	Demnum S65	260
Mg	80 mesh	TFACAC	Krytox	280
Mg	80 mesh	TFACAC	Freon E-6	290
Mg	80 mesh	HFACAC	Demnum S65	280
Mg	80 mesh	HFACAC	Krytox	290
Mg	80 mesh	HFACAC	Freon E-6	310
Zn	20 mesh	HFOD	Demnum S65	430
Zn	20 mesh	HFOD	Krytox	460
Zn	20 mesh	HFOD	Freon E-6	500
Zn	20 mesh	TFACAC	Demnum	320
Zn	20 mesh	TFACAC	Krytox	360
Zn	20 mesh	TFACAC	Freon E-6	380
Zn	20 mesh	HFACAC	Demnum	380
Zn	20 mesh	HFACAC	Krytox	410
Zn	20 mesh	HFACAC	Freon E-6	440

perfluoropolyalkylether; Krytox 143 AC and the lowest solubility in the linear perfluoropolyalkylether Demnum S-65. No changes in the reactivity of the various metals was observed through these fluids suggesting that only changes in the solubility of the metal complexes is being observed. When solutions were prepared using masses below the solubility limit of the metal complex, the analysis of the solution was within experimental error of the value predicted from the mass of metal and the mass of fluid.

Trace Metal Analysis in Demnum S-65

In our previous work, methods were developed for the analysis of wear metals in Krytox 143AC and Fomblin Z. Since Demnum S-65 has superior properties as a lubricant, it has been chosen as the basestock for further study. The validity of the previously developed methods was verified using Demnum S-65 and the method was used to determine the metal content of a series of samples from the rolling contact fatigue tester (RCF test).

Solutions of each metal FOD complex were examined using ICP to determine if the wavelengths available on the polychrometer would suffer from any serious interferences. Only the aluminum complex showed any interference. The observed interference is similar to the interference observed when oil samples are analyzed. A calibration curve for each of these metals, using concentrations from 1 ppm to 100 ppm has been measured with the results similar to those observed for Krytox 143AC. The calibration curves are linear throughout the range of concentrations studied. Estimates of the detection limit based on three times the standard deviation of the readings of the blank solution indicate that the detection limits in Demnum S-65 are slightly poorer than those observed in Krytox. Detection limits for the various metals in solution are in the range of 2-50 ppb under the conditions of this analysis, corresponding to concentrations between 10 and 250 ppb in the

original oil sample(Table II). The linear calibration data and the observed detection limits indicate that the method is usable for the determination of wear metals in Demnum S-65.

Table II. Calibration Data for Metals in Demnum S-65

Element	λ	Correlation Coefficient	Detection Limit (ppb)	Oil Detection Limit (ppb)
Al	308.2	0.999	44	220
Ba	493.4	1.000	2	10
Cd	228.8	1.000	8	38
Co	228.6	0.994	18	90
Cr	267.7	0.999	32	160
Cu	324.7	1.000	16	80
Fe	259.9	1.000	8	40
Mg	279.5	0.999	3	15
Na	588.9	0.999	20	100
Ni	231.6	0.999	24	120
Zn	213.8	0.999	30	150

A series of RCF test samples were analyzed by ICP-AES for the various trace elements. The results of these analyses are summarized in Table 3. The data presented indicate that metal is dissolved by the perfluoropolyalkylethers during the RCF test. This data is consistent with other data such as corrosion of the metal specimen used in the test. Other metals were found at concentrations below the level necessary for quantification.

Table III. Trace Metal Analysis for RCF Test Samples. Fluid is Demnum S-65 at 274 C unless otherwise noted.

MLO Number	Fe ppm	Cr ppm	Zn ppm	Cu ppm
MLO 96-33	0.9	0.1	0.3	0.3
MLO 96-34	0.3	0.0	0.2	0.1
MLO 96-35	0.4	1.0	0.1	0.2
MLO 96-36 330 C	0.2	0.0	0.1	0.1
MLO 96-38 1%PH3	4.9	0.1	0.6	1.5
MLO 96-62	1.3	0.0	0.1	0.1
MLO 96-63 2nd Pass	16.5	0.7	0.3	0.4
MLO 96-73	0.6	0.1	0.2	0.1
MLO 96-74 1% PH3	1.0	0.0	0.2	0.3
MLO 96-75	0.5	0.0	0.1	0.0
MLO 96-76	0.2	0.2	0.1	0.0

Conclusions

The research presented in this report indicates that the in situ formation of metal complexes is a valid method for preparation of standard solutions for use in ICP-AES. The reaction with β-diketones proceeds slowly to give saturated solutions of the metal complex in the fluid of choice. The concentrations of metal in the various fluids are adequate for use of the solution as a standard for ICP-AES.

Metal complexes of FOD have been shown to be usable for the analysis of wear metals in Demnum S-65 using methods similar to those developed earlier for Krytox 143AC. The calibration curves are linear with high correlation coefficient and the curves yield detection limits only slightly higher than those seen previously for Krytox.

References Cited

[1] Eisentraut, K.J.; Newman, R.W.; Saba, C.S.; Kauffman, R.E. and Rhine, W.E.; Anal. Chem.; **56**, 1086A, (1984).

[2] Eisentraut, K.J. and Sievers, R.E.; J. Am. Chem. Soc., **87**, 5254, (1965).

[3]Johnson, D.W. "Analytical Methods For The Determination Of Wear Metals In Perfluoropoly-alkylether Lubricating Oils" Final Report, 1993 AFOSR Summer Faculty Program, Research And Developement Laboratories, CA.

[4]Eisentraut, K.J. and Johnson, D.W. "Analysis of Wear Metals in Perfluorinated Fluids" U.S. Patent 5,506,677, April 1996.

[5] Grandon, D.P. *Coord. Chem. Rev.*, 1963, *4*,1.

[6] Hammond, G.S. Nonhebel, D.C. and Wu, C.S. *Inorg. Chem.* **1963**,*2*, 73.

AIRCRAFT SUPER CAPACITOR BACK-UP SYSTEM

Marian K. Kazimierczuk
Professor
Department of Electrical Engineering

Wright State University
Dayton, OH 45435

Final Report for:
Summer Faculty Research Program
Wright Laboratory

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

and

Wright Laboratory

August 1996

AIRCRAFT SUPER CAPACITOR BACK-UP SYSTEM

Marian K. Kazimierczuk
Professor
Department of Electrical Engineering
Wright State University
Dayton, OH 45435

Abstract — The feasibility of the use of a super capacitor to improve the back-up power system of an aircraft was studied. The back-up power system consists of a super capacitor charger, a bank of super capacitors, and boost dc-dc converter. A novel super capacitor charger was invented. A boost dc-dc converter was used to transfer the energy from the super capacitor to the 28-V dc bus and regulate the voltage. A new control technique introduced recently was used to achieve a good line regulation of the boost converter over a wide input voltage range. Experimental results indicate that a super capacitor and DC/DC converter can be used to improve the regulation of the bus voltage of distributed power systems.

A back-up power system with a super capacitor for aircraft applications was studied. It consists of a super capacitor charger, a super capacitor, and a dc-to-dc converter. The nominal value of the input voltage of back-up system is $V_I = 28$ V and the nominal value of the output voltage is $V_I = 24$ V. The output current is 1 A. A bank of super capacitors was build using 10 F/2.5 V Panasonic super capacitors. It consists of 10 super capacitors in series, which gives a 1 F/25 V super capacitor. Next, 3 such super capacitors were connected in parallel, resulting a 3 F/25 V super capacitor. This super capacitor was used in all experimental test of the back-up system.

The super capacitor charger is a difficult circuit to build. The voltage across the charger is

$$\Delta V = V_I - V_C. \tag{1}$$

where V_C is the voltage across the super capacitor. The input current I_I is approximately equal to the capacitor current I_C. Therefore, the power dissipated in the charger is

$$P_D = \Delta V I_C. \tag{2}$$

The amount of energy stored in the super capacitor C at the charging end

$$E_C = \frac{1}{2} C \Delta V_C = \frac{1}{2} \times 3 \times 24^2 = 826 \text{ J}. \tag{3}$$

On the other hand, this energy can be expressed in terms of the power delivered to the super capacitor $P_C = V_C I_C$ and the charging time Δt

$$E_C = P_C \Delta t = V_C I_C \Delta t. \tag{4}$$

When the voltage across the super capacitor is zero, the voltage across the charger is $\Delta V = V_I = 28$ V. Assuming that the capacitor current $I_C = 1$ A, the power loss in the charger at the beginning of the super capacitor charging is

$$P_D = \Delta V I_C = V_I I_C = 28 \times 1 = 28 \text{ W}. \tag{5}$$

This large amount of power loss cannot be safely and reliably dissipated in linear voltage regulators. Notice that the power loss in the super capacitor charger is higher than that in

battery charger. This is because the output voltage of a battery charger is always higher than zero, whereas the output voltage of the super capacitor is zero for the worst case. A novel super capacitor charger was invented and is described in the patent application "Super Capacitor Charger."

The input voltage of the boost dc-dc converter changes over a very wide range, typically from 24 to about 10 V during the super capacitor discharging. On the other hand, the voltage at the output of the super capacitor should be held constant at 28 V. Therefore, a good line regulation is required from the boost dc-dc converter. A novel and simple control technique was invented that offers an excellent line regulation of the boost converter. This invention is described in the patent application "High Reliable, Low Maintenance Back-up Power System for Fixed Loads."

CORROSION RESISTANT SOL-GEL COATINGS FOR AIRCRAFT ALUMINUM ALLOYS

Edward T. Knobbe

Associate Professor
Department of Chemistry and the
Center for Photonics and Laser Research
Oklahoma State University
Stillwater, OK 74078

Final Report submitted to RDL for:
Summer Faculty Research Program

Sponsoring Agencies:
Wright Laboratory WL/MLBT
Wright Patterson Air Force Base, Dayton, OH
and
Air Force Office of Scientific Research
Bolling Air Force Base, DC

September 1996

CORROSION RESISTANT SOL-GEL COATINGS FOR AIRCRAFT ALUMINUM ALLOYS

R.L. Parkhill and E.T. Knobbe
Department of Chemistry and the
Center for Photonics and Laser Research
Oklahoma State University

Abstract

Sol-gel films were investigated as potential replacements for chromate-based surface treatments on aircraft aluminum alloys. Aluminum alloy 2024-T3 test coupons coated with protective sol-gel films were shown to provide greatly enhanced corrosion protection compared to current protective treatments such as alodyne 1200. Although substantial cracking and pinhole defects were found in most of the films prepared, improvements of up to six orders of magnitude in corrosion resistance were found for selected films. The most promising protection improvement was found in the case of a passivating cerium doped silica subbing layer with an organically-modified silicate (ormosil) overcoat. The sol-gel bilayer was found to give corrosion protection which rivaled or exceeded the complete alodyne/primer/paint topcoat system currently in use by the Air Force.

CORROSION RESISTANT SOL-GEL COATINGS FOR AIRCRAFT ALUMINUM ALLOYS

R.L. Parkhill and E.T. Knobbe

1. Introduction

Recently, the Air Force has encountered serious problems concerning their aircraft protective coating system. Military and environmental agencies are presently implementing restrictions that require a new coating system to be developed which is capable of meeting improved environmental and lifetime standards [1]. With military estimates of increasingly limited flight hours per year and expensive replacement costs, present and future aircraft will be expected to have service lifetimes greater than 30 years, and up to 80 years in certain instances. In addition, environmental regulations have mandated a replacement of all coating processes which rely on the use of chromate passivating agents, such as strontium chromate, which are currently the standard used in alodyne 1200 treatment. Replacement protective coating must be capable of satisfying extended lifetime requirements, must be compatible with environmental requirements, and must be easily integrated into the current primer/topcoat paint systems.

Recent research has been directed the towards development of a new coating process employing alternate passivating agents. Preliminary investigations indicate that the one of the most promising approaches involve a sol-gel coating system doped with cerium passivating additives [2-5]. The studies described here in suggest that a composite sol-gel systems will provide a protective barrier capable of meeting both environmental and military standards.

2. Methodology

Over the past 25 years, sol-gel chemistry has proven to be useful over a wide range of thin film applications [6-9]. The sol gel method provides a low temperature route to the preparation of thin ceramic coatings, which are readily applied to most metallic substrates. The sol-gel method has the potential to incorporate chemically tailored features aimed at the development of corrosion resistance through the use of structurally modified chemical reactants. The sol-gel method represents a promising approach to the development of a corrosion protection surface pretreatment as a result of salient preparation and chemical aspects, including: room temperature synthesis, chemical inertness, barrier protection, and commercially demonstrated scale-up potential.

The requirement for developing a new coating system originated from primarily environmental concerns. The existing chromate system required the employment of stringent hazmat precautions during both film application and removal processes. The sol-gel coating process eliminates longterm environmental concerns attributed to the application of corrosion-resistant

coating. Once applied and cured, the film and waste products become chemically inert and require little or no additional hazmat handling during subsequent paint removal procedures.

Exposure of aluminum alloys to elevated temperatures during the passivation coating process is a primary concern. The T-series aluminum alloy preparation requires controlled heating to precipitate Al_2CuMg dispersoids in the aluminum matrix [10]. Specific control of particle size and composition determines the hardness and strength of such alloys. If annealing temperatures are exceeded, the Al_2CuMg particles undergoes additional phase separation, leading to irreversible strength loss and possible structural failure. Sol-gel coatings are compatible with T-series aluminum alloys since film application and cure temperatures are typically 150°C or less. Thus, alloy damaging overaging effects are avoided.

Silicates and organically modified silicates (ormosil) represent a promising means of corrosion protection for metallic substrates. By combining both organic and inorganic sol-gel reagents, thin films which incorporate beneficial characteristics associated with composite networks are readily deposited. The ceramic portion of such layers provide improved adhesion and stratch-resistance, while the organic network provides additional film flexibility and density. The ormosil-type films yield exceptional durability and adhesion, while providing a diffusion barrier to permeont corrosion-inducing species. of water and corrosion initiators.

Current uses of sol gel in industry support the concept that large scale coating systems can be developed which meet or exceed environmental targets. Two different techniques, spin- and dip-coating, have been used in applying commercial sol-gel coatings, while spray-coating has been demonstrated in the laboratory. Applications that use the dip-coating method have included heat mirrors and head up displays (HUD) for automobiles, while spin-coating has been used to apply scratch resistant coatings to various substrates.

3. Background

3.1. Corrosion

Military aircraft are largely constructed from 2024-T3 and 7075-T6 aluminum alloys. The alloys are composed of aluminum (>90%) with small amounts of elemental metals added to improve mechanical strength, as indicated in Table I [11]. While pure aluminum is typically considered to be chemically inert, the presence of secondary phase regions which are rich in transition metal species provide a more facile mechanism for corrosion to occur. Exposure to harsh environmental conditions such as acid rain, salt water, and thermal cycling increase the rate of corrosion in these alloy systems.

TABLE I Elemental Compositions of Aluminum Alloys 2024 and 7075.

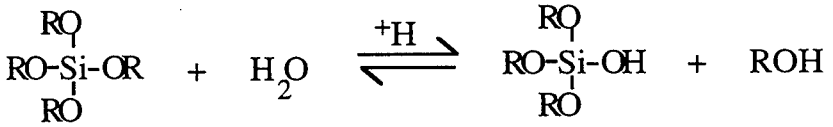
	Elemental Composition (wt.%)								
Alloy	Al	Cu	Mg	Mn	Si	Fe	Cr	Zn	Ti
AA2024	93.5	4.4	1.5	0.6	——	——	——	——	——
AA7075	Remainder	1.2	2.1	0.3	0.4	0.05	0.18-0.28	5.1-6.1	0.2

The mechanism of corrosion in aircraft aluminum alloys is not well understood, but is more generally regarded as a redox reaction involving $Cl^-_{(aq)}$, resulting in the formation of dissolved aluminum cations [12,13]. The water-soluble cations are transported away from the surface leading to a cyclic corrosion cascade. In order to block the corrosion mechanism, an impervious coating layer must be developed to eliminate the primary transport of common corrosion initiators such as $Cl^-_{(aq)}$.

3.2. Sol-gel Method

The sol-gel method consists principally of hydrolysis and condensation reactions, originating with alkoxysilane precursors to form a polymeric network of nanoporous glass[6]. The reaction sequence initially proceeds through an acid catalyzed hydrolysis reaction which produces a hydrolyzed silane and an alcoholic by-product (Figure 1a). The condensation reaction proceeds by condensation of two hydrolyzed silane precursors to produce siloxane bonds with the elimination of water (Figure 1b). The reaction sequence continues in a polymeric manner resulting in the formation of a porous silica network.

a. Hydrolysis:



b. Condensation

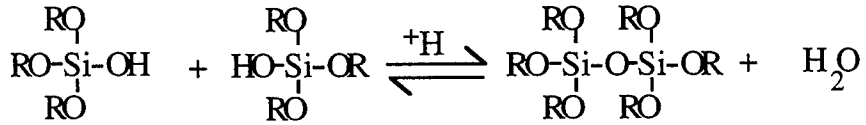


Figure 1 *Basic sol-gel polymerization reactions for producing silica.*

Ormosil systems consist of a comparable reaction sequence with the addition of an organic polymerization reaction [6,14]. In a ormosil gel, the silicate portion of the precursor undergoes hydrolysis and condensation, tending to produce a porous silica network with unreacted organic

polymer chains dispersed throughout. The organic polymerization reaction may be activated through heating or through the use of UV initiators. Polymerization of the organic components causes the porous silicate network to form a dense film, composed of intimately mixed organic and inorganic phases.

3.3. Corrosion Inhibiting Additives

The use of corrosion-inhibiting additives in coating systems, such as alodyne 1200, has been one of the primary means of developing environmental substrate passivation [13]. Currently, the most widely used and effective inhibitors are based on chromate salts such as $SrCrO_4$. Studies have confirmed the passivating nature of chromate-based surface treatments on aluminum, but have not conclusively identified the mechanism by which corrosion inhibition is achieved. In the presence of $Cl^-_{(aq)}$ ions, Cu^0 from the alloy is oxidized to Cu^+. Researchers generally speculate that $CrSO_4$ reacts with water in the presence of a galvanic electrode, giving rise to the formation of $Cr(OH)_3$. The trivalent form of chromium acts as the chemically-passivating species by acting as an electron source and initiating a reverse reduction reaction. Passivation occurs when Cu^+ is reduced backed to Cu^0 with the production of Cr^{+3} [13]. The trivalent chromium species from $Cr(OH)_3$ is also postulated to be responsible for the formation of a dense passivating Cr_2O_3 layer near the aluminum interface. Thus, corrosion is inhibited by: (1) a reduction in Cu and other transition metal loss, and (2) the formation of an inert, impermeable layer.

With new regulations restricting chromate use, alternative passivating additives are now being evaluated. Unfortunately, the absence of a clear mechanistic understanding of chromate passivation has slowed the development of modern corrosion inhibitors to supplant chromate-based media.

4. Experimental Methods

4.1. Reagents

Substrates consisted of aluminum 2024-T3 coupons which were polished with 600 grit sand paper and cleaned in an ultrasonic bath with isopropanol. Tetraethoxy silane (TEOS) and tri-sec butoxyaluminoxytriethoxysilicate (DBATES) were used as received from Aldrich and United Chemical for preparation of inorganic phase sol-gel films, respectively. Organically modified silicon alkoxide reagents were used as received from Aldrich. Cerium compounds were used in corrosion passivation experiments by direct dissolution into the aqueous sol. Spectroscopic grade solvents, including methanol, ethanol, and ethylene glycol, were used as received from Aldrich. 2,2-azobisisobutyronitrile (AIBN) and CO_2 gas were used as catalysts.

4.2. Sol-gel Preparations

Preparation of the TEOS (2:1) solution was performed as follows: 12.0 ml of TEOS was placed in a sonicator and diluted with 3 ml of ethanol. 4.4 ml of water, bubbled with gaseous CO_2 was added in a dropwise fashion until the solution became cloudy. At the onset of a cloudy appearance, additional alcohol was added until the solution cleared. These additions continued until the requisite of water was incorporated. A total of 16.4 ml of ethanol was used as co-solvent. The solution was allowed to sonicate at T<25°C for 1hr and then stirred at room temperature for 24 hrs before spin coating. The aluminosilicate was prepared in a similar manner, using instead a 1:1 water:alkoxy molar ratio. Ormosil preparations were similar, with the addition of organically modified silicon alkoxide reagents.

4.3. Coating and Curing Technique

Sol-gel films were spun onto aluminum 2024 substrates using a Solitec multiple stage spin coater. The sol gel coating was applied by a two stage spinning process. During the first stage, 2 ml of the sol solution was applied to a sample spinning at 300 rpm for 15 seconds. Immediately afterwards the sample speed was increased to between 1500-3000 rpm for 60 seconds. Samples were cured using 2- or 3-stage temperature treatments. Both the silicate and ormosil coated samples were heated to 70°C for 12 hours, and subsequently heated to 110°C for 24 hours. Silica samples underwent an additional curing at 300°C for 24 hours. The samples were then removed from the ovens and allowed to cool to room temperature before analyzing.

4.4. Characterization

Sample morphology was characterized using an OPTIPHOT-POL Nikon polarizing microscope. Elemental film composition analysis was determined by Perkin Elmer X-ray photoelectron spectroscope (XPS) equipped with Auger Spectroscopy. Film thickness was determined on a Sloan Dektak IIA profilometer. The corrosion potentials were determined by potentiodynamic polarization curves. A Gamry potentiostat was used to sweep potentials from -1.0V to 0.2V vs. SCE of coated samples immersed in a 3% NaCl solution. Specimens were typically maintained in the 3% NaCl for 90 minutes prior to potentiodynamic polarization measurement, except in the case where the effects of longer term immersion are noted. Data was analyzed using CMS100 software.

5. Results and Discussion

Table II gives an indication of the variety of film types that were prepared during the summer program. Each film was divided, for the purpose of characterization, into 4 sub-groups: single

layer inorganic, single layer ormosils, bilayer, and cerium doped systems. Each group was compared to one of three protection standard specimens: untreated 2024-T3 aluminum alloy, alodyne 1200 treated 2024-T3, or a complete paint system applied to a 2024-T3 coupon. The preliminary assessment of sol-gel protection was evaluated on the basis of film morphology, composition, and potentiodynamic polarization measurements.

Table II Sol-Gel Layers Studied

Inorganic	ORMOSIL	Bilayer Structures	Films with Inhibitor
SiO_2	S/A	SiO_2 coated with S/V/E	Ce(III) in SiO_2
2 SiO_2 :Al_2O_3	S/E		Ce(III) in S/A
	S/V/E		Ce(III) in SiO_2 with S/V/E overcoat

5.1. Inorganic Systems

Figure 2 shows an optical micrograph of a typical silica film. At 400X, the image revealed significant cracking throughout the film . Fracture widths were estimated to range from 1 to 30 μm. Areas which remained crack-free appeared featureless.

The development of cracks occurred during the curing process. We postulate that the use of elevated curing temperatures created stresses in the ceramic film due to differences in the thermal expansion coefficients between the aluminum and the silica film. Normally, films with thickness between 5 and 500 nm do not develop a sufficient stresses to cause cracking [15]. However, the silica films prepared were estimated to have thickness near 1.5 μm. Thus significant stresses arose in the drying film, resulting in film fractures.

X-ray photoelectron spectroscopy (XPS) was used to confirm the composition of the silica film. Figure 3 represents the XPS spectrum of a silica film. Spectral analysis indicated the presence of three different elements. Silicon was identified by peaks

Figure 2. *Optical micrograph of a sol-gel silica film on aluminum alloy 2024-T3 (magnification = 400X).*

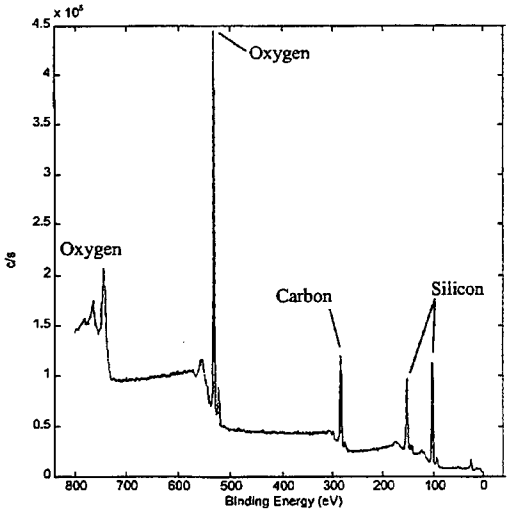


Figure 3. *XPS spectrum of a sol-gel silica film on aluminum alloy 2024-T3.*

at 100eV and 150eV, oxygen by peaks at 530eV and 750eV, and carbon by a peak at 290eV. Peak areal measurements indicated a 2:1 ratio of silicon to oxygen and traces of carbon, confirming the dominant silica phase with a small amount of residual alcohol or alkoxide.

Electrochemical analysis of the silica film indicated protection characteristics which were much better than anticipated, given the degree of film cracking observed. Figure 4 represents a protection characterization comparison between silica gel coated, alodyne 1200-treated, and untreated 2024-T3 alloy coupons. The unprotected alloy yields comparatively high current densities at all potentials, indicating high corrosion rates in the NaCl solution. The alodyne 1200 treated aluminum has a broad potentiodynamic "knee" at a log current density value of -3.5 A/cm², indicating a better than 10-fold improvement in corrosion resistance character over the untreated alloy. Hereinafter we shall use the term "corrosion protection" value, or CP value, to describe the -[Log current density (A/cm²)] knee found in the potentiodynamic plot. Accordingly, Alodyne 1200 treated 2024 test coupons were found to have a CP value of 3.6. The silica gel-coated alloy has a sharp double-knee feature corresponding to a CP value of 7.0. This peak is indicative of corrosion protection character that is 3.5 orders of magnitude better than that found for the chromate-based alodyne treatment at a comparable potential (eg., vs. E_{ref}).

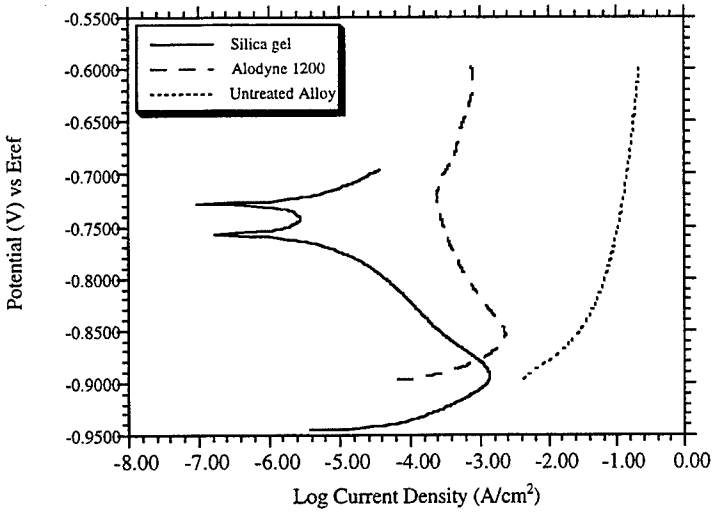


Figure 4. *Polarization curves for silica gel coated, alodyne 1200-treated, and untreated 2024-T3 alloy coupons.*

The improved protection character of sol-gel silica films over alodyne suggests the following: (1) The cracking in the silica film was either surface confined and/or protected by some type of silica passivation to develop a continuous barrier. If direct contact between the aluminum substrate and salt solution occurred over a widespread scale, the film would be expected to corrode in a fashion similar to that found for unprotected coupons did. This is clearly not the case. (2) Since both the silica film and alydyne coatings rely on a metal oxide barriers for protection, the silica layer must either be denser and/or thicker, than the alodyne treatment. The 10 fold protection increase without the use of a corrosion inhibitor suggests that the silica film provides an improved barrier to Cl^- anions penetration and water permeation.

5.2. Ormosil Films

Figure 5 contains optical micrographs of the ormosil S/V/E- and S/A-type films. Both ormosil compositions were found to yield crack-free and largely featureless films except for scattered pinholes. The S/A film, Figure 5(a), had two distinguishably different sized pinholes. The smaller pinholes, with diameters averaging approximately 16 μm, were uniformly dispersed throughout the film with a density near 2 holes/mm^2. The larger pinholes, having diameters ranging from 30-50 μm, were determined to have a distribution of approximately 43 holes/cm^2. The S/V/E film, Figure 5(b), was found to be free of the smaller pinholes, but exhibited the 30-50 μm sized pinholes with a distribution that was found to be comparable to the S/A composition.

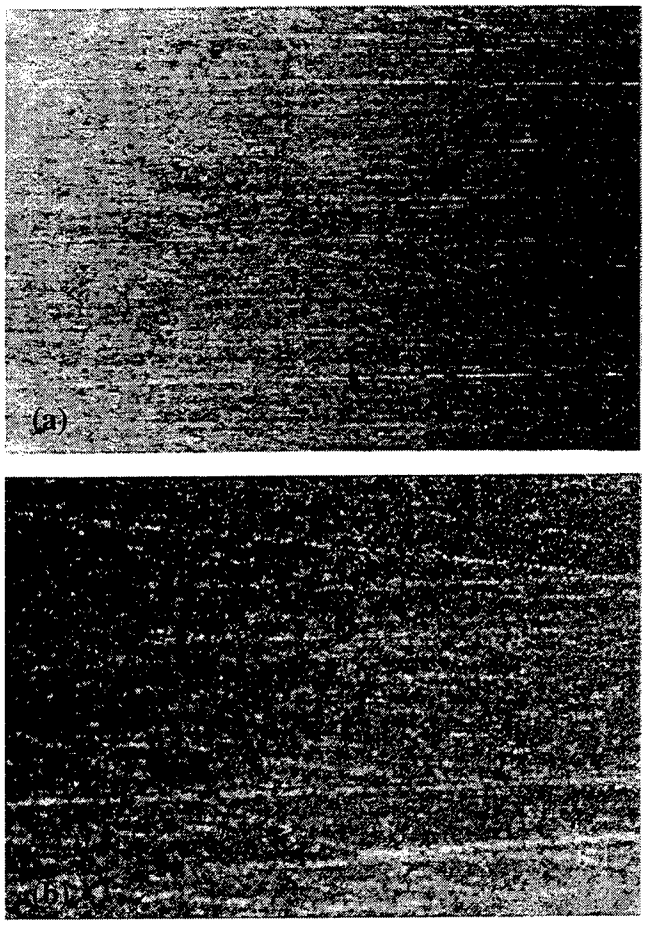


Figure 5. *Optical micrographs at 50X of two ormosil films on aluminum alloy 2024-T3; (a) S/A and (b) S/V/E*

Figure 6 represents a profilometry plot of the S/A film. The scan was recorded over a 1 mm section, using a scanning rate of 0.1 mm/s. The first 0.2 mm of the film scan was disregarded due to the defective edge formation. The deposited film thickness was averaged over a 0.2 mm section and determined to be 4.5 μm, nearly 3 times the thickness of a conventional defect-free ormosil coating.

The increased thickness along with wetting problems most probably contributed to the development of pinhole defects. It is postulated that the evaporation of excess solvent from the thick coating modified the

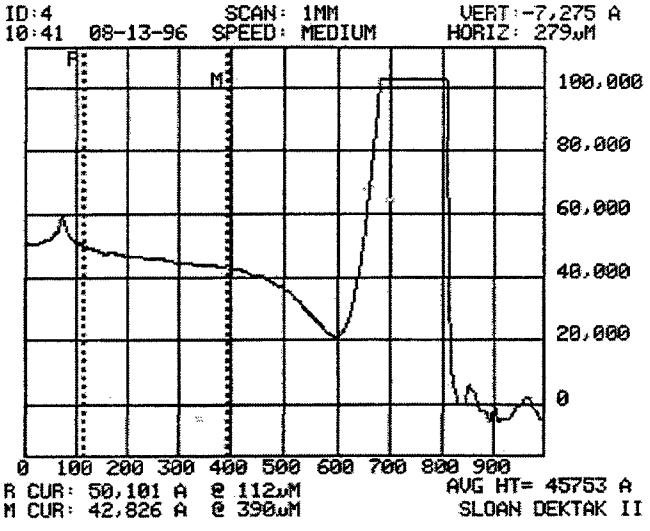


Figure 6. *Profilometry plot of S/A film on aluminum alloy 2024-T3.*

surface tension of the sol, creating voids in the film and subsequent pinhole formation. Additionally, the presence of AIBN polymerization catalyst could have contributed to the development of the small pinholes due to N_2 release upon decomposition of the azide initiator.

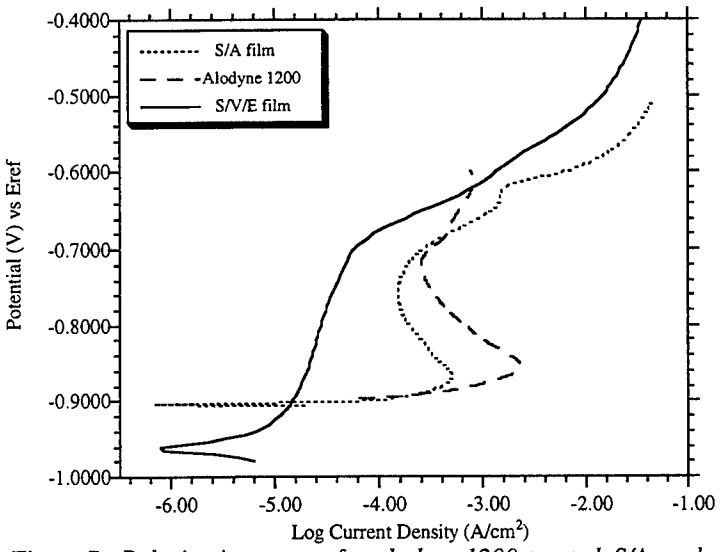


Figure 7. *Polarization curves for alodyne 1200-treated, S/A, and S/V/E coated 2024-T3 alloy coupons.*

Electrochemical evaluation of the ormosils showed greatly improved initial corrosion resistance, but a reduced stable potentiodynamic range. Figure 7 represents a comparison between the alodyne 1200, S/A, and S/V/E film coatings. The two ormosil films showed very sharp potentiodynamic knees corresponding to CP values near 6.0 The sharpness of these features is indicative of limited stability over extended potential ranges, a result which is consistent with a catastrophic film integrity failure mechanism.

The result suggests that the degradation in protection potential may be attributed to the presence of pinholes, which provided a route to premature film failure. Since the ormosil does not have a passivating mechanism, corrosion is initiated as soon as the tenuous integrity of these defect sites are compromised. Pinholes clearly behave as a weak point in the film, and their presence is expected to limit the range of potentials over which corrosion protection may be maintained.

5.3. Bilayer Films

The bilayer film systems were composed of a silica base film and a S/V/E overcoat layer. The surface of the films appeared to be more or less indistinguishable from the ormosil films shown in Figure 5. Figure 8 shows two optical micrographs taken under different magnification. Figure 8(a), taken under low magnification (50X), shows the large pinhole density present in the S/V/E film. Pinhole diameters were determined to range from 2-35 μm. Figure 8(b), taken at higher magnification (400X), shows cracking in the silica film; fissure widths were estimated to range from 1-4 μm.

Electrochemical evaluation of the 2-layer system is presented in Figure 9. Comparison between the silica-S/V/E system and the reference alodyne treatment shows a nearly 4 order of magnitude improvement over the chromate-based preparation. The silica-S/V/E coating has CP value of 7.8 and shows a comparatively slow degradation behavior over a wide potential range.

Physical defects in the multilayer system identified by optical microscopy were identical to problems associated with single layer systems. The S/V/E film developed a increase in pinhole coverage, a result which may be attributed to poor wettability of the ormosil layer on the silica film. Silica film cracking is believed to arise from stresses induced by overly-thick films.

Substantial protection improvement over single-layer compositions were attributed to the additional barrier provided by the ormosil overcoat. A comparison between the single layer ormosil and 2-layer coating suggested a dramatic reduction in corrosion species transport rates. Reduced degradation indicates that pinhole defects of the ormosil were compensated by the underlying silica layer. Thus, the bilayer system was found to offer approximately four orders of magnitude increase in corrosion protection over the conventional alodyne 1200-treated specimens.

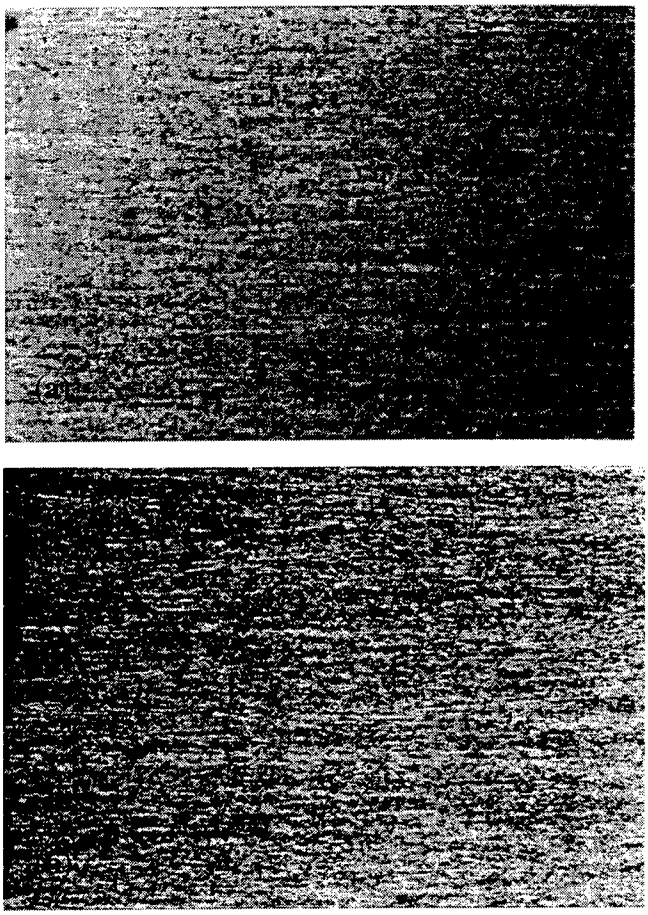


Figure 8. *Optical micrographs of a sol-gel silica film overcoated with an S/V/E film on aluminum alloy 2024-T3; (a) 50X and (b) 400X.*

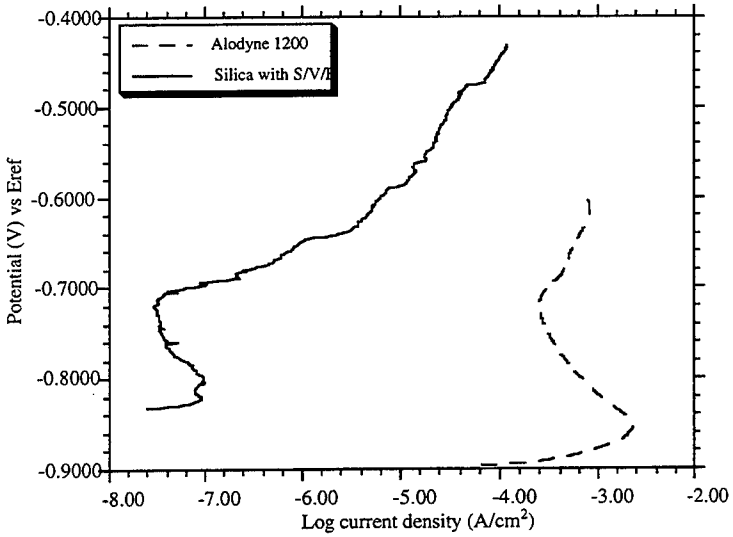


Figure 9. *Polarization curves for alodyne 1200-treated, and silica-S/V/E coated 2024-T3 alloy coupons.*

5.4. Cerium Doped Systems:

Three types of cerium doped films were prepared and characterized. Each film studied was comparable to a film previously described, with the addition of a corrosion-inhibiting component.

Figure 10 shows optical micrographs of the three different cerium doped films. Figure 10(a) shows a typical cerium doped silica

film. At a magnification of 400X, the film can be seen to have a highly fractured surface, similar to that of an undoped system. Crack widths were determined to range from 0.5-4.0 μm.

Figure 10(b) shows a cerium doped S/V/E film. The morphology appears to be indistinguishable from the previous system. At a magnification of 100X, the film can be seen to have large randomly dispersed pinhole with diameter ranging form 10-60 μm and the appearance of film cracking.

Figures 10(c) and 10(d) show a cerium doped silica film with a S/V/E overcoat. Under a magnification of 100X, the ormosil film can be seen to have a significant pinhole density. The pinholes have diameters ranging from 10-60 μm and are randomly dispersed throughout the film. Under a magnification of 400X, the silica film can be seen to have a significant amount of cracking. The crack widths were determined to range from 0.5 to 4.0 μm.

The thickness of the cerium doped silica film was determined by the profilometry plot shown in Figure 11. The scan was recorded over a 1 mm section with a scanning rate of 0.1 mm/s. The first 0.3 mm of the film scan was disregarded due to the defective edge formation. The average film thickness was determined to be 1.6 μm, corresponding to a thickness which was approximately 3-10 times that which would normally be found in crack free silica films.

X-ray photoelectron spectroscopy was used to confirm the presence of cerium in the silica film. Figure 12 shows an XPS spectrum of a cerium doped silica film. Four different elements were detected. Silicon, oxygen, and carbon were found in the dopent free sol-gel silica film(Figure 3). The fourth element, cerium, was identified by peaks at 110eV, 205eV, 500eV, and 605eV. Peak areal analysis indicated the cerium concentration to be 5 atom %.

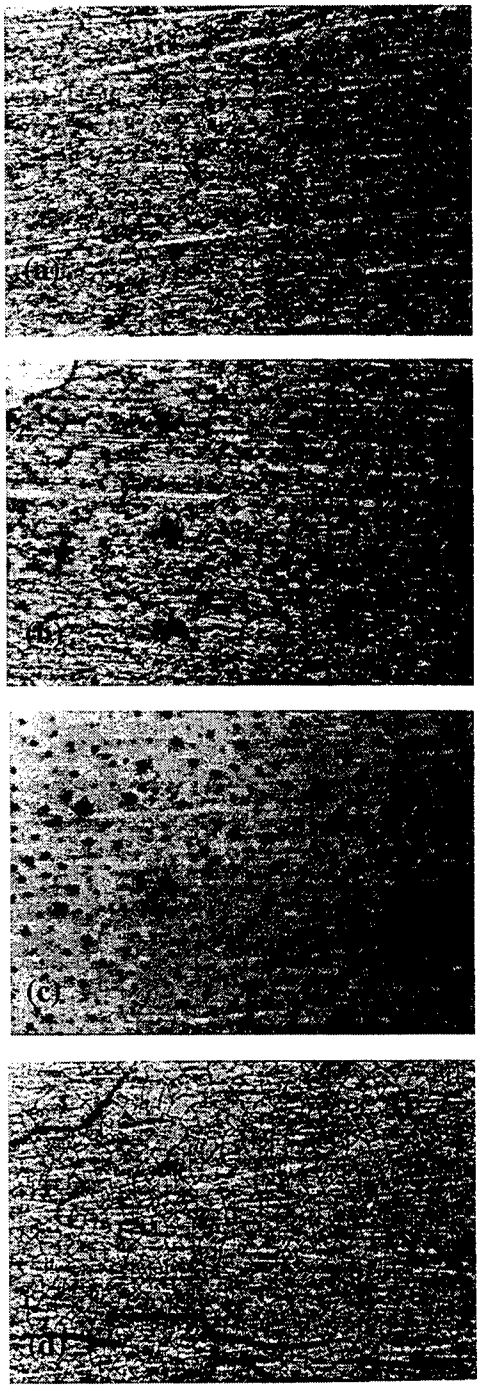


Figure 10. *Optical micrographs of three different cerium doped films on aluminum alloy 2024-T3; (a) cerium-doped silica (magnification = 400X); (b) cerium-doped S/V/E (magnification = 400X); (c) cerium-doped silica overcoated with S/V/E (magnification = 50X); (d) cerium-doped silica overcoated with S/V/E (magnification = 200X).*

Electrochemical measurements of the cerium doped systems are presented in Figures 13, 14, and 15. Figure 13 shows a comparison between the cerium-doped silica, cerium-doped S/V/E (designated silica(Ce) and S/V/E(Ce), respectively), and alodyne coatings. The silica(Ce) film exhibited a CP value of 6.50, which was slightly lower than the value found for undoped silica layers. This result was attributed to pronounced cracking found in the thick cerium-doped silica films. The S/V/E(Ce) film was found to have a CP value of approximately 8.0. This value represents an improvement in corrosion-resistance character of approximately 2 orders of magnitude over undoped S/V/E-type films and 4.5 orders of magnitude in protection over the alodyne 1200 film. The stable potentiodynamic range was also found to be dramatically improved over that of undoped S/V/E films (compare to the S/V/E curve shown in Figure 7).

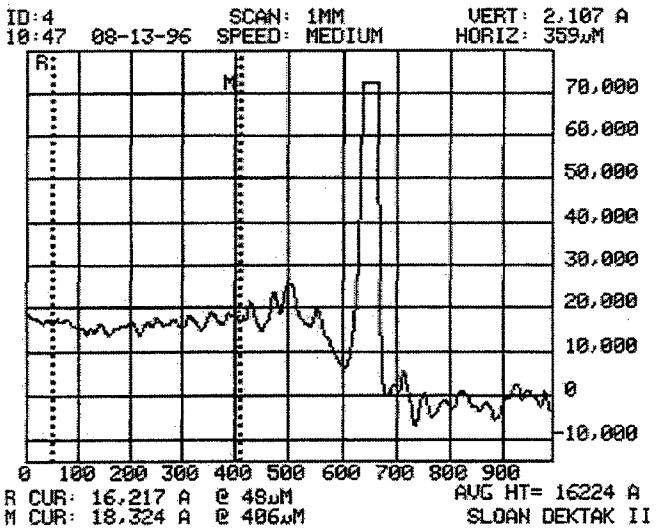


Figure 11. *Profilometry plot of a cerium doped silica film on aluminum alloy 2024-T3.*

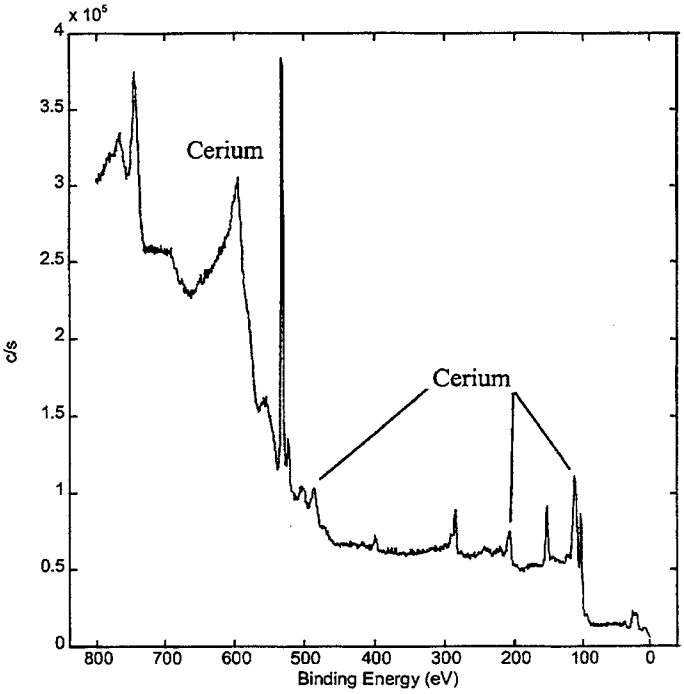


Figure 12. *XPS spectrum of a cerium doped silica film on aluminum alloy 2024-T3*

Figure 14 represents the electrochemical investigation of a sol-gel bilayer, composed of a silica(Ce) subbing layer with S/V/E overcoat, compared with the complete paint system (CPS) used on Air Force aircraft. The CPS is composed of alodyne-1200 surface treatment with an epoxy-based primer and paint top coat(s). After 90 minutes of exposure in a NaCl solution, neither sample showed reliably determinable signs of corrosion. The alodyne 1200 treated coating is not indicated in this figure, due to the comparatively large corrosion current density which yielded a plot that could not be placed on the figure with the indicated scale. However, a

comparison with the CP value found for reference alodyne-treated specimen (see, for example, figure 9) clearly indicates that the silica(Ce) with S/V/E overcoat has nearly 6 orders of magnitude improvement over the corrosion protection character of alodyne 1200.

After 4 days of exposure to a 3% NaCl solution, the coatings began to develop signs of corrosive species penetration. Figure 15 shows a polarization plot of an alodyne 1200 treated specimen, bilayer sol-gel test specimen with silica(Ce) and S/V/E overcoat, and the CPS coating. The alodyne 1200 treated coupon yields a rapid increase in current density, signifying poor corrosion resistance. The CPS curve indicates a CP value of 8.00 without signs of significant degradation over a comparatively wide potential range. The silica(Ce)-S/V/E film has a sharp potentiodynamic knee, corresponding to a CP value of 9.5, followed by a rapid increase in current density to a secondary "plateau" CP value of approximately 3.5. The fluxuation indicates initial failure followed by the formation of a passivating layer that maintains a lower level of corrosion protection. The silica(Ce)with S/V/E overcoat gives an indication of having corrosion protection characteristics that would rival the CPS system if pinholes were eliminated, possessing inhibition performance which is up to six orders of magnitude better than the conventional alodyne 1200 treatment.

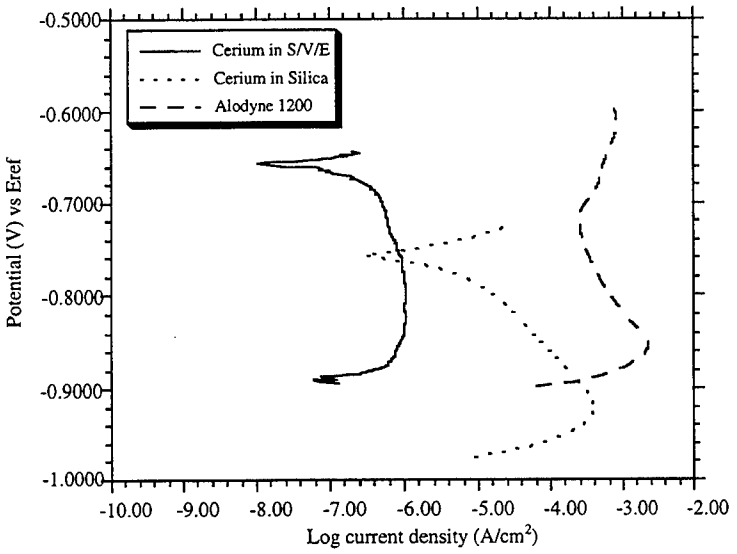


Figure 13. *Polarization curves for alodyne 1200-treated, silica(Ce), and silica(Ce)-S/V/E coated 2024-T3 alloy coupons.*

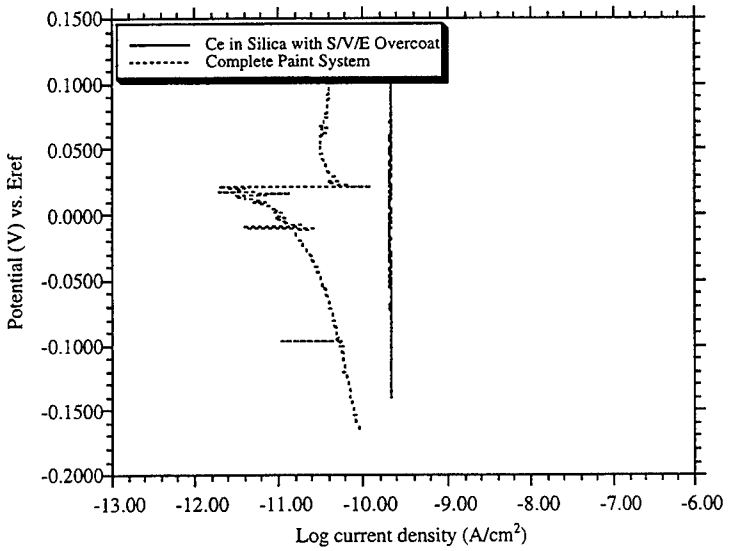


Figure 14. *Polarization curves for silica(Ce)-S/V/E overcoated and complete paint system coated aluminum alloy 2024-T3 coupons after 90 minutes of exposure to a 3% NaCl solution.*

Direct comparison of undoped and cerium-doped ormosil systems shows a dramatic improvement in the CP value and on the extended potentiodynamic stability range. Since both ormosil systems have comparable defect structures, it is reasonable to attribute improved corrosion protection performance to passivation effects derived from the cerium addition. The measured CP values for these films suggest that a defect-free single layer ormosil coating doped with cerium might be capable of providing sustained corrosion protection which is many orders of magnitude better than the current alodyne treatment.

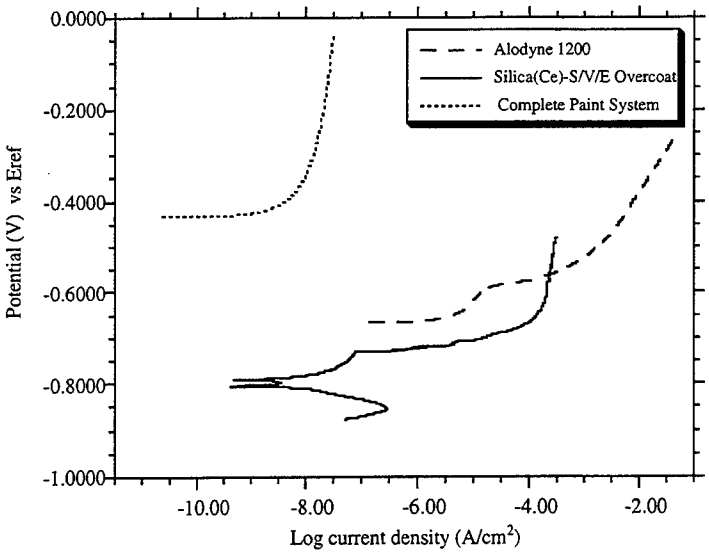


Figure 15. *Polarization curves for alodyne 1200-treated, silica(Ce)-S/V/E overcoated, and complete paint system coated aluminum alloy 2024-T3 coupons after 4 days of exposure to a 3% NaCl solution.*

Also, the comparison between the silica(Ce)-S/V/E multilayer film and the CPS highlights the tremendous promise of bilayer coatings. Even in the presence of significant film defects, the sol-gel bilayer coating was still capable of providing similar protection to that of the CPS and approximately 6 orders of magnitude better than the alodyne 1200 treatment.

6. Summary and Conclusion

The corrosion resistance of aluminum alloy 2024-T3 substrates coated with sol-gel films were assessed. Various mono- and bilayered sol-gel systems were prepared and applied by spin coating. Cerium additives were incorporated into the films as a means to replace chromate species in the passivating conversion coatings. Films were characterized by optical spectroscopy, XPS, and electrochemical analysis.

Results of the summer term project showed that sol-gel coatings provided corrosion protection values which were several orders of magnitude better than the current alodyne 1200 treatments on 2024-T3 coupons. Despite problems with inherent crack and pinhole formation, sol-gel films were up to one million times more resistant to corrosion than the conventional chromate-based surface treatment. Under the most rigorous test conditions investigated, cerium-doped silica-S/V/E ORMOSIL layers were found to be comparable in corrosion resistance character to the current complete multilayer passivation/primer/paint system used in the environmental passivation

of Air Force aircraft. Thus, sol-gel coatings were found to exhibit great potential for the replacement of chromate based conversion coatings presently used in the passivation of aircraft aluminum alloys.

7. References

1 Private Communication, Dr. Michael Donley, Coating Research Program POC, Wright Laboratory WL/MLBT , Wright Patterson Air Force Base, Dayton OH.

2. K. Kato, J. Mater. Sci. **28** (1993) 4033.

3. F.M. Seon, J. Less-Comm. Metals **148** (1989) 73.

4. S. Roure, F. Czerwinski, and A. Petric, Oxid. Met. **42** (1994) 75.

5 V.A.C. Haanappel, T. Fransen, B. Geerdink, and P.J. Gellings, Oxid. Met. **30** (1988) 201.

6. C.J Brinker and G.W. Scherer, in Sol-gel Science, Academic Press, San Diego, (1990)

7. C.J Brinker, D.E. Clark, and D.R. Ulrich, eds., in Better Ceramics Through Chemistry III, Materials Research Society (1988).

8. X.M. Chen, B. Ellis, F. Wang, and A.B. Seddon, J. Non-Cryst. Solids **185** (1995) 1.

9. T. Iwamoto and J.D. Mackenzie, J. Mater. Sci. **30** (1995) 2566.

10. J. Hatch in Aluminum: Properties and Physical Metallurgy, American Society for Metals, Metal Park, (1984)

11. K.A. Lucas and H. Clarke, in Corrosion of Aluminum-based Metal Matrix Composites, Research Studies Press LTD, Somerset, (1993)

12. C.C. Nathan in Corrosion Inhibitors, National Assoc. of Corrosion Eng., Houston (1973).

13 S.M. Cohen, Corrosion **51** (1995) 71.

14 R. Nass, E. Arpac, W. Glaubitt, and H. Schmidt, J. Non-Cryst. Solids **121** (1990) 370.

15. R.M. Guppy and A. Atkinson, in The Cracking Behaviour of Sol-gel Films, 203.

CRACKS AT INTERFACES IN BRITTLE MATRIX COMPOSITES

Michael C. Larson, Assistant Professor
and
Herbert F. Miles, II, Graduate Research Assistant
Department of Mechanical Engineering

Tulane University
New Orleans, LA 70118-5674

Final Report for:
Summer Faculty Research Program and
Graduate Student Research Program
Wright Materials Laboratory

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

and

Ceramics Group
Wright Materials Laboratory

August 1996

CRACKS AT INTERFACES IN BRITTLE MATRIX COMPOSITES

Michael C. Larson, Assistant Professor
and
Herbert F. Miles, II, Graduate Research Assistant
Department of Mechanical Engineering
Tulane University

Abstract

Interfaces play a key role in the toughness of brittle matrix composites. This study is revealing how friction, toughness, and roughness act in concert to determine the extent of interfacial sliding which may occur near the tip of an impinging matrix crack. The study is bolstered by experiments of cracks at frictional, rough interfaces which reveal the three-dimensional nature of the crack-interface interaction. Critical crack angles are measured at interfaces in dual DCDC specimens.

CRACKS AT INTERFACES IN BRITTLE MATRIX COMPOSITES

Michael C. Larson and Herbert F. Miles, II

1.0 INTRODUCTION

The objective of this work is to link a mechanics-based computational modeling technique with a materials-based experimental program aimed at identifying and optimizing toughening mechanisms in ceramic composites. The authors have been engaged in applying the surface integral method to model fractures at interfaces between brittle constituents; researchers at the Wright Materials Laboratory have shown through their experiments that interfaces play the key role in determining the crack resistance in technologically important brittle matrix composites. By coupling these two, we are seeking to refine our mechanistic representation of the interfaces in order to predict composite performance and to design optimal interfaces. In this spirit, two projects were pursued during the research period: a computational examination of the trade-off between toughness, roughness, and friction at interfaces in determining the extent of interfacial sliding and matrix crack spacing, and an experimental investigation into the three-dimensional aspects of cracks at frictional interfaces.

2.0 RESEARCH DESCRIPTION

2.1 Interfacial Sliding at Interfaces which Possess Friction, Toughness, and Roughness

2.1.1 Motivation

In brittle composites, the matrix is assumed to crack under design conditions with interfaces providing a measure of pseudo ductility to preserve the brittle fibers. The length of the interfacial sliding zone at the periphery of a matrix fracture plays an essential role in determining the global toughness of the composite. This is manifest in two ways: the sliding relieves the stress concentration which is transferred to the fibers and hence regulates fiber failure, and the sliding influences matrix fracture spacing which in turn dictates energy dissipation.

2.1.2 Method

A two-dimensional surface integral method, similar to the displacement discontinuity method described by Crouch and Starfield[1], is used to investigate the behavior of a crack intersecting a planar interface possessing toughness, roughness, and friction. The necessary singular fundamental solutions are derived from the elasticity solution of a point force in an infinite medium (Wiles and Curran[2]).

---

1. Crouch, S. L. and Starfield, A. M. *Boundary Element Methods in Solid Mechanics*, George Allen and Unwin, London: 1976
2. Wiles, T. D. and Curran, J. H. *A General 3-D Displacement Discontinuity Method*, 103-10.

The total relative displacement at any point of the cracked surfaces is obtained by summing the effects of a continuous distribution of the fundamental solutions over the entire cracked domain. This procedure leads to the governing singular integral equation

$$\sigma_{ij}(\hat{x}) = \int_{\partial\zeta} \Gamma_{ijk}(\hat{x}, \hat{\zeta})\delta_k(\hat{\zeta})dA \quad , \tag{EQ 1}$$

where σ_{ij} are the stresses at a field point, x is the location of the field point, Γ_{ijk} are the kernel solutions, δ_k are the crack displacement, and ζ is a vector which spans the fracture surfaces (all referenced to a fixed Cartesian frame).

The approximate numerical solution to the governing integral equation is obtained by transforming it to a system of algebraic equations with the nodal displacements as unknown quantities. The opening and sliding displacements are approximated over discrete elements of the fracture surface by interpolation functions which depend upon values at specified nodal points. The crack domain is divided into E discrete elements, i.e.,

$$\sigma_{ij}(\hat{x}) = \sum_{e=1}^{E} \int_{\partial\zeta} \Gamma_{ijk}(\hat{x}, \hat{\zeta})\delta_k^{(e)}(\hat{\zeta})dA \quad . \tag{EQ 2}$$

Crack displacements, δ, are calculated by boundary collocation. Interpolation functions relate an assumed displacement variation to the nodal values:

$$\delta_k^{(e)}(\hat{\zeta}) = \sum_{n=1}^{N} h(\hat{\zeta})^{(n)}\hat{\delta}_k^{(n)} \quad . \tag{EQ 3}$$

Constant displacement interpolations are used for the majority of the elements which compose the fractures in order to reduce the computational time needed per simulation. Elements with specialized high order interpolations are necessary, however, for near-tip regions having large displacement gradients. This higher order interpolation is based upon the first two terms of the asymptotic elastic crack displacement field governed by the stress intensity factor, K:

$$\delta_i = K_i 4\frac{(1-\nu^2)}{E\sqrt{2\pi}}\rho^{1/2} + B_i(\nu, E, n)\rho^{3/2} \ , \tag{EQ 4}$$

with ρ being the distance from the crack front and the subscript i is not a summation index but rather corresponds to the fracture mode, i = I or II.

Gauss quadrature is used to numerically integrate the distribution of Green's functions for the two-dimensional displacement discontinuities in a homogeneous, isotropic, linear elastic material. For frictional sliding, the normal stress is compared to the shear local to each element. If the local normal traction is tensile, then the crack elements open and are free to slide; if the current local normal traction is compressive, and the shearing traction exceeds the normal traction multiplied by a friction coefficient, then the element is closed and slides against a restraining shearing traction. The slip length is defined as the length that the crack has traveled up the interface.

2.1.3 Results

To reveal the near-tip behavior of interfacial sliding, that will not be influenced by the finite geometry of the matrix crack, we limit the absolute sliding length to be less than 8% of the half crack length. This rule of thumb was determined by keeping the difference between the numerically determined stress field around the crack tip within 5% of the analytical asymptotic stress field. The agreement with results presented in Dollar and Steif[3] for the relationship between the normalized slip length and the Coulomb coefficient of friction are shown in Figure 1. Roughness is assumed to be approximated by a triangular arrangement of surface integral elements with a period, P, and amplitude, A. Good correlation was found in the results using the same roughness angle, ϕ, defined as

$$\phi = \operatorname{atan}\frac{2A}{P} \qquad \text{(EQ 5)}$$

as long as the period of roughness was no more than ten percent of the sliding length.

Preliminary results show that, for certain interfaces with no inherent strength (i.e., no toughness), increases in the roughness angle may lead to longer sliding lengths, as can be seen in Figure 2. More analysis must be done to confirm if this behavior is actually expected or if it is a numerical artifact. One important note about these results is that the sliding zones are not continuous, closed sliding, but rather are interspersed with open and sticking portions. Sample sliding lengths as a function of toughness, friction, and roughness are displayed in Figure 3.

2.1.4 Future Work

We will proceed with parametric studies aimed at both refining the results and determining the applicable ranges for the various parameters. We will address the issues of convergence, accuracy, and

3. Dollar, A., and Steif, P. S., "A Tension Crack Impinging Upon Frictional Interfaces." *Journal of Applied Mechanics* **56** 291-8 (1989).

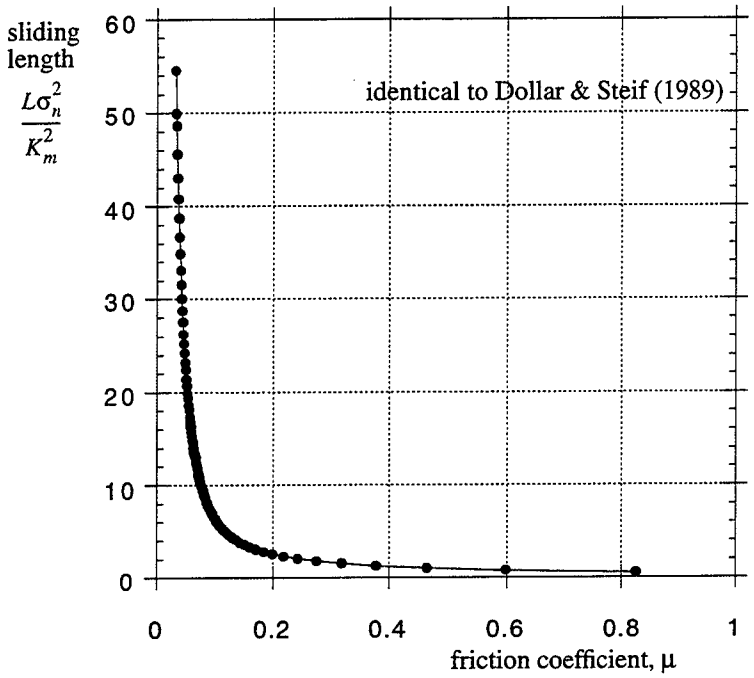


Figure 1. Effect of friction on sliding length of interfacial toughness.

uniqueness for the iterative scheme employed under conditions of monotonic, quasi-static loading. These will be pursued with regard to element sizes, numerical integration schemes, loading increments, etc.

2.2 Experiments of Three-Dimensional Cracks at Frictional Interfaces

2.2.1 Motivation

The goal of this research effort will be to measure experimentally how different interfacial surface roughness, coupled with variable friction and bonding, act to limit the crack opening displacement of an impinging fracture. These measurements will be compared with computational predictions derived from the two-dimensional and three dimensional surface integral method. If the results can indeed be replicated, it will bolster confidence in using the computational tools to give design guidance for interfaces. In addition, we seek to determine the steady-state angle at which a three dimensional crack front propagates when it intersects a rough, frictional interface. This quantity will be useful as a parameter to embed into computational simulations where there are often numerical problems associated with surfaces.

2.2.2 Method

The dual double cleavage drilled compression (dual-DCDC) specimens used in this experiment, a variation on the geometry used by Janssen[4], were fabricated from plate glass. The dimensions of both the

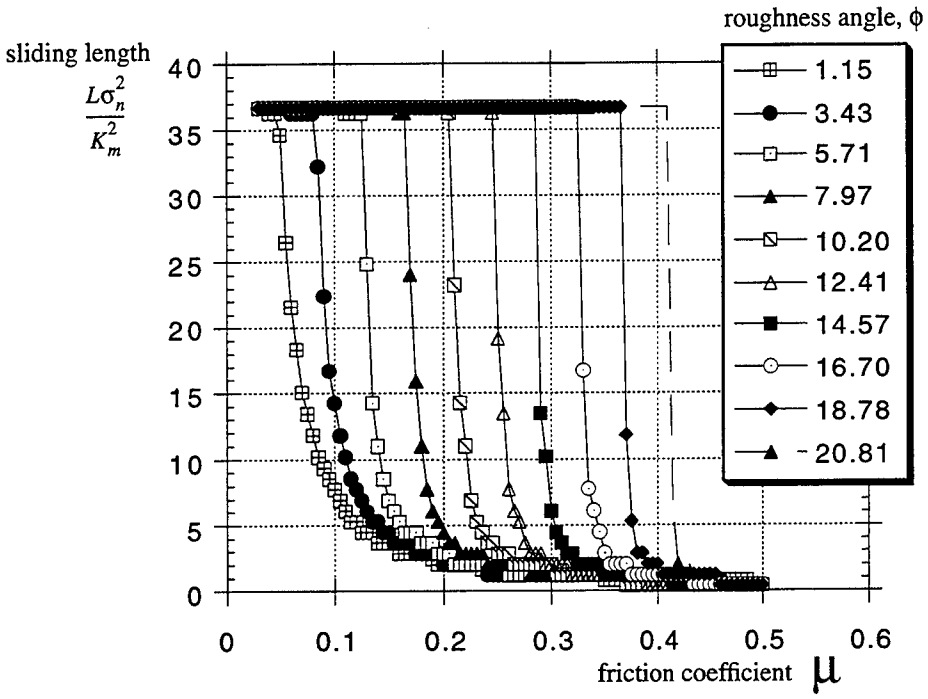


Figure 2. Interfacial sliding length variation for combinations of interfacial friction and roughness for interfaces with no toughness.

drilled and undrilled halves were 2.54 cm X 10.16 cm X 0.9525 cm. The crack starter holes, drilled in the center of the 0.9525 cm X 10.16 cm faces, have a diameter of 0.3175 cm. A half containing a hole was butted up against a non-drilled specimen with the sides in contact possessing a known roughness. The pre-stress normal to the interface was controlled using a hydraulic piston assembly that pressed the two specimens together. The dual-DCDCs were compressed in an MTS Sintech 20/G machine.

As the specimen is compressed, a crack initiates at the drilled hole due to the stress concentration located there. The stress concentration is similar to that of a remote stress applied to a hole in an infinite sheet (Warren[5]). Due to the specimen geometry, a stable crack can be propagated perpendicular to the

4. Janssen, C., "Specimen for Fracture Mechanics Studies on Glass," *Proceedings of the Tenth International Congress on Glass*, The Ceramic Society of Japan, 23-30 (1974).
5. Warren, W. E., "Theoretical Analysis of the Double Cleavage Drilled Compression Specimen," *International Journal of Fracture*, **33** 223-35 (1987).

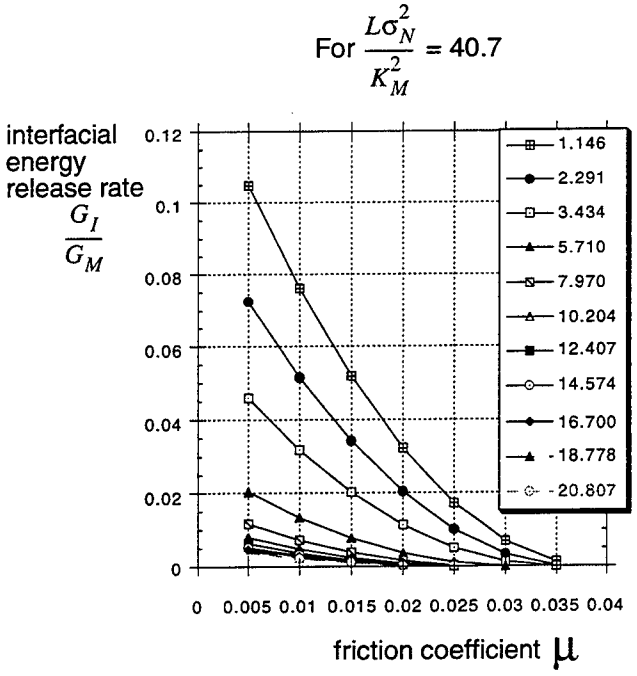


Figure 3. Sample sliding length as a function of toughness, friction, & roughness.

controlled interface. Because of the brittle nature of glass, we assume that local crack extension is at a constant fracture toughness and any variation in the crack front shape as it traverses the specimen is a consequence of the roughness and friction at the interface (see Figure 4).

While the cracks are propagating, the crack opening displacements are determined using crack opening interferometry (COI).[6] COI is based on the concept that when light is sent through two opposing surfaces the reflected beams interfere with one another creating interference fringes along the specimen. These fringes can then be used to determine the amount that a crack is opening or closing. One can represent the combination of the two beams by adding the two light vector magnitudes together where each light vector is represented by

$$E = a\cos(\phi - \omega t) \qquad \text{(EQ 6)}$$

and their combination is

6. Leichti, K. M., "On The Use of Classical Interferometry Techniques in Fracture Mechanics," in *Experimental Techniques in Fracture*, 95-124 (1994).

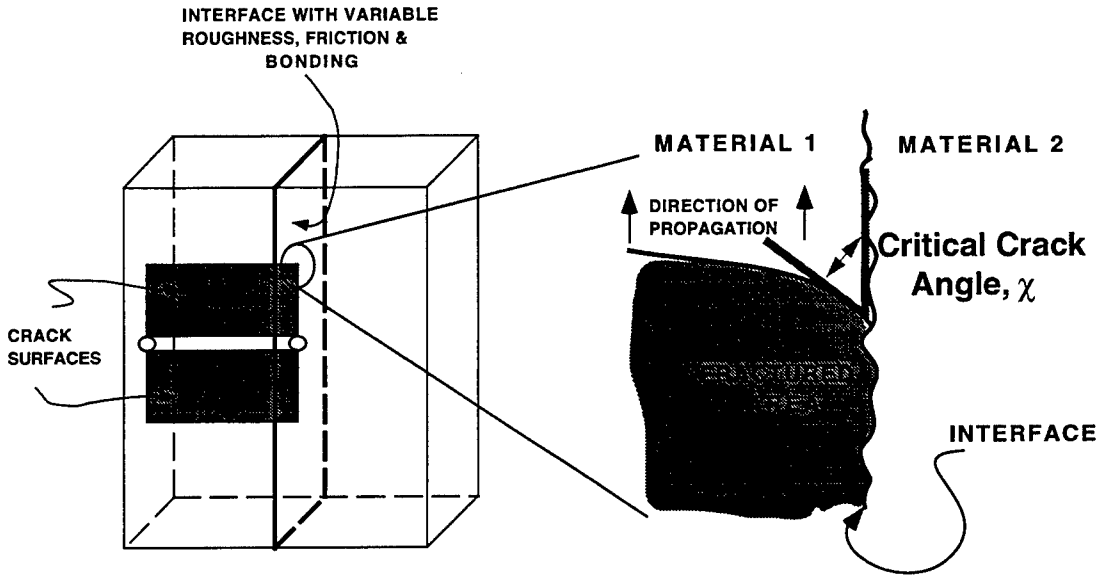


Figure 4. Dual DCDC specimen showing the crack bowing angle χ at the frictional, rough interface.

$$E_{total} = \bar{a}\cos\left(\frac{1}{2}(\phi_1 + \phi_2) - \omega t\right) \quad \text{(EQ 7)}$$

where

$$\bar{a} = a\sqrt{2[1 + \cos(\phi_2 - \phi_1)]} \quad \text{(EQ 8)}$$

is the amplitude of the resulting wave. The phase angle difference is related to the displacement δ by

$$\phi_2 - \phi_1 = 2\pi\frac{\delta}{\lambda} \quad \text{(EQ 9)}$$

Subbing into the previous equation and knowing that the intensity is proportional to the square of the amplitude, one gets

$$I \approx 4a^2\left(\cos\left(\pi\frac{\delta}{\lambda}\right)\right)^2 \quad \text{(EQ 10)}$$

It can be seen from this equation that the dark fringes occur at where the intensity is a minimum (I=0) and the light fringes occur where the intensity is at a maximum.

$$I_{min} = 0 \Rightarrow \delta = (2n+1)\frac{\lambda}{2} \quad \text{(EQ 11)}$$

$$I_{max} \Rightarrow \delta = n\lambda \quad \text{(EQ 12)}$$

Thus, destructive interference, or dark fringes, can be observed where

$$h = \frac{n\lambda}{2n_2 \cos\gamma} \quad \text{(EQ 13)}$$

and constructive interference, or bright fringes, can be seen where

$$h = \frac{(2n-1)\lambda}{4n_2 \cos\gamma} \quad \text{(EQ 14)}$$

where n_2 is the refractive index of the fluid between the crack surfaces. For these experiments, the fluid is air and the refractive index is equal to 1.003. For a wavelength λ the resolution is equal to $1/4\lambda$.

The mode I stress intensity factor for any fringe can then be found from the normal crack opening displacements by

$$K_I = \frac{n\lambda}{16n_2 \cos\gamma}\left(\frac{E}{(1-\upsilon^2)}\right)\left(\frac{2\pi}{r}\right)^{\frac{1}{2}} \quad \text{(EQ 15)}$$

2.2.3 Preliminary Results

All of the DCDC specimens were loaded to around 3600 pounds with pop-in occurring between 2800-3300 pounds. For the three sets of roughness (polished, 10 micron, 30 micron), a range of lateral compression was applied (from 0-500 psi). The rougher the interface between a set of specimens, the more slippery the materials appeared. The crack bowed at a much lower lateral compression for the polished specimens than the abraded specimens. For all specimens that exhibited this crack pinning nature, a steady state angle between the crack front and the interface was observed. The higher the roughness, the smaller the angle between the crack front and interface was. The angle also seemed to depend on the amount of lateral compression.Crack kinking was also observed if the specimens were put under a Mode II as well as a Mode I loading. These specimens exhibited growth out of plane and a significant pinning at the interface. All of

these specimens failed at a significantly lower load than the pure Mode I loading, however significant crack retardation was achieved.

	50 psi	**60 psi**	**100 psi**	**197 psi**	**500 psi**
Polished	63-64.1	61.8-62.5	N/A	57.4-60	
12 micron		64.6-67	39.6-44	46.2	
30 micron		67	67	67	34-39.1

TABLE 1. Measured crack angles, χ, as a function of the nominal normal pressure at the interface and the roughness.

2.2.4 Future Work

We are setting about to gather additional data so as to draw more general conclusions from the tests and to compare with three-dimensional computer simulations. In addition, we will seek to confirm and quantify the repeatability of the tests.

ANALYSIS AND DESIGN OF GAIN SCHEDULED MISSILE AUTOPILOTS

Douglas A. Lawrence
Associate Professor
School of Electrical Engineering
and Computer Science

Ohio University
Athens, Ohio 45701

Final Report for:

Summer Faculty Research Program
Wright Laboratory
Armament Directorate
Eglin Air Force Base

Sponsored by:

Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Wright Laboratory
Armament Directorate
Eglin Air Force Base

August 1996

ANALYSIS AND DESIGN OF GAIN SCHEDULED MISSILE AUTOPILOTS

Douglas A. Lawrence
Associate Professor
School of Electrical Engineering
and Computer Science
Ohio University

Abstract

The application of a recently developed methodology for gain scheduled control system synthesis to the design of a pitch channel missile autopilot is described. Conditions are given under which a nonlinear gain scheduled autopilot exists that linearizes to a prescribed family of linear autopilots designed over a wide operating envelope. Based on this, a nonlinear gain scheduled autopilot is constructed that makes direct use of aerodynamic missile data in tabular form. Nonlinear simulation results indicate satisfactory performance over the flight envelope.

ANALYSIS AND DESIGN OF GAIN SCHEDULED MISSILE AUTOPILOTS

Douglas A. Lawrence

1. Introduction

Controlling the behavior of a missile or aircraft over a wide operating envelope is a challenging problem in large part due to the presence of significant nonlinear effects. Existing flight control systems typically employ some form of gain scheduling to compensate for these nonlinearities and yet preserve a linear controller architecture. Documented studies of gain scheduling methods applied to flight control problems have appeared in the literature. [2] [5] [8] [9] [11] [13] [15]

Recently there have been numerous theoretical studies on gain scheduling. Controller design for linear parameter-varying (LPV) systems is currently receiving a great deal of attention. [1] [10] [12] [14] Since it is apparent that gain scheduled control systems are nonlinear, recent work has approached the gain scheduling design process from a nonlinear systems viewpoint. [3] [7]

The gain scheduling methodology proposed in [7] can be described in broad terms by the following steps:

- Determine a nonlinear plant's family of constant operating points, parameterized by a designated set of scheduling variables.
- For the corresponding family of linearized plants, design a family of linear controllers to meet prescribed design goals at each constant operating point.
- Construct a gain scheduled controller that linearizes to the appropriate linear controller at each constant operating point.
- Check nonlocal performance of the resulting nonlinear control system.

In this methodology, the emphasis is on the linearization requirement of the third step. Specifically, technical conditions are derived for the existence of a nonlinear controller with this property. Further, it is shown that failure to satisfy these conditions will introduce so-called *hidden coupling terms* in any attempt to schedule the family of linear controllers constructed in the second step. These hidden coupling terms can potentially degrade system performance.

This report describes the application of the above gain scheduling methodology to the design of a pitch channel autopilot for a hypothetical tail-controlled missile. A primary objective of this exercise is to identify any special features of the design process that might be applicable to missile autopilot design in general. The remainder of this report is organized as follows. The nonlinear missile model is presented in the next section. The gain scheduled autopilot design is described in Section 3. Simulation results are

discussed in Section 4. Finally, concluding remarks and suggestions for future investigation are given in Section 5.

2. Missile Description

The missile's longitudinal behavior is characterized by the following variables:

V_m	missile speed in *meters/sec*
α	angle-of-attack in *radians*
q	pitch rate in *radians/sec*
δ	tail fin deflection in *radians*
θ	pitch attitude angle in *radians*
η_z	normal acceleration in *meters/sec*^{2}
h	altitude in *meters.*
M	mach number.

The short-period longitudinal aerodynamics are described by

$$\begin{aligned} \dot{\alpha} &= \frac{1}{V_m}\left[-\frac{\bar{q}S}{m}C_L[h,M,\alpha,\delta] + g\cos(\theta-\alpha)\right] + q \\ \dot{q} &= \frac{\bar{q}Sd}{I_{yy}}C_M[h,M,\alpha,q,\delta] \end{aligned} \tag{1}$$

and the remaining longitudinal aerodynamics are described by

$$\begin{aligned} \dot{V}_m &= -\frac{\bar{q}S}{m}C_D[h,M,\alpha,\delta] - g\sin(\theta-\alpha) \\ \dot{\theta} &= q \\ \dot{h} &= V_m\sin(\theta-\alpha). \end{aligned} \tag{2}$$

Normal acceleration is given by

$$\eta_z = \frac{\bar{q}S}{m}C_Z[M,\alpha,\delta] + g\cos(\theta). \tag{3}$$

The aerodynamic coefficients C_L, C_D, C_M, and C_Z characterize the aerodynamic lift force, drag force, pitching moment, and body normal force, respectively, of the missile. Also, dynamic pressure, $\bar{q}$, is determined from

$$\bar{q} = \frac{1}{2}\rho V_m^2 \tag{4}$$

where ρ is the altitude-dependent air density. The tail fin actuator is modeled by

$$\frac{d}{dt}\begin{bmatrix}\delta \\ \dot{\delta}\end{bmatrix} = \begin{bmatrix}0 & 1 \\ -\omega_a^2 & -2\zeta_a\omega_a\end{bmatrix}\begin{bmatrix}\delta \\ \dot{\delta}\end{bmatrix} + \begin{bmatrix}0 \\ \omega_a^2\end{bmatrix}\delta_c \tag{5}$$

where δ_c is the commanded tail fin deflection in *radians.* Additional missile and actuator parameters are listed in Table 1.

3. Gain Scheduled Autopilot Design

The first step in the design process involves calculating the missile's family of constant operating points or equilibria. This poses an immediate difficulty in that true equilibria cannot exist for a ballistic missile since the condition $\dot{V}_m = 0$ cannot be achieved without thrust. Even if this is ignored and constant airspeed is assumed, achievable equilibria for the remaining variables correspond to small normal acceleration values. This is not desirable since any attempt to track relatively large normal acceleration commands would result in correspondingly large deviations from such constant operating points, thus violating the underlying assumption upon which a linear approximation is based. As an alternative, we consider only short-period equilibria. Specifically, we are interested in constant solutions to the algebraic equations resulting from the constraints $\dot{\alpha} = \dot{q} = 0$:

$$\begin{aligned} 0 &= \frac{1}{V_m{}^\circ}\left[-\frac{\bar{q}^\circ S}{m}C_L[h^\circ, M^\circ \alpha^\circ, \delta^\circ] + g\cos(\theta^\circ - \alpha^\circ)\right] + q^\circ \\ 0 &= \frac{\bar{q}^\circ Sd}{I_{yy}}C_M[h^\circ, M^\circ, \alpha^\circ, q^\circ, \delta^\circ]. \end{aligned} \tag{6}$$

For any such solution, the corresponding normal acceleration value is

$$\eta_z{}^\circ = \frac{\bar{q}^\circ S}{m}C_Z[M^\circ, \alpha^\circ, \delta^\circ] + g\cos(\theta^\circ). \tag{7}$$

We note that solutions to (6) will not yield zero derivatives in (2). We therefore refer to solutions to (6) as *quasi-static operating points* to reflect the fact that these are not equilibrium points for the overall longitudinal aerodynamics. This is somewhat worrisome since in the next step we propose to design linear autopilots on the basis of linearizations computed about such operating points. To see the potential danger in performing linear designs about non-equilibria, consider the simple first-order system

$$\dot{x} = f(x,u) = x^2 + 2x + 2 + u \tag{8}$$

and the operating point $(x^\circ, u^\circ) = (0,0)$, which is not an equilibrium. Nevertheless, on the basis of

$$A = \frac{\partial f}{\partial x}(x^\circ, u^\circ) = 2 \qquad B = \frac{\partial f}{\partial u}(x^\circ, u^\circ) = 1 \tag{9}$$

we select the linear control law $u = -4x$ which yields the closed-loop system

$$\dot{x} = f_{cl}(x) = x^2 - 2x + 2 \tag{10}$$

having

$$A_{cl} = \frac{\partial f_{cl}}{\partial x}(x^\circ) = -2. \tag{11}$$

One then might expect stable closed-loop behavior in a vicinity of $x^\circ = 0$. However for $x(0) = x^\circ = 0$ the closed-loop trajectory is given by $x(t) = 1 + \tan(t - \pi/4)$ which escapes to infinity in finite time.

Acknowledging the potential hazards of our approach, we continue with the design and recognize that, since (6) constitutes two equations in six unknowns, there is a continuous family of solutions which can be parameterized in several ways. Here we parameterize by altitude, mach number, angle-of-attack, and pitch attitude angle and determine pitch rate, tail fin deflection, and normal acceleration as functions of these variables to obtain the relationships

$$\begin{aligned} q^\circ &= q^\circ[h^\circ, M^\circ, \alpha^\circ, \theta^\circ] \\ \delta^\circ &= \delta^\circ[h^\circ, M^\circ, \alpha^\circ, \theta^\circ] \\ \eta_z{}^\circ &= \eta_z{}^\circ[h^\circ, M^\circ, \alpha^\circ, \theta^\circ]. \end{aligned} \tag{12}$$

In addition, the actuator dynamics yield

$$\delta_c{}^\circ = \delta^\circ[h^\circ, M^\circ, \alpha^\circ, \theta^\circ]. \tag{13}$$

For the second step in the design process, we focus on the short period aerodynamics (1), the actuator dynamics (5), and normal acceleration (3) for which (12) and (13) characterize a parameterized family of constant operating points. For the associated parameterized family of linearizations, we consider a parameterized family of linear autopilots of the form

$$\begin{aligned} \dot{x}_I &= \eta_c - \eta_z \\ \delta_c &= K_1(h^\circ, M^\circ, \alpha^\circ)\, x_I - K_2(h^\circ, M^\circ, \alpha^\circ)\, \eta_z - K_3(h^\circ, M^\circ, \alpha^\circ)\, q \\ &\qquad - K_4(h^\circ, M^\circ, \alpha^\circ)\, \delta \;\; - K_5(h^\circ, M^\circ, \alpha^\circ)\, \dot{\delta} \end{aligned} \tag{14}$$

where η_c is a commanded normal acceleration. The scheduled gains were generated by McDonnell Douglas Aerospace via their proprietary *AUTOGAIN* software at the design points characterized by the altitude, mach number, and angle-of-attack values listed in Table 2. As an indication of the performance and stability robustness achieved by these linear designs we consider the operating point corresponding to 10,000 *meter* altitude, mach 3, 20 *degrees* angle-of-attack, and 20 *degrees* pitch attitude. For the associated linearized short-period aerodynamics, actuator dynamics, and linear autopilot a closed-loop unit step response is given in Fig. 1 and an open-loop bode plot, with the loop open between the autopilot and actuator, is given in Fig 2. These plots indicate desirable transient response behavior and excellent classical stability margins and are representative of the linear performance achieved over the entire operating point envelope.

The next step in the design process is to construct a nonlinear autopilot of the form

$$\begin{aligned} \dot{x}_I &= \eta_c - \eta_z \\ \delta_c &= c(x_I, \eta_z, q, \delta, \dot{\delta}, h, M, \alpha, \theta) \end{aligned} \tag{15}$$

that schedules the family of linear autopilots in the sense that the following two requirements are met.

Requirement 1. At each quasi-static operating point, the nonlinear autopilot generates the correct constant control value. Specifically, there must exist a function $x_I{}^\circ[h, M, \alpha, \theta]$ that satisfies (suppressing arguments)

$$c(x_I{}^\circ, \eta_z{}^\circ, q^\circ, \delta^\circ, 0, h^\circ, M^\circ, \alpha^\circ, \theta^\circ) = \delta^\circ. \tag{16}$$

Requirement 2. At each quasi-static operating point, the nonlinear autopilot linearizes to the appropriate linear autopilot. Specifically, the following partial derivative identities must hold.

$$\begin{aligned}
\frac{\partial c}{\partial x_I}(x_I{}^\circ,\eta_z{}^\circ,q^\circ,\delta^\circ,0,h^\circ,M^\circ,\alpha^\circ,\theta^\circ) &= K_1(h^\circ,M^\circ,\alpha^\circ)\\
\frac{\partial c}{\partial \eta_z}(x_I{}^\circ,\eta_z{}^\circ,q^\circ,\delta^\circ,0,h^\circ,M^\circ,\alpha^\circ,\theta^\circ) &= -K_2(h^\circ,M^\circ,\alpha^\circ)\\
\frac{\partial c}{\partial q}(x_I{}^\circ,\eta_z{}^\circ,q^\circ,\delta^\circ,0,h^\circ,M^\circ,\alpha^\circ,\theta^\circ) &= -K_3(h^\circ,M^\circ,\alpha^\circ)\\
\frac{\partial c}{\partial \delta}(x_I{}^\circ,\eta_z{}^\circ,q^\circ,\delta^\circ,0,h^\circ,M^\circ,\alpha^\circ,\theta^\circ) &= -K_4(h^\circ,M^\circ,\alpha^\circ)\\
\frac{\partial c}{\partial \dot{\delta}}(x_I{}^\circ,\eta_z{}^\circ,q^\circ,\delta^\circ,0,h^\circ,M^\circ,\alpha^\circ,\theta^\circ) &= -K_5(h^\circ,M^\circ,\alpha^\circ)
\end{aligned} \tag{17}$$

and

$$\begin{aligned}
\frac{\partial c}{\partial h}(x_I{}^\circ,\eta_z{}^\circ,q^\circ,\delta^\circ,0,h^\circ,M^\circ,\alpha^\circ,\theta^\circ) &= 0\\
\frac{\partial c}{\partial M}(x_I{}^\circ,\eta_z{}^\circ,q^\circ,\delta^\circ,0,h^\circ,M^\circ,\alpha^\circ,\theta^\circ) &= 0\\
\frac{\partial c}{\partial \alpha}(x_I{}^\circ,\eta_z{}^\circ,q^\circ,\delta^\circ,0,h^\circ,M^\circ,\alpha^\circ,\theta^\circ) &= 0\\
\frac{\partial c}{\partial \theta}(x_I{}^\circ,\eta_z{}^\circ,q^\circ,\delta^\circ,0,h^\circ,M^\circ,\alpha^\circ,\theta^\circ) &= 0.
\end{aligned} \tag{18}$$

The identities in (18) reflect the fact that the family of linear autopilots does not exhibit linear terms in the variables that parameterize the linear gains and the quasi-static operating points. The following theorem establishes necessary and sufficient conditions for the existence of a nonlinear autopilot that meets these two requirements.

Theorem. There exists a nonlinear autopilot (15) that meets Requirements 1 and 2 if and only if there exists a function $x_I{}^\circ[h,M,\alpha,\theta]$ that satisfies

$$\begin{aligned}
\frac{\partial \delta^\circ}{\partial h} &= K_1(h^\circ,M^\circ,\alpha^\circ)\frac{\partial x_I{}^\circ}{\partial h} - K_2(h^\circ,M^\circ,\alpha^\circ)\frac{\partial \eta_z{}^\circ}{\partial h} - K_3(h^\circ,M^\circ,\alpha^\circ)\frac{\partial q^\circ}{\partial h} - K_4(h^\circ,M^\circ,\alpha^\circ)\frac{\partial \delta^\circ}{\partial h}\\
\frac{\partial \delta^\circ}{\partial M} &= K_1(h^\circ,M^\circ,\alpha^\circ)\frac{\partial x_I{}^\circ}{\partial M} - K_2(h^\circ,M^\circ,\alpha^\circ)\frac{\partial \eta_z{}^\circ}{\partial M} - K_3(h^\circ,M^\circ,\alpha^\circ)\frac{\partial q^\circ}{\partial M} - K_4(h^\circ,M^\circ,\alpha^\circ)\frac{\partial \delta^\circ}{\partial M}\\
\frac{\partial \delta^\circ}{\partial \alpha} &= K_1(h^\circ,M^\circ,\alpha^\circ)\frac{\partial x_I{}^\circ}{\partial \alpha} - K_2(h^\circ,M^\circ,\alpha^\circ)\frac{\partial \eta_z{}^\circ}{\partial \alpha} - K_3(h^\circ,M^\circ,\alpha^\circ)\frac{\partial q^\circ}{\partial \alpha} - K_4(h^\circ,M^\circ,\alpha^\circ)\frac{\partial \delta^\circ}{\partial \alpha}\\
\frac{\partial \delta^\circ}{\partial \theta} &= K_1(h^\circ,M^\circ,\alpha^\circ)\frac{\partial x_I{}^\circ}{\partial \theta} - K_2(h^\circ,M^\circ,\alpha^\circ)\frac{\partial \eta_z{}^\circ}{\partial \theta} - K_3(h^\circ,M^\circ,\alpha^\circ)\frac{\partial q^\circ}{\partial \theta} - K_4(h^\circ,M^\circ,\alpha^\circ)\frac{\partial \delta^\circ}{\partial \theta}.
\end{aligned} \tag{19}$$

Proof. For necessity, suppose there is a nonlinear autopilot that satisfies Requirements 1 and 2. Then differentiating (16) with respect to h, M, α, and θ and substituting the identities in (17) and (18) yields (19). For sufficiency, suppose there is a function $x_I{}^\circ[h,M,\alpha,\theta]$ that satisfies (19) and consider

$$
\begin{aligned}
c(x_I,\eta_z,q,\delta,\dot{\delta},h,M,\alpha,\theta) \;=\; & \delta^\circ[h,M,\alpha,\theta] + K_1(h,M,\alpha)\left[x_I - x_I{}^\circ[h,M,\alpha,\theta]\right] \\
& - K_2(h,M,\alpha)\left[\eta_z - \eta_z{}^\circ[h,M,\alpha,\theta]\right] - K_3(h,M,\alpha)\left[q - q^\circ[h,M,\alpha,\theta]\right] \\
& - K_4(h,M,\alpha)\left[\delta - \delta^\circ[h,M,\alpha,\theta]\right] - K_5(h,M,\alpha)\,\dot{\delta}.
\end{aligned}
\tag{20}
$$

It is easy to see that (16) and (17) hold for any choice of $x_I{}^\circ[h,M,\alpha,\theta]$ and it remains to check that (18) is satisfied. Differentiating (20) with respect to h and evaluating at a quasi-static operating point yields

$$
\begin{aligned}
&\frac{\partial c}{\partial h}(x_I{}^\circ,\eta_z{}^\circ,q^\circ,\delta^\circ,0,h^\circ,M^\circ,\alpha^\circ,\theta^\circ) \;= \\
&\frac{\partial \delta^\circ}{\partial h} - K_1(h^\circ,M^\circ,\alpha^\circ)\frac{\partial x_I{}^\circ}{\partial h} + K_2(h^\circ,M^\circ,\alpha^\circ)\frac{\partial \eta_z{}^\circ}{\partial h} + K_3(h^\circ,M^\circ,\alpha^\circ)\frac{\partial q^\circ}{\partial h} + K_4(h^\circ,M^\circ,\alpha^\circ)\frac{\partial \delta^\circ}{\partial h}
\end{aligned}
\tag{21}
$$

which is zero due to the first identity in (19). The remaining identities in (18) can be verified in the same way.

For the particular autopilot given by (20), each identity in (18) is satisfied if and only if the corresponding identity in (19) holds. Failure to satisfy any one of the identities in (19) will result in a so-called *hidden coupling term* with respect to the associated variable. Unfortunately, (19) constitutes a set of partial differential equations in the unknown function $x_I{}^\circ[h,M,\alpha,\theta]$ for which a solution will typically not exist. It is always possible, however, to satisfy one of the identities in (19). For example, since the linear autopilot family was designed based on a short-period approximation of the longitudinal aerodynamics and a hidden coupling term with respect to angle-of-attack would therefore seem to be the most troublesome, the third identity in (18) can be satisfied by taking

$$
\begin{aligned}
\frac{\partial x_I{}^\circ}{\partial \alpha}[h,M,\alpha,\theta] \;&=\; \frac{1}{K_1(h,M,\alpha)}\left[K_2(h,M,\alpha)\frac{\partial \eta_z{}^\circ}{\partial \alpha} + K_3(h,M,\alpha)\frac{\partial q^\circ}{\partial \alpha} + \left(1+K_4(h,M,\alpha)\right)\frac{\partial \delta^\circ}{\partial \alpha}\right] \\
x_I{}^\circ[h,M,\alpha,\theta] \;&=\; \int_0^\alpha \frac{\partial x_I{}^\circ}{\partial \alpha}[h,M,\xi,\theta]\,d\xi.
\end{aligned}
\tag{22}
$$

In order to implement the nonlinear gain scheduled autopilot described by (15) and (20), the functions in (12) and (22) must be computed. Closed-form expressions are extremely difficult, if not impossible, to obtain largely because the aerodynamic coefficients are represented in tabular form. Consequently, the following numerical approach was taken:

1. To simplify matters, altitude was fixed at 10,000 *meters* and, since solutions to (6) were only weakly dependent on pitch attitude, θ was set to 0 *radians.*
2. On an $\alpha \times M$ grid, (6) was solved numerically for q° and δ° and $\eta_z{}^\circ$ was set according to (7).
3. Splines with respect to α were constructed for $1/K_1, K_2, K_3, (1+K_4), \eta_z{}^\circ, q^\circ, \delta^\circ$ for each discrete M value.

4. A spline with respect to α was constructed for $\partial x_I{}^\circ/\partial\alpha$ according to the first equation in (22) for each discrete M value.
5. A spline with respect to α was constructed for $x_I{}^\circ$ according to the second equation in (22) for each discrete M value.
6. Two-dimensional look-up tables were constructed for $x_I{}^\circ, \eta_z{}^\circ, q^\circ, \delta^\circ$ for implementation in the nonlinear simulation.

The use of splines greatly facilitated the differentiation and integration operations called for in (22). Also, this procedure can be repeated for a discrete set of altitude and/or pitch attitude values to arrive at higher dimensional look-up tables for the required functions. In summary, the gain scheduled autopilot implemented in the nonlinear simulation is given by (15) and (20) where the autopilot gains are implemented as three-dimensional look-up tables in altitude, mach number, and angle-of-attack and the functions $x_I{}^\circ, \eta_z{}^\circ, q^\circ, \delta^\circ$ are implemented as two-dimensional look-up tables as described above.

In the fourth step of this gain scheduling methodology, nonlocal performance over the entire operating point envelope is analyzed under the assumption that all exogenous signals acting on the system are sufficiently slowly-varying. [4] [6] Here there is some concern over the fact that linear designs were performed about operating points that are not equilibria for the overall longitudinal aerodynamics. However, gain scheduled pitch channel autopilots are traditionally designed using a short-period approximation in a manner similar to that described here and work well in practice. This outcome is most likely due to two main factors. First, the coupling from the short-period aerodynamics into the remaining longitudinal aerodynamics is fairly weak so the variables governed by (2) can be treated as exogenous inputs in (1). Second, the variables governed by (2) and externally generated acceleration commands are sufficiently slowly-varying. In this case, it seems that the stability analysis previously applied to gain scheduled control systems applies here as well, at least approximately. This is only conjecture at this point and a rigorous analysis is required. However, the simulation results presented in the next section support this conclusion.

4. Simulation Results

A *Simulink*™ simulation was constructed for the closed-loop system consisting of the aerodynamics (1) and (2), the actuator dynamics (5), and the gain scheduled autopilot (15). Simulation results are presented here for an initial condition corresponding to the quasi-static operating point parameterized by:

altitude	10,000 *meters*
mach number	3
angle-of-attack	0 *radians*
pitch attitude	0 *radians*

and a commanded normal acceleration consisting of a 1 to -40 g step at $t = 0$ *sec* and a -40 to 0 g step at $t = 15$ *sec*. Various responses are shown in Figs. 3 through 9. The normal acceleration response in Fig. 3 indicates transient response characteristics that are in good agreement with linearized responses throughout the operating envelope. However, even with integral-error action in the autopilot, zero steady-state error for constant commanded normal acceleration is not achieved. This is a direct result of the fact that a true steady-state or equilibrium condition is never reached as indicated in Figs. 4 through 9. In particular, as the autopilot attempts to track a -40 g command over the interval from 0 to 1.5 *sec*, we see that altitude increases from 10,000 to almost 10,250 *meters*, airspeed decreases from mach 3 to mach 2.1, and pitch attitude increases from zero to approximately 70 *degrees*. The time variations in these variables appear to act as perturbations away from a steady-state condition with zero tracking error. In contrast, since a relatively small angle-of-attack is required for a 0 g normal acceleration condition as seen in Fig. 7, the reduced drag that results corresponds to a slower decrease in airspeed as seen in Fig. 5. Also, pitch attitude is almost constant. These slower time variations result in smaller normal acceleration tracking error over the interval from 1.5 to 4.0 *sec*.

5. Concluding Remarks

In this report, the application of a recently developed methodology for gain scheduling to a missile autopilot design problem has been described. Based on the experience gained from this investigation, it appears that this approach is well-suited to missile autopilot design for the following reasons. First, an analytical nonlinear model is not required. Specifically, the calculation of operating point data and linearized models can be easily performed numerically which permits aerodynamic coefficients to be represented in tabular form. Also, the calculation of other quantities required in the autopilot implementation can be done numerically, and the use of splines appears to be useful when certain computations call for differentiation and integration. Second, the wealth of tools and experience for linear control system design can be incorporated into the design process and both the traditional approach of scheduling a finite number of linear point designs along with rapidly emerging techniques for true parameterized linear controller synthesis can be accommodated.

The gain scheduling methodology employed in this work presumes that linear controller designs are computed based on plant linearizations about constant operating points or equilibria. In the autopilot design studied here, this was not possible and the proposed alternative was to use so-called quasi-static operating points consisting of short-period equilibria assuming all other variables are constant. This appears to be common practice and results in autopilots that perform well. Since in general such an approach comes without stability guarantees, a rigorous stability analysis is required to pinpoint the stability mechanisms observed in missile behavior. Such an investigation might not only explain the performance of existing autopilots, but also suggest design modifications to realize improved performance.

The autopilot design problem described in this report can be extended in several ways. First, only a pitch channel autopilot was investigated in this work and a full six degree-of-freedom design would be a true test of the proposed gain scheduling methodology. Second, the autopilot described here was implemented under the assumption that all the required variables were available for feedback. The design of a gain scheduled dynamic output feedback autopilot is a natural and more challenging extension.

6. References

[1] Apkarian and R.J. Adams, "Advanced Gain-Scheduling Techniques for Uncertain Systems," submitted to the *IEEE Transactions on Control Systems Technology,* 1996.

[2] R.A. Hyde and K. Glover, "The Application of Scheduled H-Infinity Controllers to a VSTOL Aircraft," *IEEE Transactions on Automatic Control,* Vol. AC-38, No. 7, pp. 1021-1039, 1993.

[3] I. Kaminer, A. Pascoal, P. Khargonekar, C. Thompson, "A Velocity Algorithm for the Implementation of Gain-Scheduled Controllers," *Automatica,* Vol. 31, No. 8, 1995.

[4] M. Kelemen, "A Stability Property," *IEEE Transactions on Automatic Control,* Vol. AC-31, No. 8, pp. 766-768, 1986.

[5] M.G. Kellett, "Continuous Scheduling of H-Infinity Controllers for a MS760 Paris Aircraft," In *Robust Control System Design Using H-Infinity and Related Methods,* P.H. Hammond (Ed.), pp. 197-223, Institute of Measurement and Control, London, 1991.

[6] D.A. Lawrence and W.J. Rugh, "On a Stability Theorem for Nonlinear Systems with Slowly Varying Inputs," *IEEE Transactions on Automatic Control,* Vol. AC-35, No. 7, pp. 860-864, 1990.

[7] D.A. Lawrence and W.J. Rugh, "Gain Scheduling Dynamic Linear Controllers for a Nonlinear Plant," *Automatica,* Vol. 31, No. 5, pp. 381-390, 1995.

[8] R. Nichols, R. Reichert, W.J. Rugh, "Gain Scheduling for H-Infinity Controllers: A Flight Control Example," *IEEE Transactions on Control Systems Technology,* Vol. 1, No. 2, pp. 69-79, 1993.

[9] A. Ostroff, "High-Alpha Application of Variable-Gain Output Feedback Control," *AIAA Journal of Guidance, Control, and Dynamics,* Vol. 15, pp. 491-497, 1992.

[10] A. Packard, "Gain Scheduling via Linear Fractional Transformations," *Systems & Control Letters,* Vol. 22, pp. 79-92, 1994.

[11] R.T. Reichert, "Dynamic Scheduling of Modern-Robust-Control Autopilot Designs for Missiles," *IEEE Control Systems Magazine,* Vol. 12, No. 5, pp. 35-42, 1992.

[12] S. Shahruz and S. Behtash, "Design of Controllers for Linear Parameter-Varying Systems by the Gain Scheduling Technique," *Journal of Mathematical Analysis and Applications,* Vol. 168, pp. 195-217, 1992.

[13] J.S. Shamma and J.R. Cloutier, "Gain-Scheduled Missile Autopilot Design Using Linear Parameter Varying Transformations," *AIAA Journal of Guidance, Control, and Dynamics,* Vol. 16, No. 2, pp. 256-263, 1993.

[14] J.S. Shamma and M. Athans, "Analysis of Gain Scheduled Control for Nonlinear Plants," *IEEE Transactions on Automatic Control,* Vol. AC-35, No. 8, pp. 898-907, 1990.

[15] White, J. Wozniak, D. Lawrence, "Missile Autopilot Design Using a Gain Scheduling Technique," *Proceedings of the 26th Southeastern Symposium on System Theory,* pp. 606-610, Athens, Ohio, 1994.

Symbol	Description	Value
S	surface area	0.0214 meters^2
m	mass	50.2 kg
g	acceleration due to gravity	9.8 meters/sec^2
d	reference length	0.1651 meters
I_{yy}	pitch moment of inertia	51.14 kg-meter^2
ζ_a	actuator damping ratio	0.7
ω_a	actuator natural frequency	2π 50 radians/sec

Table 1. Missile and Actuator Parameters

Altitude (*meters*)	Mach Number	Angle-of-Attack (*degrees*)
0	0.1	0
3,048	0.6	5
10,668	0.8	10
	1.0	20
	1.15	30
	1.5	40
	2.0	50
	3.0	60
	5.0	

Table 2. Linear Autopilot Design Points

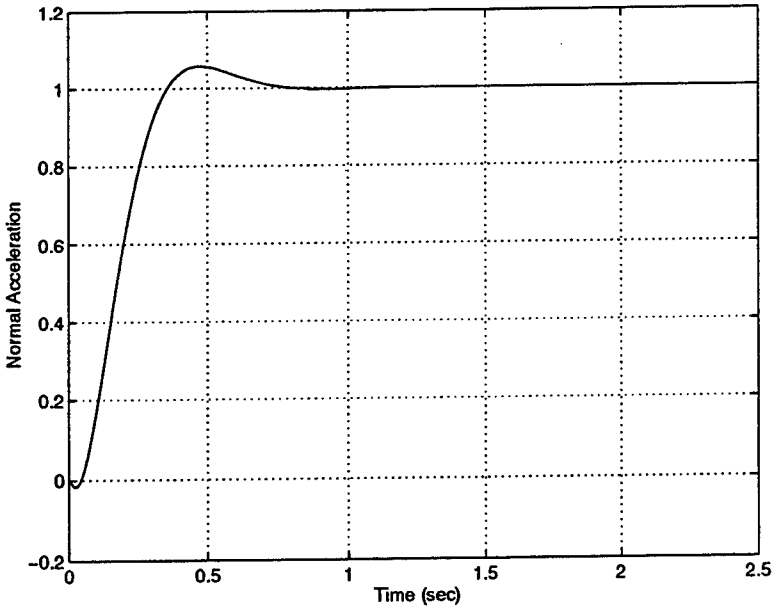


Fig. 1 Linearized closed-loop unit step response

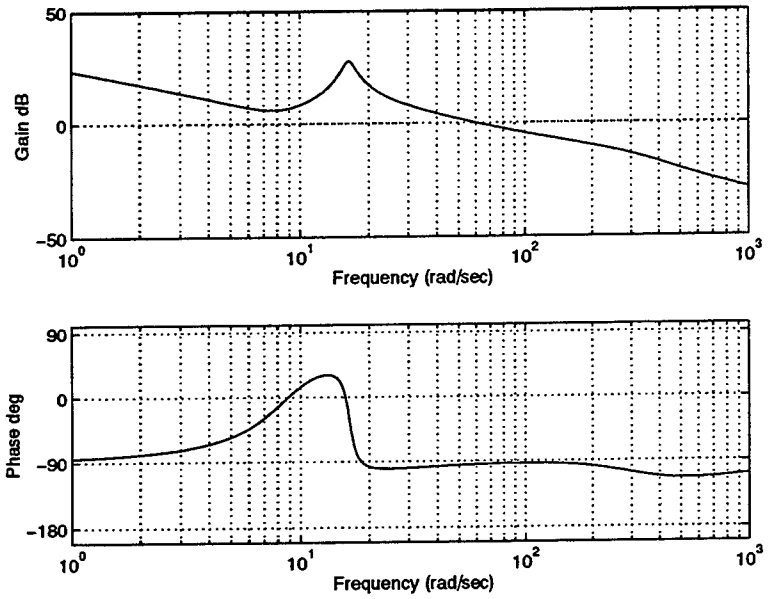


Fig. 2 Linearized open-loop frequency response.

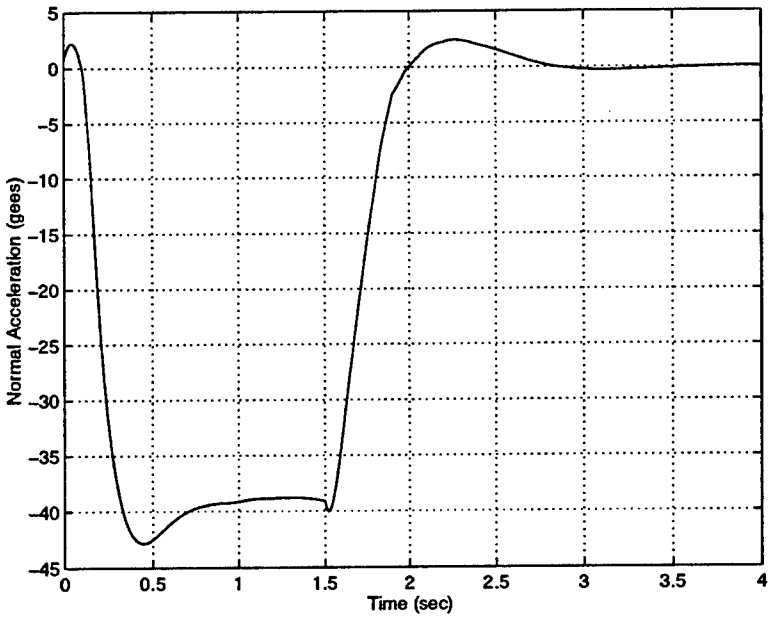


Fig. 3 Normal acceleration response.

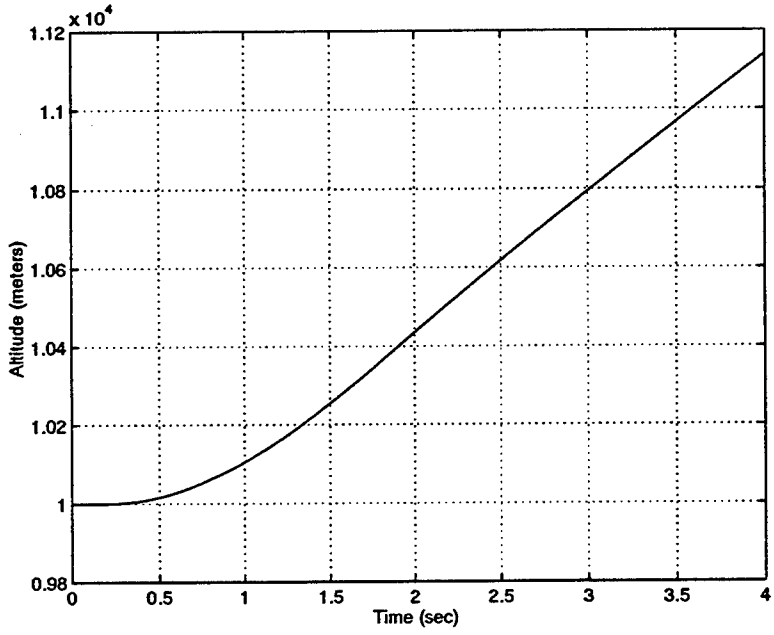


Fig. 4 Altitude response.

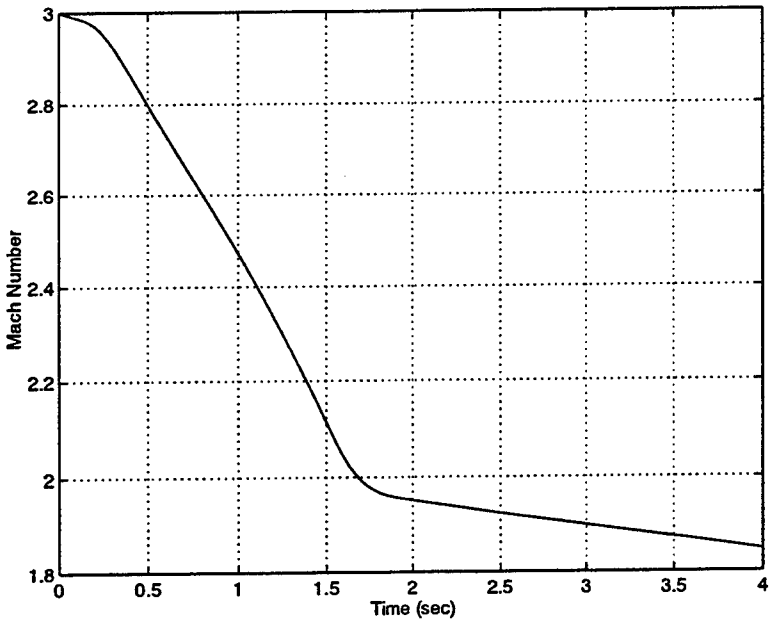


Fig. 5 Mach number response.

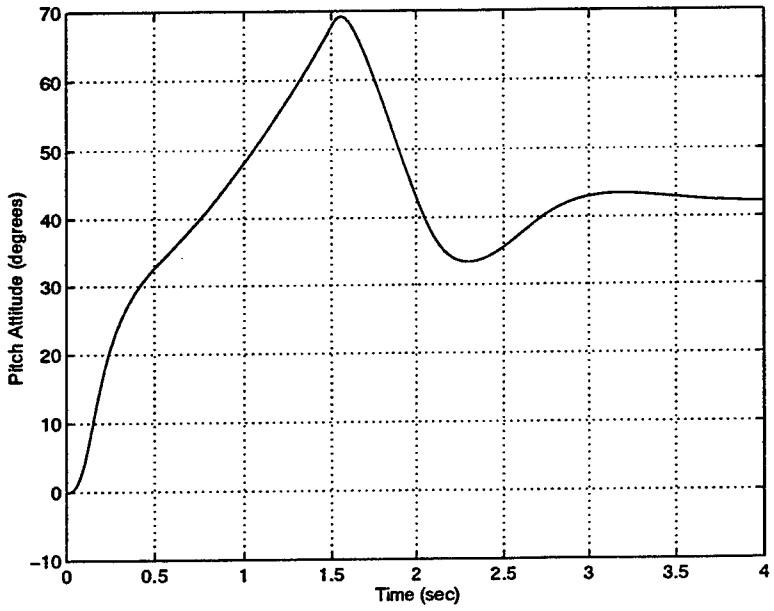


Fig. 6 Pitch attitude response

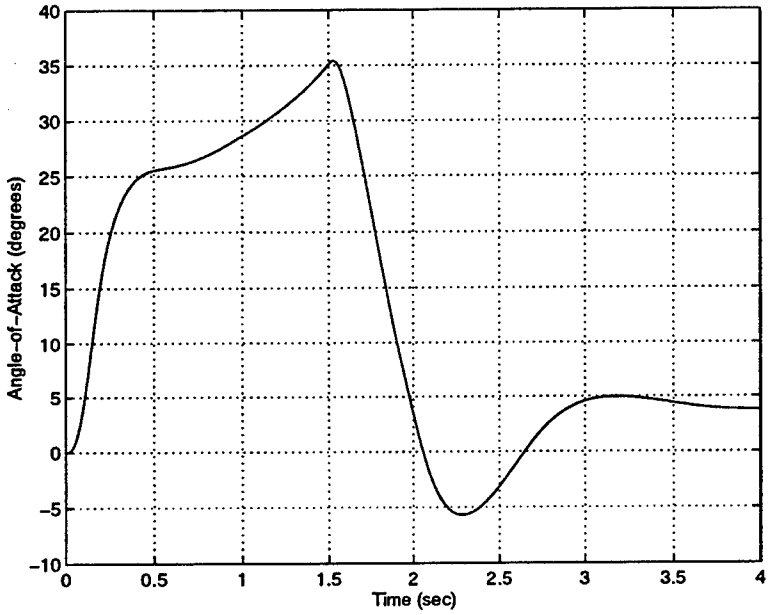


Fig. 7 Angle-of-attack response.

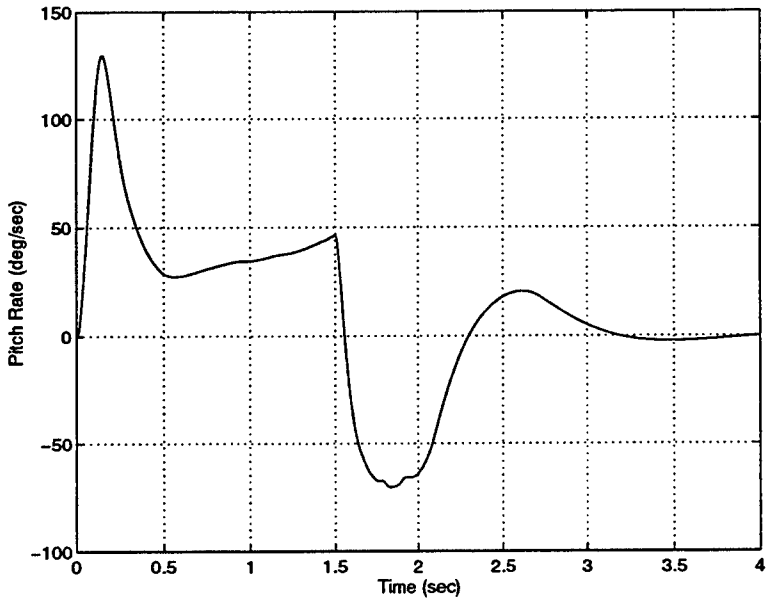


Fig. 8 Pitch rate response.

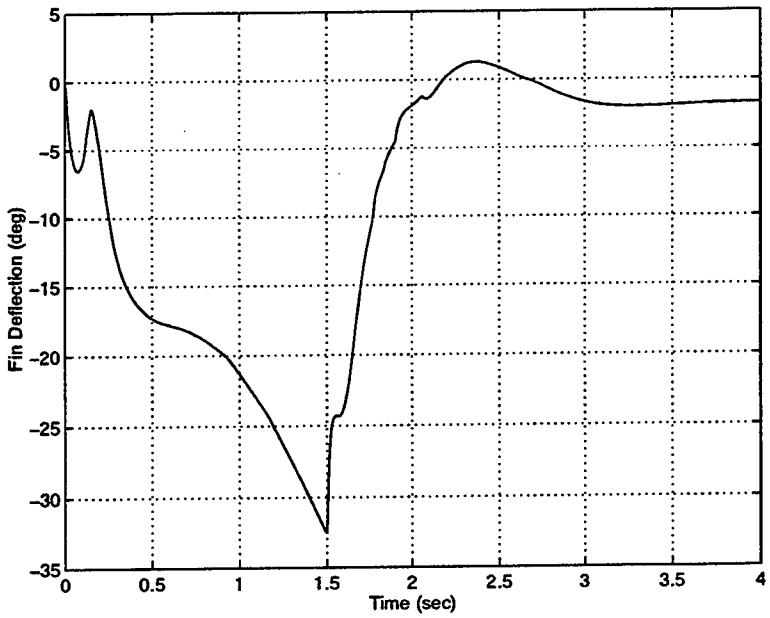


Fig. 9 Fin deflection response

DETERMINATION OF 3D DEFORMATIONS, FORCES AND MOMENTS OF AIRCRAFT TIRES WITH A SYNCHRONIZED OPTICAL AND ANALOG SYSTEM

Junghsen Lieh
Associate Professor
Mechanical & Materials Engineering

Wright State University
3640 Colonel Glenn Highway
Dayton, Ohio 45435

Final Report for:
Summer Faculty Research Program
Wright Laboratory

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

and

Wright Laboratory

September 1996

DETERMINATION OF 3D DEFORMATIONS, FORCES AND MOMENTS OF AIRCRAFT TIRES WITH A SYNCHRONIZED OPTICAL AND ANALOG SYSTEM

Junghsen Lieh
Associate Professor
Mechanical & Materials Engineering
Wright State University

Abstract

The use of optical sensors for geometry and vibration measurements has received a great deal of attention. A vision system may be placed in a remote distance with passive or active light sources such that the motion history can be detected. With a high-speed image processing unit, the measured data may be displayed on a PC in real time. This report describes the measurement procedures and results for aircraft tire properties with Wright State's Optotrak and analog system. The current setup can detect 3D positions of pre-defined points. Infrared light emitted from the markers is received by the vision sensor and digitized for display. The optical system is synchronized with analog (non-optical) devices, such as load cells and pressure transducers. Tires were tested on two machines. In the first experiment, the tire (KC-135 and F-16) was mounted on a Tire Force Machine (TFM), where the vertical, lateral and longitudinal (fore-aft) forces were applied. The system recorded the tire deformation, force, moment, TFM tolerance and table movement as the table travels. The side slip and the lock-up braking force verses the normal load and various yaw angles and their relation with 3D tire deformations were measured. The second experiment involved the use of an 84" dynamometer. An F-16 tire was mounted on the axle where free rolling and brake tests were conducted. The data for the dynamo testing included tire deformations, drum speed, normal force and brake torque.

The measurement and data acquisition were successful. However, due to the collection of tremendous data sets and a limited time for analysis, this report can only present some preliminary results in the global coordinate system. Detailed studies will be conducted in the near future if the time and fund are available.

DETERMINATION OF 3D DEFORMATIONS, FORCES AND MOMENTS OF AIRCRAFT TIRES WITH A SYNCHRONIZED OPTICAL AND ANALOG SYSTEM

Junghsen Lieh

Introduction

A CCD camera contains one or more rows of closed spacing electrodes which apply voltage pulses for storing and transferring signal charges in the semiconductor. The device is an electric image sensor with light sensitive pixels. Since it was proposed in 1970, the application of CCD becomes more and more popular. Numerous researchers are continuously improving the device and further explored its commercial potentials. Today different types of CCD cameras can be found in the markets ranging from medical, biomechanics, manufacturing, traffic to structural and control areas.

A vision sensor is capable of observing geometry and movement of an object. The images can be recorded with passive or active light sources (regular, laser or infrared) reflecting or emitting from the object. A computerized vision sensor contains the following advantages: 1) The light source can be detected at a remote distance; 2) The sensor has a high spatial resolution (for example 2048 pixels per array) with various fields of view; 3) The measurements may be displayed in real time with hundreds of frames per second; 4) The system is portable requiring little human labor. By placing a CCD system at a remote distance from an object can prevent the operator from getting too close to the hazardous environment. The measured 3D displacements may be converted into velocity and acceleration vectors of the object movement or into a local coordinate system for further analysis. This data can be used for modal analysis (such as frequencies and amplitudes with Fast Fourier Transform). It can also be utilized for visualization and control signals.

The objective of the research work is aimed at exploring the use of Wright State's optical and analog system for measurement of aircraft tires under different operating conditions. Active IRED markers (up to 256 points depending on the desired sampling rate) were mounted on the tire, wheel and table. The optical signals were synchronized with non-optical sensors including load cells and pressure transducers. Two testing machines were unitized in the summer, i.e., Tire Force Machine (TFM) and 84" dynamometer. The tests on TFM include free rolling with yaw

angles, lateral, relaxation, side slip, normal loading, and lock-up braking. The second experiment involved the use of an 84" dynamometer, where an F-16 tire was mounted on the axle for free rolling and brake tests.

Methodology

This section describes the instrument, and the procedure of measurements. The description of a synchronized optical and analog system is provided. The measurements of KC-135 and F-16 tires on the Tire Force Machine and 84" dynamometer are illustrated.

1. Synchronized Optical and Analog System

The schematic of the measurement system is shown in Figure 1, which includes the coordinate systems. The system includes

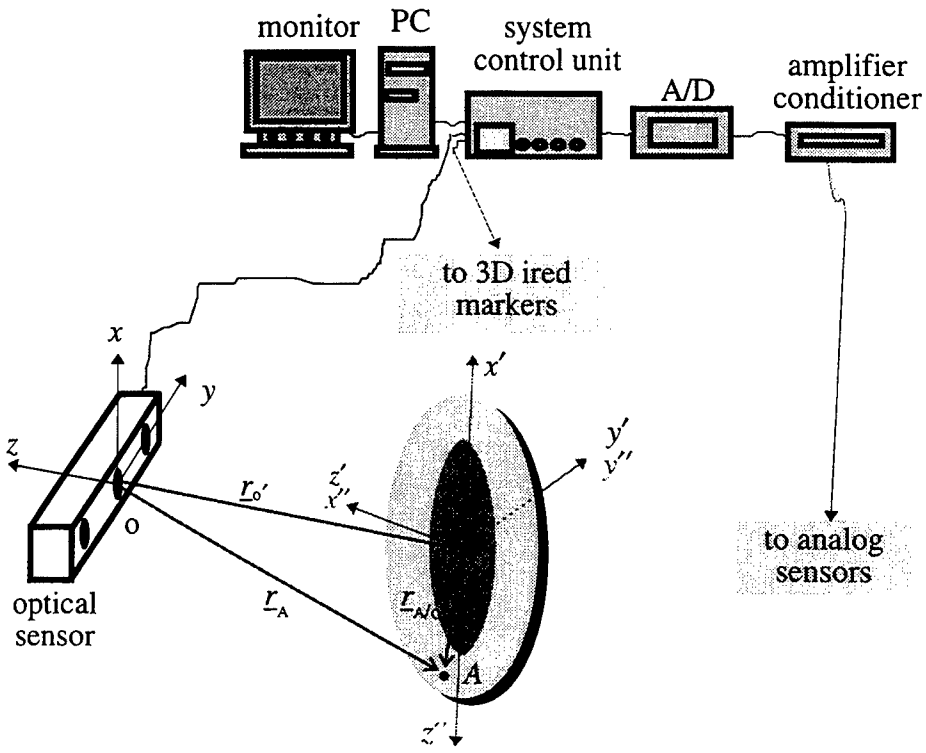
- A 3D position sensor consists of a 3 lens assembly. The depth of field of the current setup ranges from 3 to 20 meters. The accuracy is dependent on the depth of field, for example, the accuracy is approximately 0.05~0.1 mm in x and y directions and 0.1~0.15 mm in z axis when measured from 3 to 4 meters distance.

- An image control unit whose maximum marker and frame sampling rates are 3,000 and 400 Hz, respectively. The frame rate is approximately the ratio of the setup marker sampling rate to the number of markers used. Both light intensity and frequency are controlled by this unit. The maximum number of markers to be detected is 256.

- A data acquisition unit synchronizes 16 analog channels (non-optical) for transducers. A power supply, a system conditioner or an amplifier are needed between the analog sensors and the data acquisition unit.

- A PC with an interface board and a color monitor. The interface board is connected to the system image control unit. Data collected is displayed in the monitor and stored in the hard drive.

- Multiple-channel strobers with infrared markers are linked sequentially and connected to the system image control unit. The infrared signal emitted from the markers are collected by the 3-lens optical sensor and processed by the system image control unit then sent to the PC for display and storage.



o-xyz = optical (global) coordinate system
o′-x′y′z′ = intermediate coordinate system
o′-x″y″z″ = wheel coordinate system

Figure 1. Schematic of the optical and analog system for tire measurement.

2. KC-135 Tire on Tire Force Machine (TFM)

A KC-135 tire (SN 60150607) with a undeformed dimension of 48.6" (dia) × 16.8"(width) was inflated to 170 psi. Two test plans were performed, i.e., free rolling and lateral deformation. In both test plans, the normal load was set to 42,000 lb., and the sampling rate was 5 Hz. 17 markers are used with 10 on the tire, 4 on the wheel and TFM frame, and 3 on the table. For the free rolling test, the tire plane is aligned with the fore-aft direction of the table as shown in Figure 2. Yaw angles of 0, 2, 4, 6° were included in the plan. For the lateral deformation test plan, the tire was rotated 90° as illustrated in Figure 3. The table was moved forward to build up the lateral force until it reached the preset value. The table is then moved backward to reach the negative preset force and finally returned to its zero lateral force position. In addition these test plans, the tire sidewall deflection vs. normal load test was conducted.

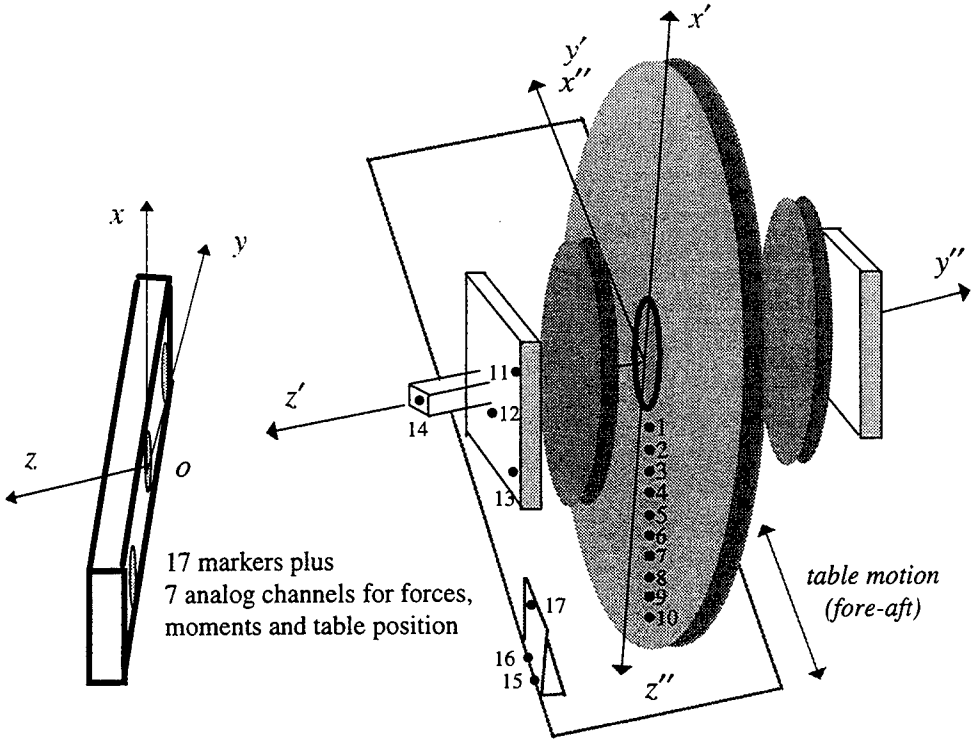


Figure 2. KC-135 Tire Test Plan 1: Free rolling testing.

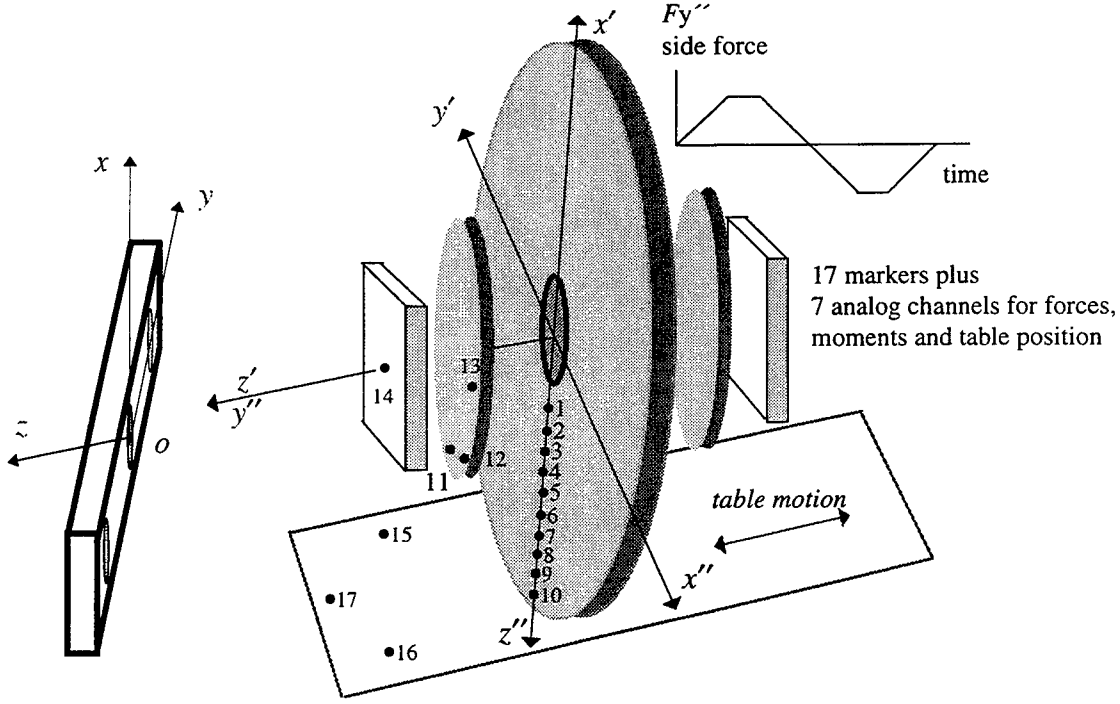


Figure 3. KC-135 Tire Test Plan 2: Lateral deformation testing.

In addition to the optical positions, 7 analog channels recorded the tire forces, moments and table position, i.e., F_x'', F_y'', F_z'', M_x'', M_y'', M_z'', and *pos*, respectively. The maximum side force used for the test plan 2 includes 2k, 4k, 6k, and 8k lb.

3. F-16 Tire on Tire Force Machine (TFM)

A F-16 Goodyear baseline radial block 50 tire (27.75×8.7R14.5/24PR SN 33381603) was inflated to 300 psi. Four different test plans were performed, i.e., (1) lateral-relaxation deformation, (2) slip lateral-relaxation deformation, (3) fore-aft braking (lock up) test, (4) fore-aft free rolling over pressure and slip sensors. In all test plans, the same analog channels were adopted as for the KC-135 tire.

In the first test plan, the normal load was set to 20,000 lb. 18 markers were used including 5 on the tire profile, 4 on the wheel and TFM frame, 3 on the moving table, and 6 on the tire periphery (see Figure 4). The table traveled back and forth, as shown in Figure 4, to induce the lateral force. Peak lateral forces of 8,000, 6,000, 4,000 and 2,000 lb. were applied. Test plan 2 is similar to test plan 1, however, only 3 markers were mounted on tire profile as shown in Figure 5. The table was moved to provide adequate lateral forces in order to induce slip between the table and tire. The normal forces of 20,000, 17,500, 15,000, 12,500 10,000, 7,500 and 5,000 lb. were applied in each run .

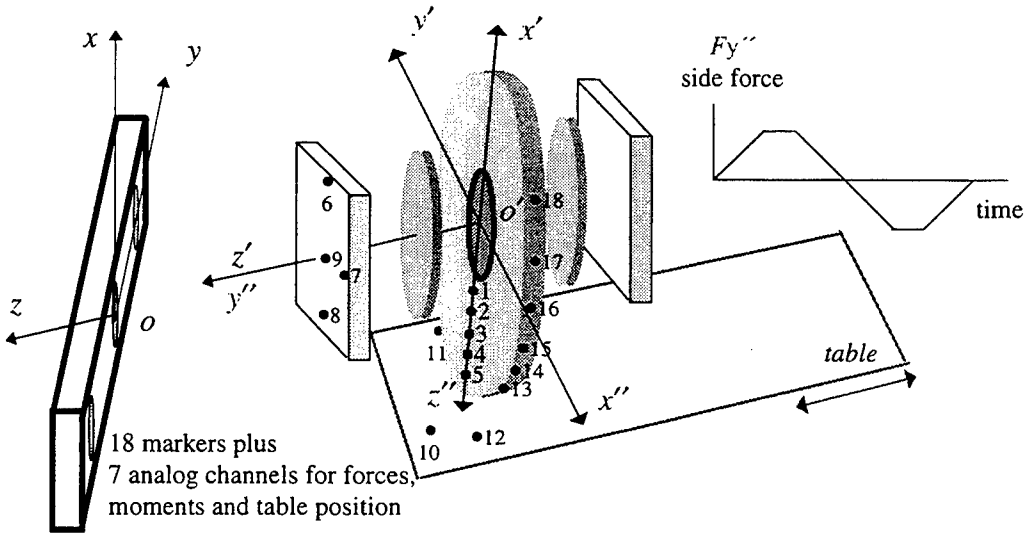


Figure 4. F-16 tire test plan 1: lateral-relaxation test.

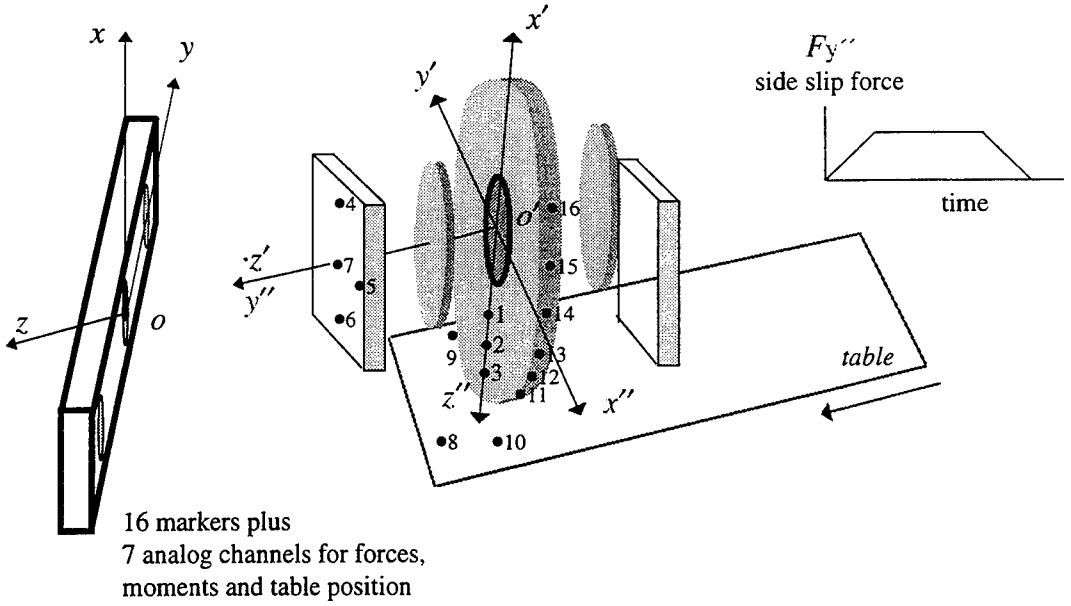


Figure 5. F-16 tire test plan 2: slip lateral-relaxation test.

In the lock-up brake and free rolling tests (plans 3 and 4), 12 markers were used as shown in Figures 6 and 7. During brake, the table was moved forward far enough to allow the table and the tire slip then moved backward for slip again, finally returned back to the origin. The normal force applied for each run was 15,000, 12,500, 10,000, 7,500 and 5,000 lb., respectively. In the free rolling test, the pressure and slip sensors were added and recorded by the Daytronics system. A normal force of 20,000 lb., and yaw angles of 0, 2, 4, 6° were used. After finishing the free rolling test, the pressure and slip sensors were moved 1/4" toward the camera. The free rolling test was then repeated.

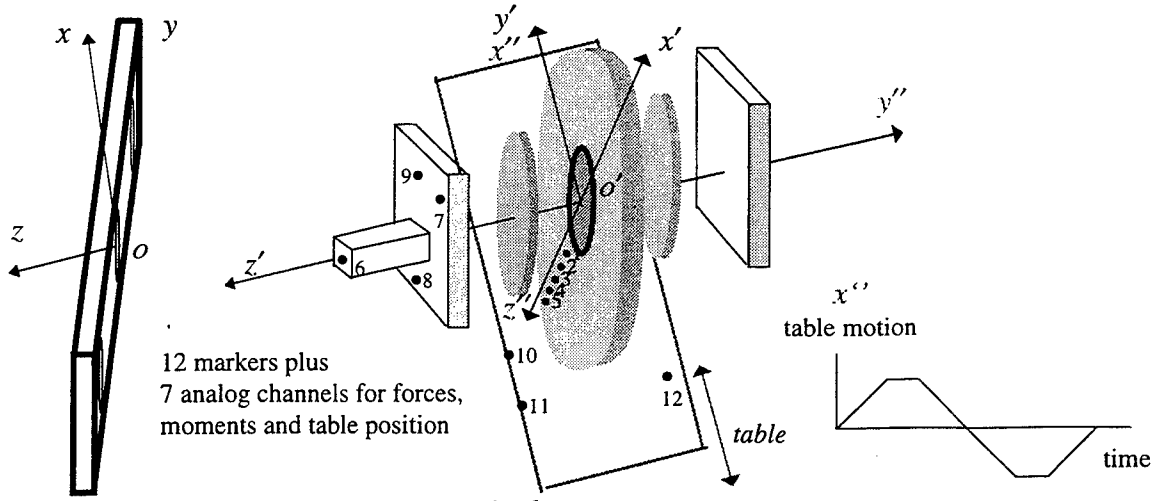
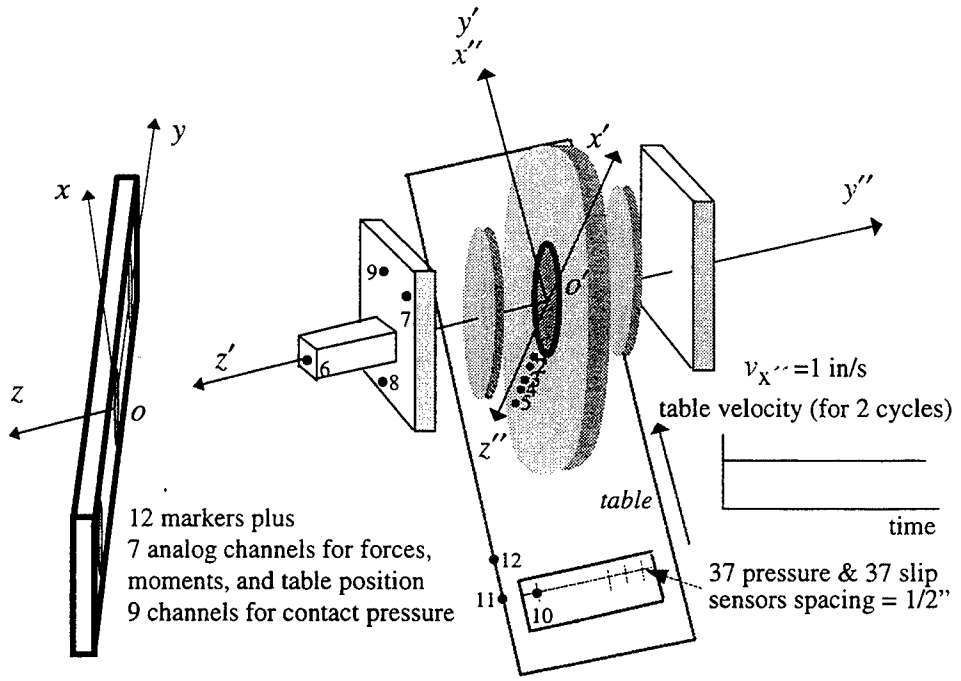
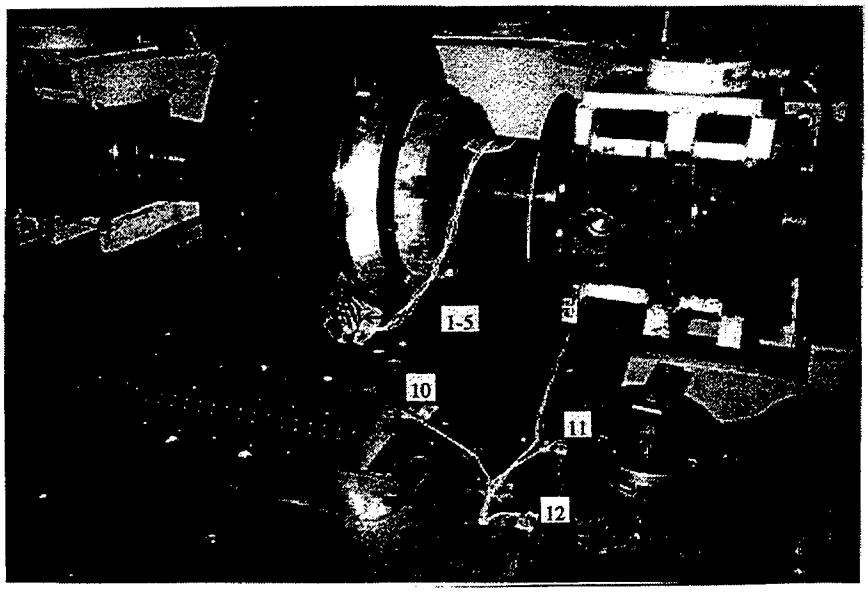


Figure 6. F-16 tire test plan 3: lock-up brake test.



(a) schematic



(b) photo picture

Figure 7. F-16 test plan 4: free rolling test.

4. F-16 Tire on 84" Dynamometer

The same F-16 tire inflated to 300 psi was mounted on the 84" dynamometer as shown in Figure 8. Five markers were used, including 3 on the tire and 2 on the wheel. They were aligned in a line through the center line of the wheel. Free rolling and brake tests were conducted, both with 10,000 lb. normal load. 50 and 75 mph were used in the free rolling test. In the brake test, the drum speed of 50 mph was used. The brake torque of 4,000 and 6,000 ft-LB was applied until the drum was stopped. Since the torque measured through the offset center was the system torque, therefore, it was necessary to obtain the real brake torque with the formula:

$$T_b = T_s * R_r / (R_r + e)$$

where T_b is the brake torque, T_s is the measured system torque, R_r is the radius of the rolling tire (12.73"), and e is the offset (6"). The system torque for the brake torque of 4,000 and 6,000 ft-lb should be 70,700 and 106,000 in-lb, respectively.

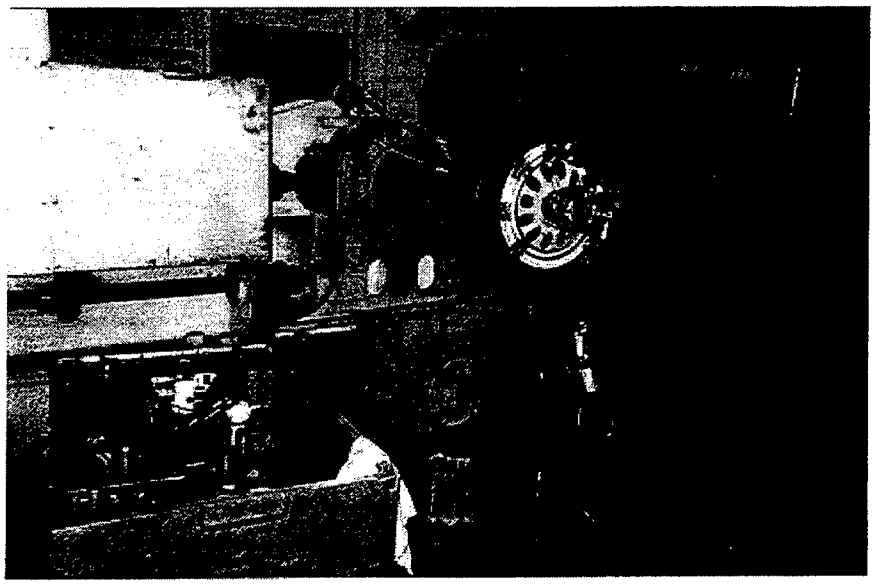


Figure 8. Test of the F-16 tire on 84" dynamometer.

Preliminary Results in Global Coordinate System

Due to the collection of a large number of data files during the summer and a limited time for analysis, this report provides only a few preliminary results based on the global (optical) coordinates. Figure 9 shows the KC-135 sidewall deformation as the normal load was increased from 0 to 42,000 lb. The KC-135 tire sidewall deformation history for the free rolling test with 0^{o} and 4^{o} yaw is given in Figures 10 and 11. Because of the tolerance, the tire deformation data has been contaminated by the TFM movement. The sloppy of the center line of the TFM axle is illustrated in Figures 12 and 13. To obtain precise results, the data collected in the global coordinate system has to be transformed to the wheel coordinates. Table 1 lists the maximum tolerance observed during the tire lateral deformation test. The tire sidewall deformation history during the lateral deformation test (plan 2) is shown in Figure 14. The hysteresis can be observed from this figure, i.e., the tire did not return to its original position (or zero lateral force position).

Table 1. TFM movement (axle) during the lateral deformation test (plan 2).

Maximum longitudinal motion	≈ 15 mm (0.6")
Maximum lateral motion	≈ 35 mm (1.4")
Maximum vertical motion	≈ 5 mm (0.2")

Figure 15 shows the F-16 tire sidewall deformation as the lateral force was applied back and forth (test plan 1). Hysteresis appeared as that of KC-135 tire. The tire relaxation line is given in Figure 16. The TFM movement during the lock-up brake test (plan 3) is provided in Figures 17 and 18, which is similar to the movement for the test of the KC-135 tire. The sidewall deformation history for the free rolling test (plan 4) is illustrated in Figure 19. The tire contact pressure collected by the ODAU (in voltage) for the free rolling test is shown in Figure 20. It was concluded that the pressure collected by ODAU matches that recorded by the Daytronics system.

Summary

The measurement of 3D tire deformations and forces using the Wright State's Optotrak system has been demonstrated in the summer research program. The experiment was successful.

The device was concluded to be suitable for both static and dynamic tire measurements (including deformations, and corresponding forces and moments). All data measured were cast in the global (optical) coordinate system. The TFM was observed to have a large tolerance whose value could be larger than the tire deformation itself. To eliminate the data contamination, wheel coordinates should be used.

Recommendations

After careful evaluations, the following recommendations are considered necessary:

1. The TFM experienced a large sloppy in the mechanism, therefore, all the tire deformations measured in optical coordinates should be converted into the wheel coordinate system. The relative position of point A relative to the wheel center (o') can be expressed in the o-xyz (global) coordinate system (see Figure 1)

   $$\underline{r}_{A/o'} = \underline{r}_A - \underline{r}_{o'}$$

   where $\underline{r}_A$ and $\underline{r}_{o'}$ are measured by the optical system. When expressed in the intermediate (o′-x′y′z′) coordinates, the position vector becomes

   $$\underline{r}'_{A/o'} = \underline{R}_1 (\underline{r}_A - \underline{r}_{o'})$$

   where $\underline{R}_1$ is the rotational matrix from o-xyz to o′-x′y′z′ satisfying the transformation of unit vectors between o′-x′y′z′ and o-xyz coordinate systems:

   $$\underline{e}_{o'} = \underline{R}_1 \underline{e}_o$$

   The final data will have to be converted into the wheel coordinate system (o′-x″y″z″), i.e.,

   $$\underline{r}''_{A/o'} = \underline{R}_2 \underline{r}'_{A/o'}$$

2. To avoid data contamination, a plexiglass (which could vibrate) should not be utilized between the optical system and the object. It is obviously that the plexiglass will distort the position as the light passes through the glass.

3. The analysis of 3D tire data is time consuming. It is recommended that a special program (in C++) be developed to simplify the process of post-analysis.

4. Additional synchronized ODAU channels are needed such that more analog information may be collected.

5. To develop a comprehensive tire model, more tests under different conditions (such as ABS, varying slip and skid combining with yaw/camber angle and alignment changes on TFM and dynamometer) should be conducted. In addition, more tires should also be tested.

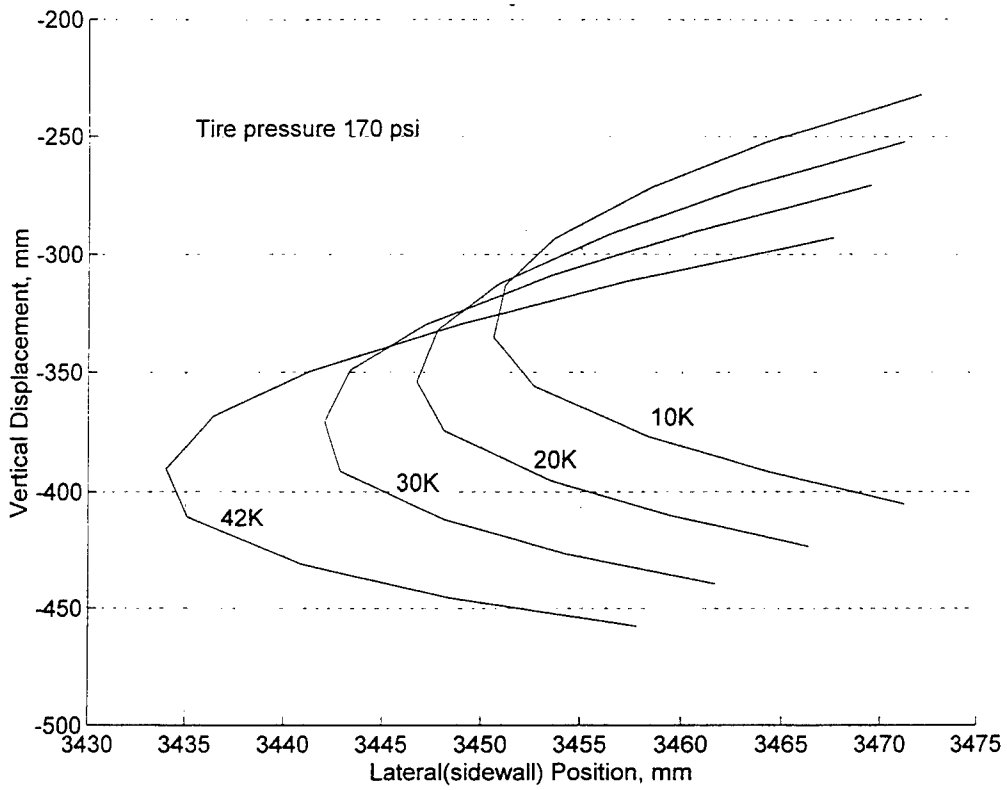


Figure 9. KC-135 tire sidewall deformation vs. normal loading.

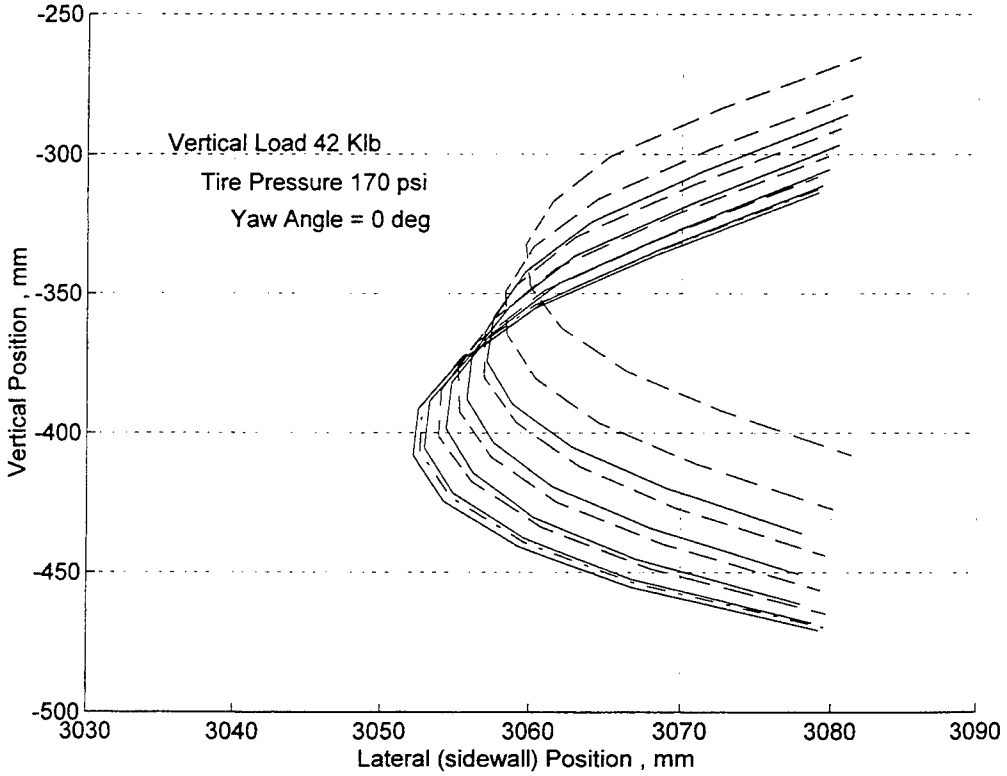


Figure 10. KC-135 tire sidewall deformation history for free rolling with 0^{o} yaw (KC-135 test plane 1).

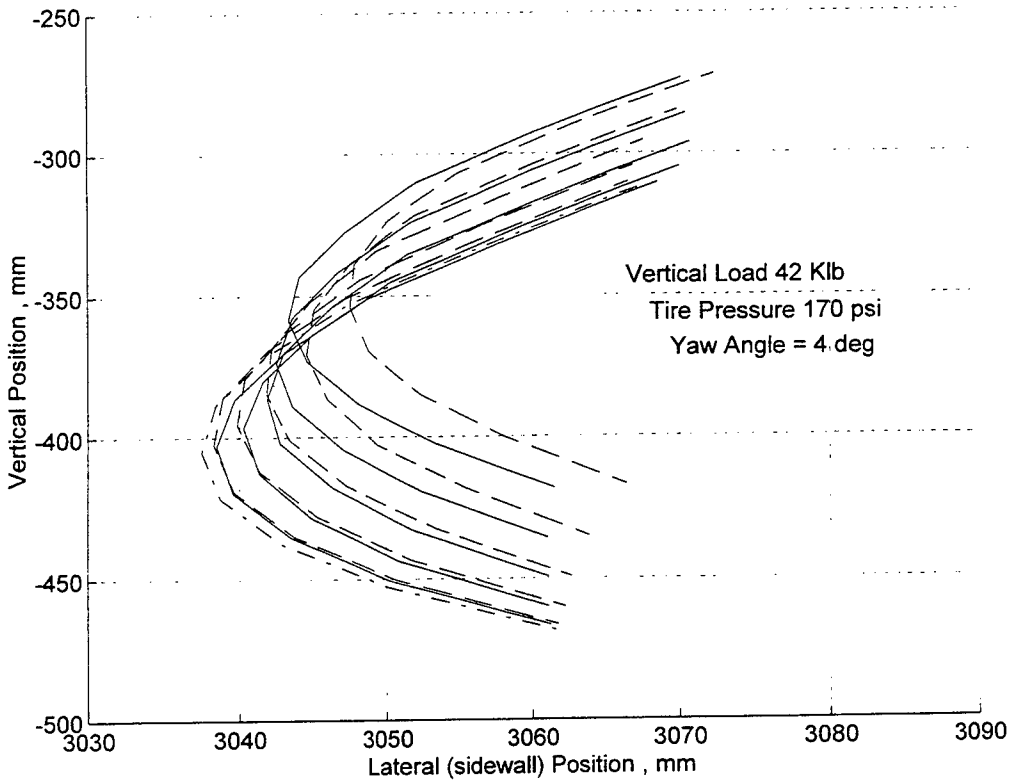


Figure 11. KC-135 tire sidewall deformation history for free rolling with 4° yaw (KC-135 test plane 1).

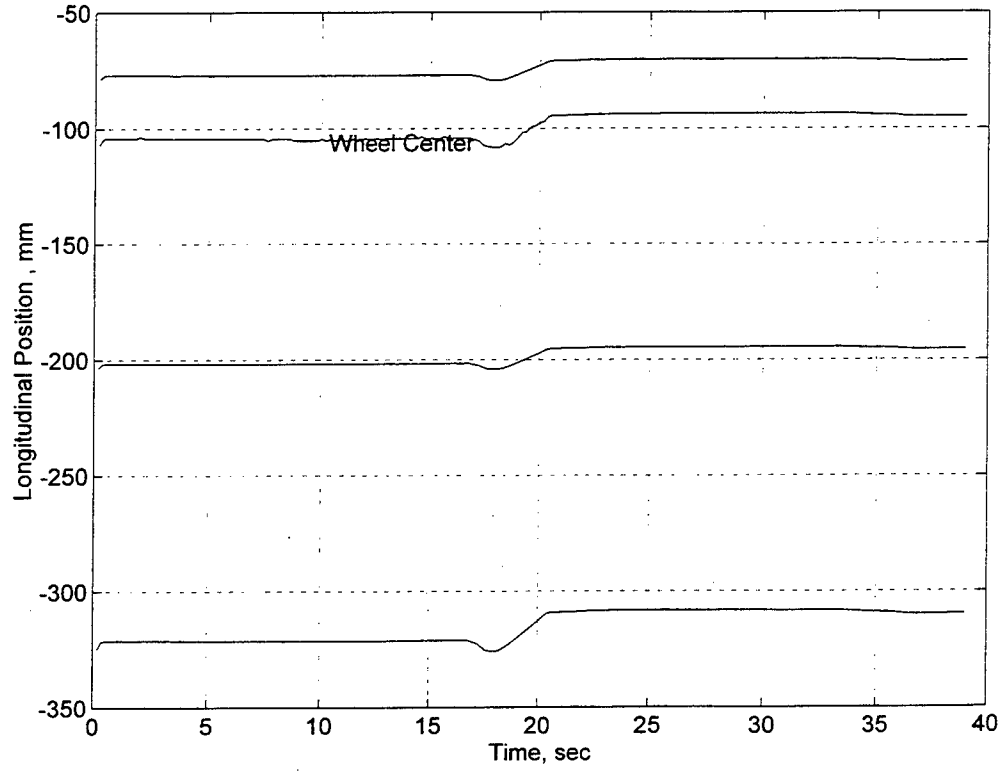


Figure 12. Longitudinal TFM frame and wheel movement during KC-135 tire lateral deformation test (KC-135 test plan 2, peak lateral force = 8,000 lb.).

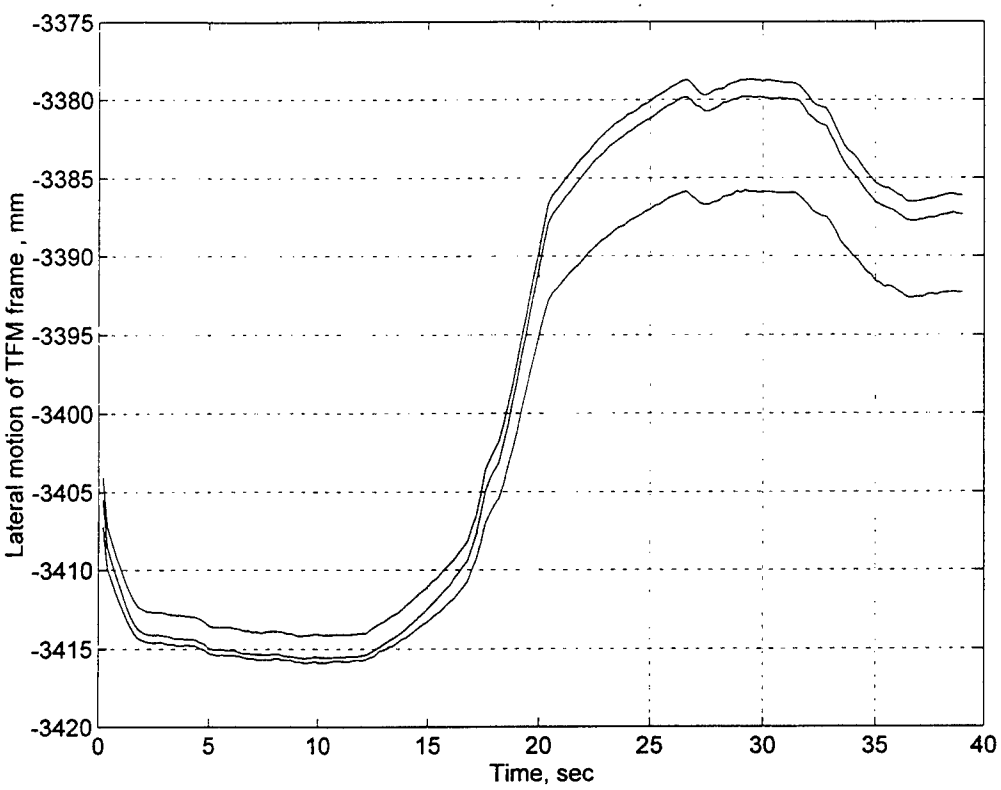


Figure 13. Lateral TFM frame and wheel movement during KC-135 tire lateral deformation test (KC-135 test plan 2, peak lateral force = 8,000 lb.).

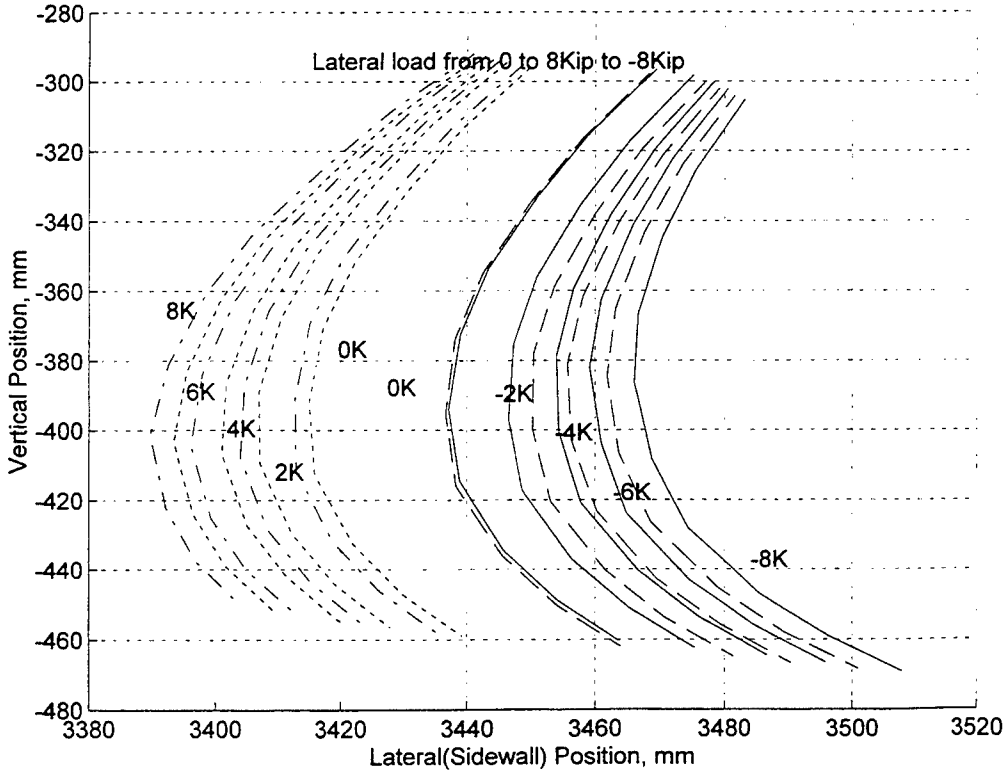


Figure 14. KC-135 tire sidewall deformation under the variation of lateral force (KC-135 test plan 2).

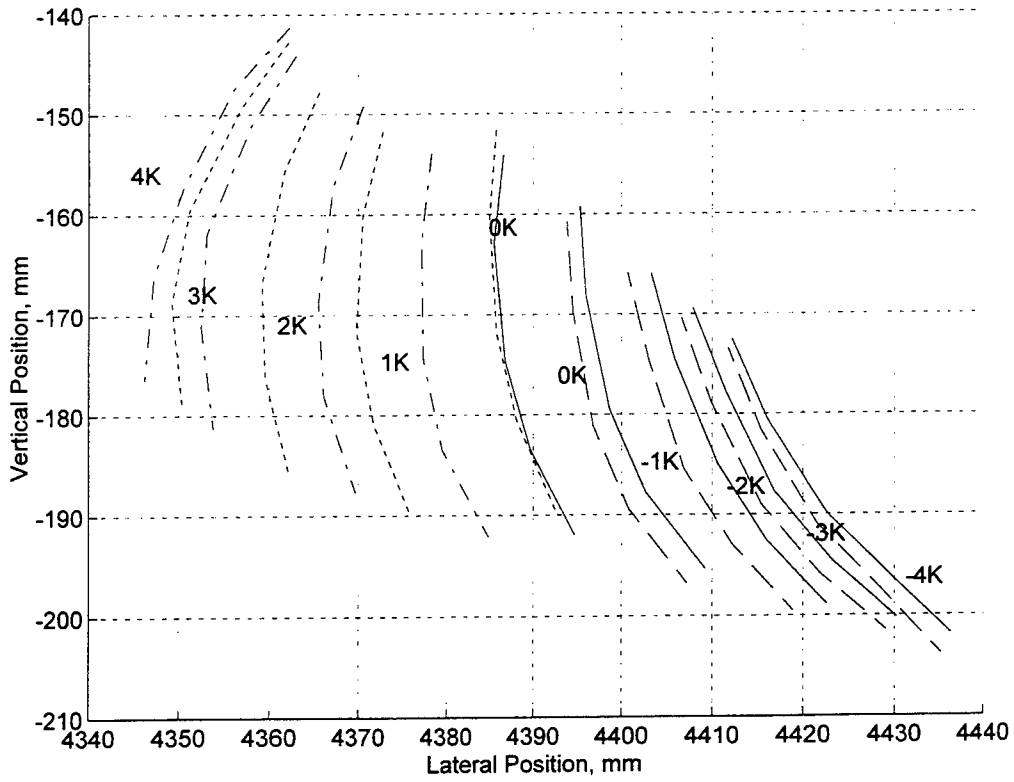


Figure 15. F-16 tire sidewall deformation under the variation of lateral force (F-16 test plan 1).

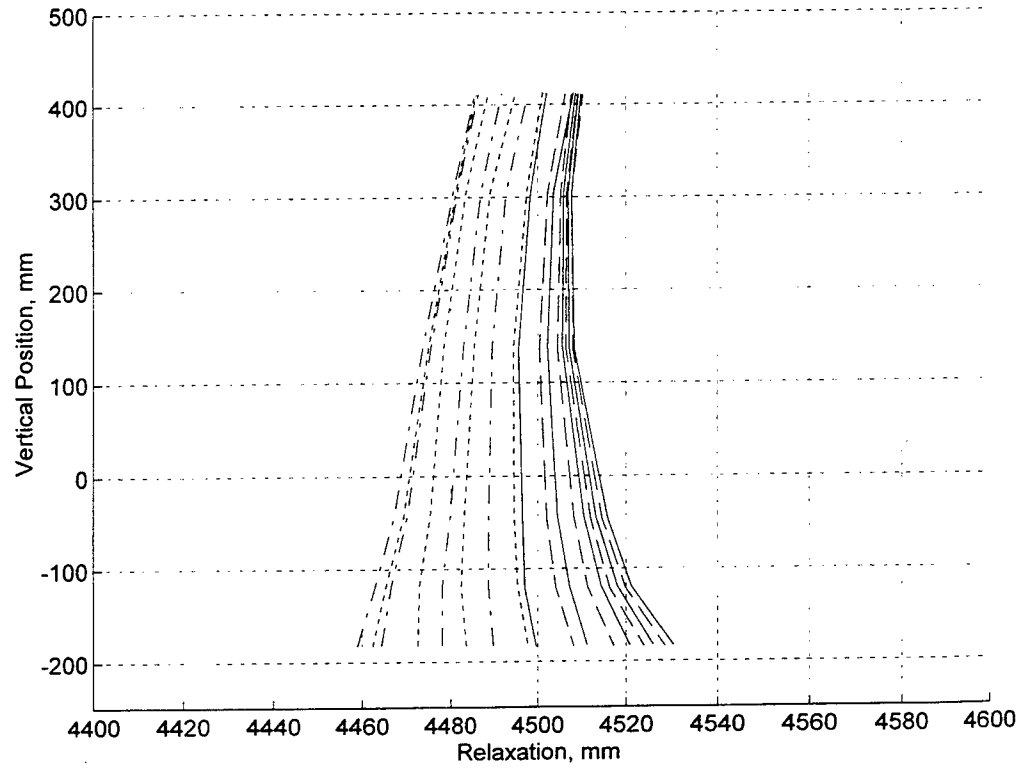


Figure 16. F-16 tire relaxation under the variation of lateral force (F-16 test plan 1).

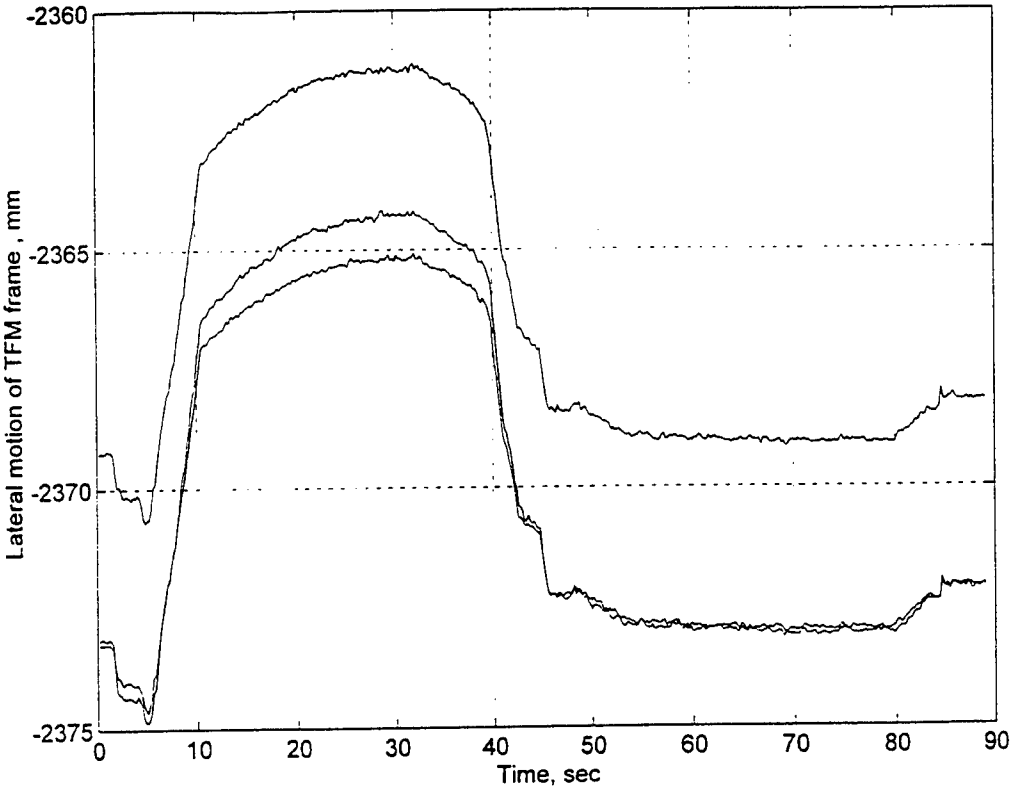


Figure 17. Lateral TFM frame and wheel movement during F-16 tire fore-aft lock-up brake test (F-16 test plan 3 with normal force = 15,000 lb. and 4° yaw).

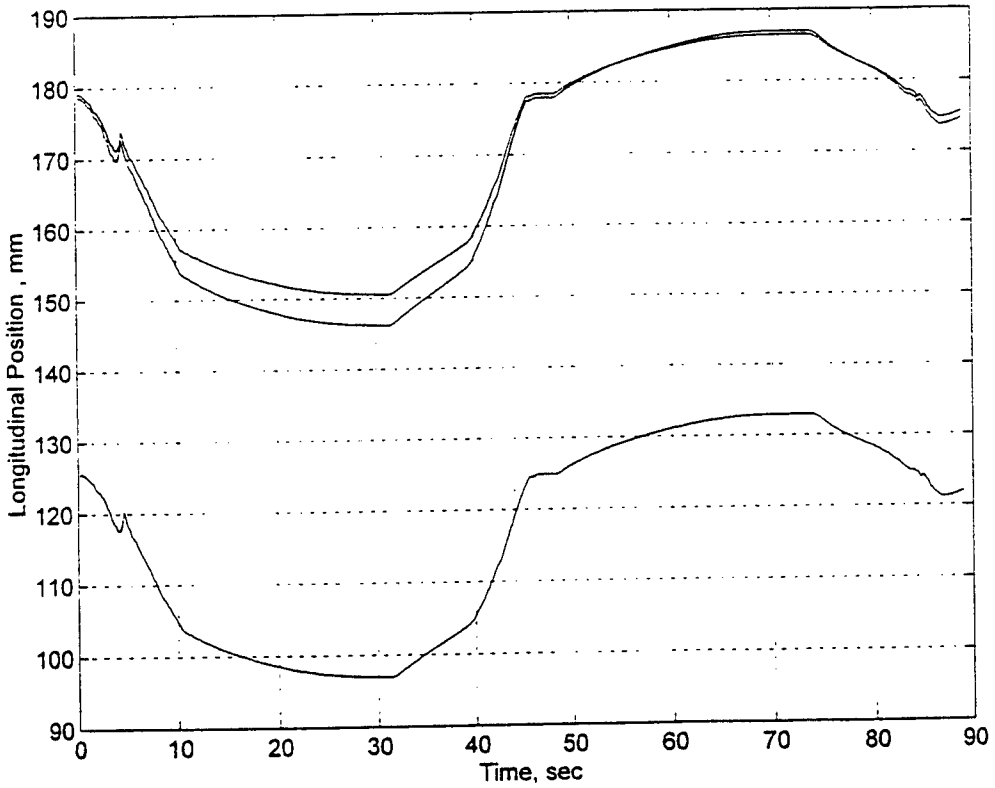


Figure 18. Longitudinal TFM frame and wheel movement during F-16 tire fore-aft lock-up brake test (F-16 test plan 3 with normal force = 15,000 lb. and 4° yaw).

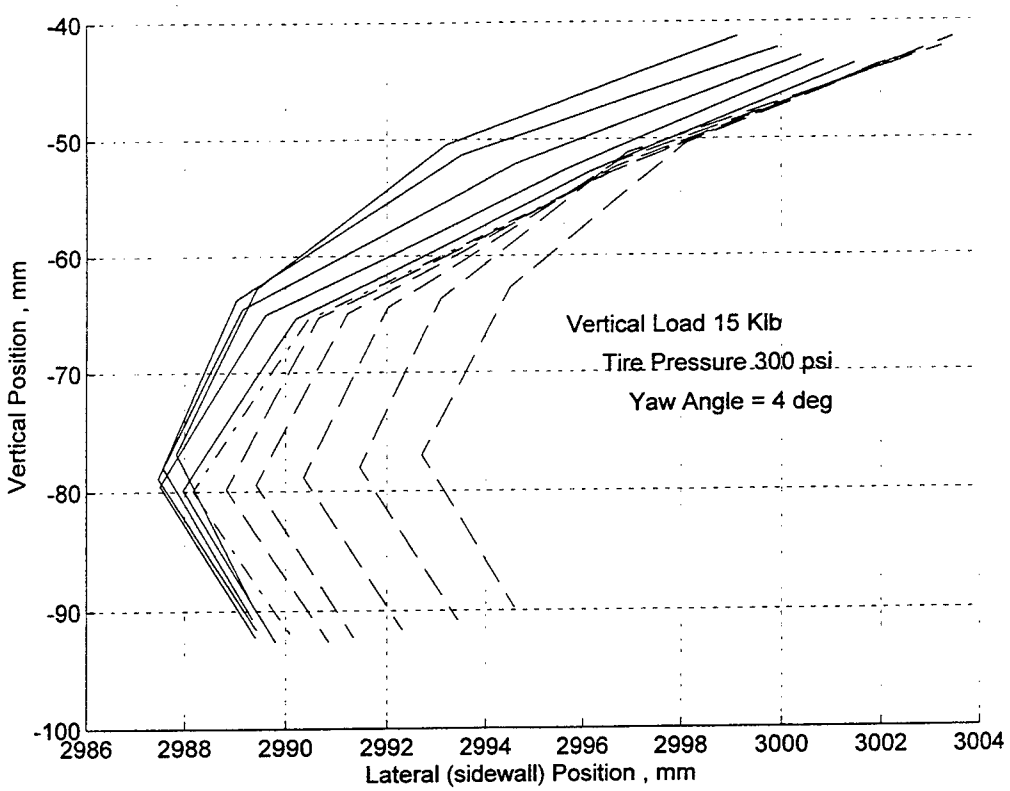


Figure 19. F-16 tire sidewall deformation history for free rolling (F-16 test plane 4 with normal force = 15,000 lb. and 4° yaw).

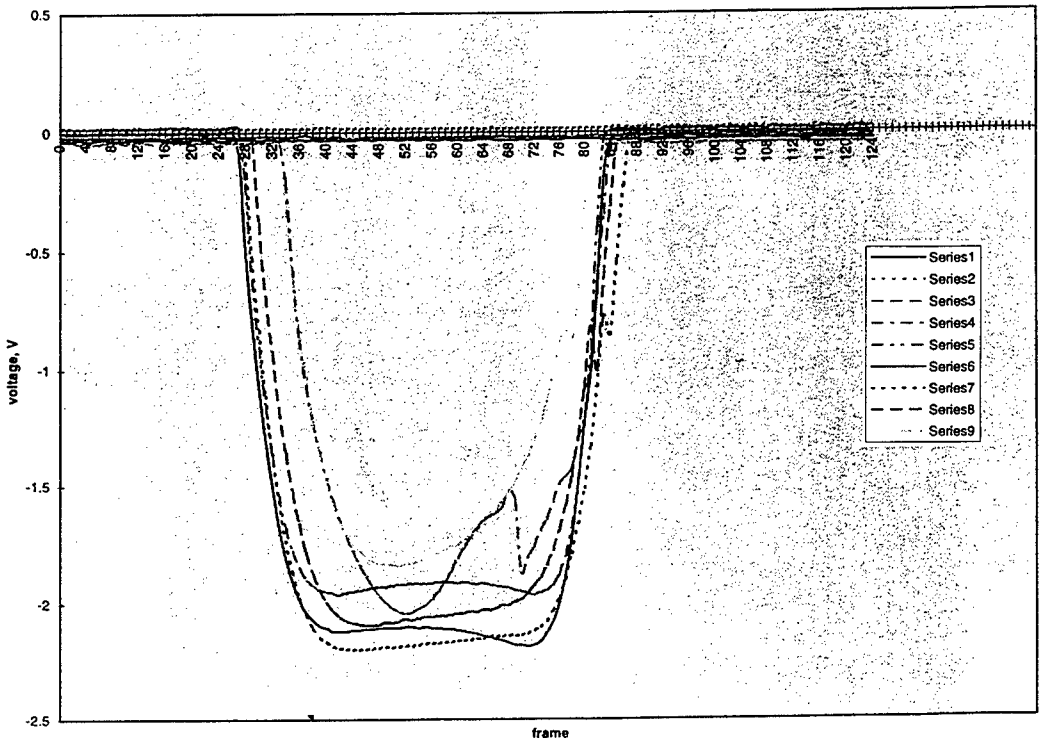


Figure 20. F-16 tire contact pressure for free rolling (F-16 test plane 4 with normal force = 15,000 lb. and 4° yaw, 1volt=200psi).

ACKNOWLEDGMENT

The research work was conducted at the Landing Gear Division of WPAFB. The guidance from John Medzorian and assistance from other WPAFB and SRL staff are greatly appreciated.

NEURAL NETWORK TECHNOLOGY FOR PILOT-VEHICLE INTERFACE AND DECISION AIDS

Chun-Shin Lin
Associate Professor
Department of Electrical Engineering

University of Missouri-Columbia
Columbia, MO 65211

Final Report for:
Summer Faculty Research Program
Wright Laboratory

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

and

Wright Laboratory

July 1996

Abstract

As the sophistication of adversary weapon systems and tactics advances, effective operation of an aircraft in air combat becomes a complicated task. A large amount of information is available to the pilot and many quick decisions and responses need to be made. Many problems in this area involve the development of machine intelligence. Expert systems have been often the choice in solving such problems. Design of an expert system involves the acquisition of knowledge from experts, expression of the knowledge into rules or procedures, and implementation of the intelligent system that uses these rules and procedures. Collecting and organizing the knowledge into a rule base is not an easy job. In contrast to the knowledge-based technique, example-based learning has received increased attention in the past two decades. Neural networks are the major mechanism for example-based learning. One advantage of example-based learning is that, rather than explicit knowledge, cases of desired behavior are collected. This type of learning makes generalizations from observed examples. Even with relatively few examples, an adequate decision could be generated from a well developed example-based learning structure. The Pilot Vehicle Interface Technology Section of Wright Laboratory has shown interest in potential applications of the example-based technology to pilot-vehicle interfaces (PVI) and decision aids. The author was supported under the Summer Faculty Fellowship Program for an 8-week period to investigate the issue. In this report, promising areas in the PVI and related fields to which the neural network technology could be applied are discussed. Topics focus on neural networks for rule development, air combat maneuvering, and voice input. A new special neural network, which can handle problems with high dimensional input space and could be adequate for many PVI and decision aid problems is also introduced.

1. Introduction

As the sophistication of adversary weapon systems and tactics advances, effective operation of an aircraft in air combat becomes a complicated task. A large amount of information is available to the pilot and many quick decisions and responses need to be made. To increase the efficiency and reduce the workload of pilots, cockpit automation aims at

- integration on-board and off-board data sources and efficiently presenting timely and pertinent information to the pilot
- providing decision aids based on available information
- performing pilot tasks with approval.

Available information must be adequately filtered, organized, and probably interpreted. Inference may need to be made in order to generate recommendations and decision aids for the pilot. Many problems in this area involve the development of machine intelligence. Expert systems have been often the choice in solving such problems. Design of an expert system involves the acquisition of knowledge from experts, expression of the knowledge into rules or procedures, and implementation of the intelligent system that uses these rules and procedures. Collecting and organizing the knowledge into a knowledge base is not an easy job. If some situation has not been considered in the design phase, it is likely that the intelligent system will not be able to handle that unexpected situation. For instance, one may have the rule "if the ground is wet, then it has rained." The statement is often but not always true. For a large system, the knowledge base is hardly complete and is often hard to modify. In contrast to the knowledge-based technique, example-based learning has received increased attention in the past two decades. Neural networks [1-5] are the major mechanism for example-based learning. One advantage of the example-based learning is that, rather than explicit knowledge, cases of desired behavior are collected. This type of learning makes generalizations from observed examples. Even with relatively few examples, an adequate decision could be generated from a well developed example-based learning structure. This is an important merit and characteristic of the example-based technique.

There is no doubt that the expert system technique will remain being an important technique for implementing intelligent systems. However, applying the neural network technology in areas that expert systems are less adequate should further improve the capability and reliability of the intelligence in cockpit automation.

The Pilot Vehicle Interface Technology Section of Wright Laboratory has shown interest in potential applications of the example-based technology to PVI and decision aids. The author was supported under the Summer Faculty Fellowship Program for an 8-week period to investigate the issue. The study started from a general survey in the PVI-related areas. The public access catalog of the Wright Laboratory Technical Library and CD ROMs were the major sources in starting this study. Since the PVI and decision aids involve technologies in a broad area and the 8-week study period is relatively short, the results from this study are considered preliminary.

In this report, a brief introduction to neural networks is given in Section 2. A general overview on PVI, decision aids, and cockpit automation is provided in Section 3. Section 4 discusses promising areas in the PVI and related fields to which the neural network technology could be applied. Topics focus on neural networks for rule development, air combat maneuvering, and voice input. Section 5 discusses a new special neural network, which can handle problems with a high dimensional input space and could be adequate for many PVI and decision aids problems. Section 6 gives a conclusion.

2. Brief Introduction to Neural Networks

Artificial neural networks are massive arrays of simple processors that execute in parallel. The interest in this field stems from

- the difficulty of using traditional computer architectures for performing complicated tasks in real-time, and
- the desired capability of learning and generalization from examples

One of the most important capabilities of artificial neural networks is to learn and implement an input-output mapping from examples. Multilayer neural networks (MNNs) [1][2] and basis function networks (BFNs) [3][4][6][7] are widely used neural networks that can perform such learning. Cerebellar Model Articulation Controller (CMAC) [5][8] and functional link networks (FLNs) [9] are other types of neural networks possessing the same learning ability. In this section, a brief introduction to MNNs and BFNs is given.

2.1 Multilayer neural networks

Multilayer neural networks (MNNs) are the best known type of feedforward neural networks. An MNN is a layered network with each layer of the network consists of computing neurons. Figure 1 shows a multilayer neural network with an input layer, hidden layers, and an output layer. Neurons in the input layer only act as buffers for distributing the inputs to neurons in the first hidden layer. Neuron j in the output or a hidden layer sums up its weighted inputs and computes its output y_j as

$$y_j = f(\sum w_{ji} x_i) \tag{1}$$

where x_i is an input from neuron i, w_{ji} is the weight for the connection from neuron i to neuron j, and f is an activation function, which can be a simple threshold function, a sigmoid or a hyperbolic tangent function.

A method called error backpropagation [1] is often used to obtain the derivatives of an error function with respect to the weights in the network. During training, data are provided to the neural network. Derivatives are evaluated and used to update the connection weights. The neural network is expected to be able to respond appropriately after it has had a sufficient exposure to the training data.

One advantage of the multilayer neural network is the computational simplicity. However, it is known that MNNs often encounter difficulty in learning. For a slightly complicated mapping, it is hard to predict how long the learning will take and even whether the learning will converge to an acceptable result.

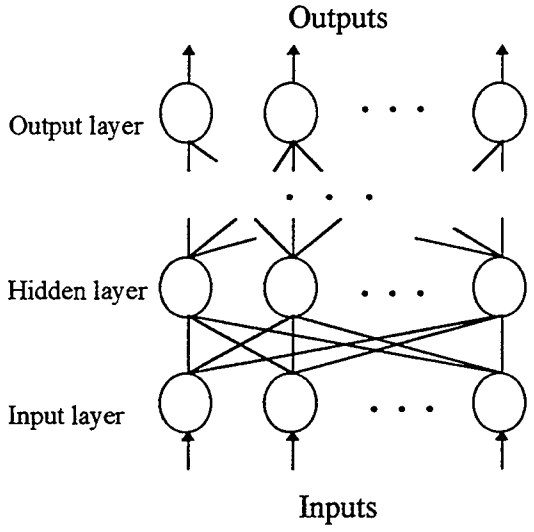


Figure 1. The multilayer neural network.

2.2 Basis function networks

A BFN uses a set of basis functions denoted by $\psi(\mathbf{x}, \mathbf{c}_k)$, where $\mathbf{x}$, referred to as the input vector, is a vector of system variables needed in the representation and $\mathbf{c}_k$ specifies the location of the kth basis function. The unknown system may be described as a function of a linear combination of $\psi(\mathbf{x}, \mathbf{c}_k)$

$$BFN(\mathbf{x}) = \sum_{k=1}^{n} [v_k \ \psi(\mathbf{x}, \mathbf{c}_k)] + b \tag{2}$$

where b denotes a bias and v_k is the weight.

The most commonly used basis for BFNs is the Gaussian function. One typical arrangement is to have $\mathbf{c}_k$ either evenly or randomly distributed in the input space of the function to be learned. A more sophisticated basis function may have the input scaleable and the function rotatable. Figure 2 shows one kind of BFN. Each row is a hidden neuron. The input vector entering the kth neuron is translated by $\mathbf{c}_k$, rotated by $\mathbf{R}_k$, and then scaled by $\mathbf{D}_k$. $\mathbf{D}_k$ is a diagonal matrix, of which each diagonal element is a scaling factor. The transformed vector is used as the input to the basis function ψ, the result of which is multiplied by weight v_k.

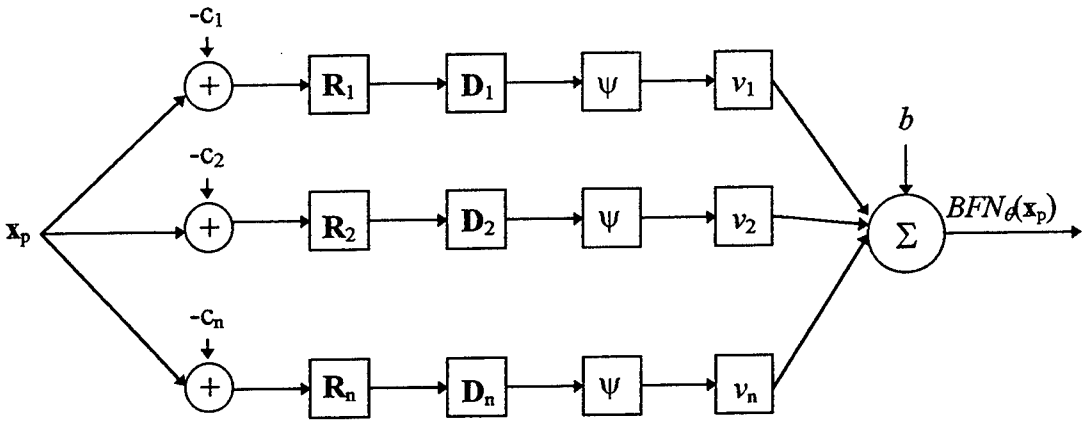


Figure 2. RBFN and its learning parameters.

The network output is computed as

$$BFN_\theta(\mathbf{x}) = \sum_{k=1}^{n} \{v_k \ \psi[\mathbf{D}_k\mathbf{R}_k(\mathbf{x}-\mathbf{c}_k)]\} + b \tag{3}$$

Learning tries to minimize the following cost function:

$$COST(\theta) = \frac{1}{2}\sum_{p}[BFN(\mathbf{x}_p)-y^{(t)}_p]^2 \quad (4)$$

where $y^{(t)}_p$ is the target value for the input $\mathbf{x}_p$ and θ represents the parameters $\mathbf{c}_i$, $\mathbf{R}_i$, $\mathbf{D}_i$, ν_i and b. The partial derivative of the cost function with respect to each component can be derived [6][7] for learning usage.

One major merit of BFNs is the capability of local learning. Learning in a local area is less likely to affect the whole space. However, one possible difficulty is that the necessary number of basis functions may be big.

3. An Overview on PVI and Decision Aids

The PVI and PVI-related areas may be divided into the following three categories:

1. Instrumentation and human-machine communication
2. Information processing and aids to the pilot.
3. Supporting research, modeling and simulation

An overview on these three categories is provided below.

3.1 Instrumentation and human-machine communication

Older cockpits contained a lot of gauge style indicators and control switches. Digital technology has created a lot of changes in today's modern cockpits. The following are some major technologies and R&D topics in this area:

Cockpit display hardware

Major display hardware technologies include CRT, liquid crystal display, electroluminescent display, vacuum fluorescent display, and plasma display [10]. Differences between them include response time, operating temperature range, light emission, life-span, cost to produce, etc. The cockpit display hardware is one fundamental issue in PVI. References [10] and [11] have discussions on the current technologies to be researched.

Voice technology

Speech is the most natural form of communication. It frees the hands and eyes when used as a control method. Other advantages include reduction of workload, better operator mobility, works in dark environment, etc. Disadvantages of the speech technology include [11][12]

- Various physical conditions such as acceleration, vibration and noise can alter speech characteristics [12].

  Under fast acceleration, in access of 6G, speech articulators in the vocal tract can be displaced. Vibration will induce breathing irregularities and affect the speech sound. The body tension resulting from the effort to deal with vibration, coupled with interference to the movement of the lower jaw can influence the voice. Regarding the noise, people tend to increase the volume of their speech in noisy environments and cause the change of speech sound.
- Psychological conditions can alter speech characteristics.

  Stress has clear effects on the vocal apparatus, which, in turn, produces effects on the speech.
- Other audio range emissions can interfere with the desired source.

For voice input, accuracy and reliability need more improvement. A study on connected speech for cockpit applications [13] was performed in Wright Lab. The study evaluated the accuracy of ITT-290 Speech Recognition System using simple and complex phrases. The phrases may be like "change waypoint eight" or "north three four seven six point one two". Study in speech recognition for cockpit applications is still ongoing in the Wright Lab.

Other human-machine communication approaches and technologies

In addition to traditional approaches, other human-machine communication related technologies recommended or under study include eye and head tracking, direct writing on the retina, hand gesturing, etc. [11].

Display management

Display management deals with when, what and how the information should be provided to the display. This is a more software and human factors related topic. The identification of pertinent information, selection of display formats, and the placement of the format on the display screens could be performed by a neural network used as a classifier.

3.2 Information processing and aids to the pilot

Aids to a pilot can be accomplished in three major ways: (1) providing recommendations (2) monitoring the pilot work, and (3) taking over the pilot tasks under approval. The first two often rely on artificial intelligence and expert systems. The third one may be implemented by computer procedures.

One well-known representative example that offers proper, effective assistance to the pilot is the Pilot's Associate (PA) system [14] first developed by Lockheed. Figure 3 shows the modules of the Pilot's Associate system. The PA consists of the system status (SS) module, the situation assessment (SA) module, the mission planner (MP), the tactics planner (TP) and the pilot vehicle interface (PVI). The system status module assesses the internal state of the aircraft. The situation assessment module assesses the state of the world external to the aircraft. The mission planner performs route planning and assesses the impact of world states on the long term objectives of the mission. The tactics planner creates tactical plans for short-term actions and maneuvers. The pilot vehicle interface interprets pilot intentions and provides for transfer of information between the other modules and the pilot via display configurations.

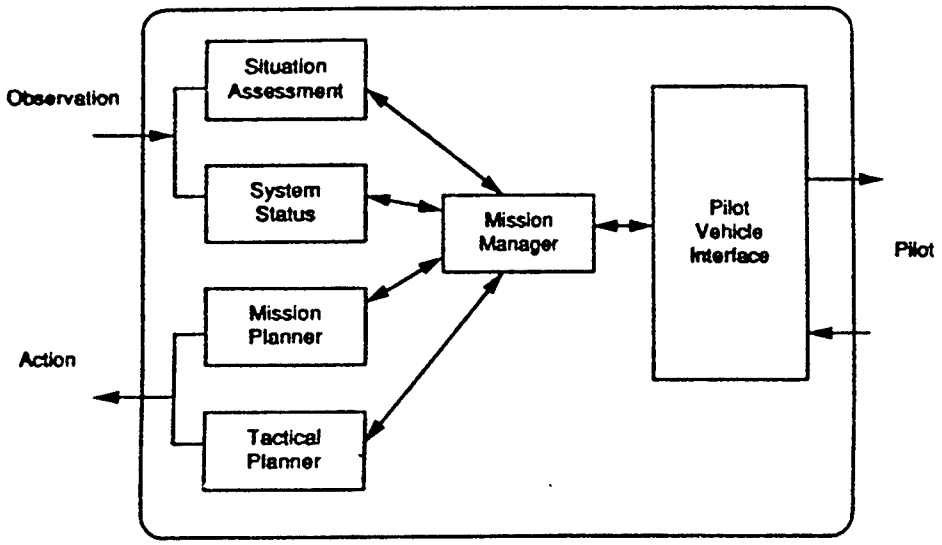


Figure 3. Modules of the Pilot's Associate System

The modules of the PA provide an overall picture of needed functions to performing an air combat. The PA performs the functions and provides recommendations to pilots. Other important issues include at least the following:

- Error monitoring
- Intelligent interface
- Exception handling
- Pilot intent
- Situation awareness
- Adaptive aider

3.3 Supporting research, modeling and simulation

Some other issues are also important to the design of the PVI. They are not hardware or software products but are important to the design. Simulation is one example that is important to evaluate new technologies before they are implemented into the cockpit. Simulation provides an economic, effective, and safe means to design and validate systems. Another example is the study of human factors [15]. Human-in-the-loop issues will affect the design and must be studied in order to optimize the performance and safety.

The simulation facilities include the ground-based simulators and in-flight simulators. The ground-based facilities in the Flight Dynamics Directorate of the Wright Laboratory include

- a large amplitude motion simulator for flight control development and flying quality research and development
- cockpit and mission simulators for supporting technology integration and assessment in simulated mission environments
- several manned combat stations

In-flight simulators are used to support aircraft development, flying qualities research, and test pilot training. Models used in simulation can be constructed using neural networks driven by examples of desired behavior.

4. Potential Applications of Neural Networks

A list of about 150 articles in PVI and related areas has been found from the Wright Laboratory Technical Library database. We studied the abstracts for all of them and studied the complete articles for selected ones. The purpose was to understand the nature of the PVI problems and evaluate the possibility of improving the technologies through the use of neural networks and

example-based learning. This section discusses the study results. This study was conducted over an eight-week period and is considered preliminary.

4.1 Neural networks for learning and rule development

PVI decision aids are usually expert systems. Not many currently available pilot decision aids use neural nets.

One potential application of the neural network technology in decision aids is for learning and rule development from examples. The development of a typical expert system involves the extraction of knowledge from experts and expressing the knowledge as rules. This is considered a difficult job, which may require 12 -18 months [16]. Modification of the knowledge base at a later time may be difficult. However, the advantage is that the system has a perfect explanation ability. On the other hand, a neural network can be developed by a learning algorithm using example data. The procedure may take just a few weeks. Considerably less involvement of experts is required. But there is a difficulty in identifying the purpose of a neuron because the knowledge is implicitly expressed. The advantages and disadvantages of rule-based and neural approaches to the expert system design are complementary. Thus it is natural to suggest the integration of advantages from both.

Sima [16] has discussed two ways in integrating these two approaches. The first one is to create useful neural networks from rules and example data. References [17] to [19] studied the development of neural networks from explicit rules. In the military applications, Stottler and Henke [20] had studied the conversion of a tactical decision aid (TDA) expert system into a neural network based one. Goldsmith et al [21] had investigated the use of expert systems for air combat to develop neural networks for airplane maneuvering. It is also possible to develop a neural network initially using rules and then examples [22-24]. These techniques are useful to convert knowledge-based systems into neural net based systems or to speed up the development by pre-training the neural network using data generated from rules. Potentially, the ability of a rule-based system to generalize can be improved by employing neural networks.

Another way of integration is to create additional rules from example data with the aid of neural learning to complete the knowledge base. Satio and Nakano [25] proposed the heuristics of rule extraction from data and from a trained neural network for medical diagnosis applications. A multilayer neural network was trained on 300 patients. The rules were extracted from the trained

neural network. The authors claimed that the extracted knowledge was considered reasonable by doctors. Enbutsu, Baba, and Hara [26] introduced a so-called causal index as a measure of the causal relationship between input and output neurons. The index was used to extract rules from trained neural networks.

We believe that these above hybrid techniques are applicable to military decision aid application. For large decision aids systems, it is often difficult to thoroughly include all possible occasions into the design. It is also the case that the trend in modern problems for which automated solutions are sought tend to be unstructured. Such problems do not lend themselves to algorithmic solution. Neural networks are one tool available to solve such problems. Modification of a decision aid for neglected occasions often needs to be done. Such work requires information on examples that the expert system fails to successfully handle. The neural network technique may be adopted to learn the needed skill and be integrated with the rule-based system to handle exceptional situations. The technique can be even used to help derive the necessary rules from the examples to assist the modification of the original expert system, thus improving the capability of the expert system. Another similar but not quite the same problem happens in the case with multiple knowledge subsystems. Recommendations or conclusions from different knowledge-based subsystems may conflict. A neural network may be trained to resolve the conflict problem from encountered examples.

4. 2 Air combat maneuvering

The purpose of air combat maneuvering is to escape from possible attacks. Maneuvering may be applied to the airplane engagement or the escape from a missile.

Schvaneveldt [21] had a study on neural network models of air combat maneuvering. Neural networks were proposed for modeling and recommending the maneuvering in a one-on-one engagement. In one part of the study, the Air Combat Expert Simulation (ACES) was used for the development of the neural network model. The neural network tried to learn to select the best as well as the top three types of maneuvering. There were 17 possible types of maneuvering considered in that part of the study.

Since the neural network learned from an expert system, one would not expect that it outperformed the expert system. The selections from the ACES and the trained neural networks were compared to the selections by pilots. There were observable differences between the pilot

group, ACES, and the neural network. These selections have not been tested on the flight simulator so that which one is the best is not known. However, one may see that it should be easy to refine the neural network by further training the network using new examples. These examples would come from experiences in manned or unmanned simulations. Pilots can try their selected maneuvering on the simulator. The result can be used to train the neural net to achieve the optimal performance. However, it would be relatively difficult to modify the expert system utilizing the collected examples.

In another part of the study in [21], a neural network was used to predict the *g*-forces and the roll rate. The implicit usage is to decide the best *g* and roll according to the current airspace situation for possible autopiloting. One difference from the former method introduced in the previous paragraph is that this one considers the modeling of time-varying maneuvering variables and the former considers only the type of maneuvering. This part of the study used maneuvering data collected from pilots in one-on-one engagement simulations. Results of one good pilot who defeated the hostile airplane three times out of four engagements were used for neural network training. A neural network was trained to learn the decisions made by this pilot. The study is somewhat preliminary. With data from only four engagements, the training patterns were not likely to cover the entire space of the air combat situations. Another possible issue is that, for the same situation, the same pilot may take different actions. The issue could be critical but had not been addressed. Although the study seems not thorough, the introduced approach does provide some ideas and reference for further research on this application as well as possible applications to maneuvering to escape from a missile attack.

Another study on maneuvering is against the missile. Harmon et al [27][28] used a 2-D simulator to show how reinforcement learning could be used to find the ways for the airplane to escape from a missile attack. The used model is a simple one. How this technique performs for a 3-D and more complicated (realistic) model is an interesting question.

4.3 Voice input

Speech is the most natural form of communication. It frees hands and eyes when used as a control method. Voice control techniques will allow the pilot to command and query the aircraft and transfer control of the aircraft to automated systems during incapacitating emergencies.

Hidden Markov models (HMMs) have been widely used for continuous speech recognition [29][30(ch.3)]. The approach uses a finite number of states for modeling a class of sound, which can be a sentence, a word or a sub-word. The total number of models will be much smaller if the words or the sub-words (instead of sentences) are to be modeled. The speech is usually divided into short periods (10~30ms). Features, usually the spectral vector, are obtained for each interval. The model transitions from state to state (see Figure 4 for illustration). Each transition is assigned a probability ($P(q_i|q_j)$), which is one parameter of the model. For each state, the observation (for instance the spectral vector) is stochastic. Thus the probability to have some specific observation (i.e. $P(x_n|q_i)$) is another type of parameter. In training, the parameters are estimated. Either the maximum likelihood criterion or Viterbi criterion (an approximate one) is used in iterative parameter estimation. Two estimation algorithms using the two criteria are called Forward-backward algorithm and Viterbi algorithm, respectively. During speech recognition, the probability P(the sequence of observation|the set of parameters for class i) is evaluated and used to determine the class for a time series of observations. Efficient search algorithms are used to search for the class with the maximum or approximate maximum likelihood. The approach is so popular because of the inherent statistical framework, the ease and availability of training algorithms for estimating the parameters from the speech data, the flexibility of the resulting systems (easy to change size, architecture, etc.), and the ease of implementation.

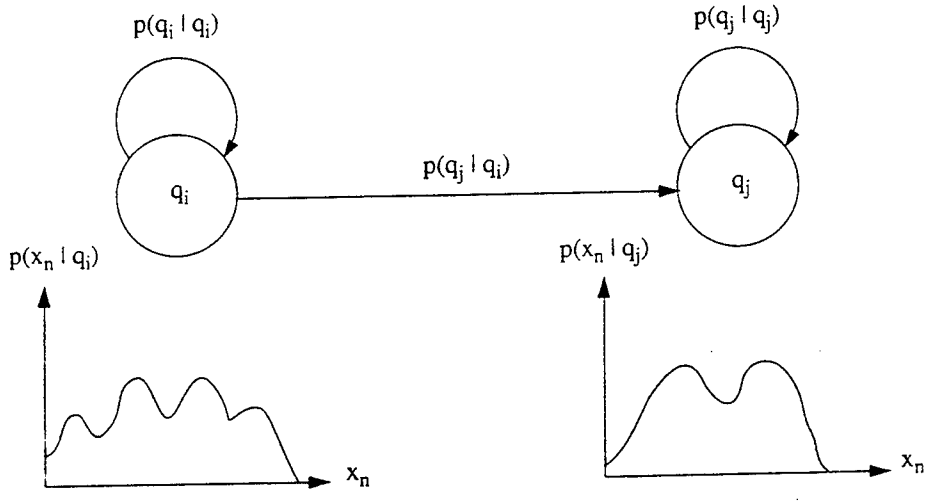


Figure 4. A schematic of a two state HMM

In recent years, there has been an increased interest in approaches based on neural networks [30(ch. 5)][31]. Advantages and reasons to use neural networks include:

- only weak hypotheses about statistical distribution are necessary
- humans understand speech through the use of biological neural networks
- neural networks are adaptive and can learn and generalize for new data.
- neural networks can generate any kind of nonlinear boundaries between classes
- neural networks are highly parallel for high performance.

The simplest way to perform sequence recognition is to turn the temporal pattern into a spatial pattern at the input of a multilayer neural network. A tapped delay line can be used at the input layer of a multilayer neural network for this purpose. The network can be trained using the error backpropagation algorithm [1]. Some modified models [30(ch.5)] have feedback from either the hidden layer or the output layer to the input. These models have good performance on short, isolated speech units, but are not effective for large scale continuous speech recognition.

Motivated by the success of HMM algorithms for speech recognition problems, neural network implementations of HMMs are of interest. The fundamental HMM algorithm is mapped to a connectionist framework. The neural networks have been often suggested for modeling the output probability density for each state in an HMM.

A lot of research efforts have been devoted to the speech recognition problem. But there are still gaps in performance we can address and that we would like to achieve. For the pilot-vehicle interface, the vocabulary size may be small but problems such as noise, and voice change at high G, and under stress are additional issues.

4.4 Other potential applications

Two of the most promising areas for the neural network technology are classification and interpolation. Thus, all PVI problems that involve the classification and interpolation might be good candidate areas for neural network applications. Situation awareness, pilot intent, exception handling, and hand gesturing are some examples.

5. A Special Neural Network Structure

This section discusses the drawbacks of current major neural network technologies and introduces a new type that can handle high-dimensional problems and has good learning behavior.

The neural networks including multilayer neural networks (MNNs) [1][2], basis function networks (BFNs) [3][4], and CMAC [5] are able to learn static mappings. In applications, during the training phase, cases of desired behavior are collected and the neural network learns to generalize the implicit information from observed examples. The trained neural network is expected to generate correct output for given inputs in later usage. Neural networks that perform static mapping have broad applications in pattern recognition, system identification, control, etc.

MNNs and BFNs are widely used neural networks for static mapping. However, they both have some drawbacks and limitations. It is known that MNNs often encounter difficulty in learning. For a slightly complicated mapping, it is hard to predict how long the learning will take and whether the learning will converge to an acceptable result. Another type of neural network, BFNs, often use the Gaussian function as the basis function (although others are used as well). Since a Gaussian function provides fitting in a local area, the learning convergence is a less difficult problem compared to that in MNNs. In addition, learning is likely to only alter local information and thus, will be less likely to destroy the previously learned information. However, the number of basis functions may become enormous for problems with many input variables or a complex learning space that needs a large number of examples. To increase the fitting power of each basis function in order to reduce the number of basis functions needed, many researchers suggested to make the Gaussian function scaleable in each dimension and rotatable in the input space [6][7]. The trade-off is the increased learning difficulty. In addition to MNNs and BFNs, CMAC is another structure capable of learning static mappings. CMAC is a table look-up neurocomputing technique. It distributively stores mapping information in memory, such as RAM or EPROM. Neighboring data points in the input space share some information stored in the memory. This creates the generalization and on the other hand reduces the needed memory space. While the learning is usually local, it is fast and has been proved to always converge [8]. The CMAC can be viewed as a basis function network with the use of a plateau function (constant basis). Thus, it has a similar drawback like BFNs; the memory size grows very fast while the number of input variables increases.

A new neural network that consists of small CMACs is recently introduced by Lin and Li [32]. The new structure can help solve the fast size-growing problem as well as the learning difficulty existent in currently available types of neural networks. Figure 5 shows the proposed neural network structure. The network output *SOPNN*( ) is the sum of the outputs from a set of submodules. The output of the neural network has the sum-of product form $\sum_{i=1}^{N_p} \prod_{j=1}^{N_v} f_{ij}(x_j)$ and the function $f_{ij}(x_j)$ is expressed in a form as $\sum_{k=l}^{l+n_e-1} w_{ijk} B_{jk}(x_j)$ where $B_{jk}()$ is a single-variable basis function, w_{ijk}'s are weight values stored in memory, l is the quantized element number for x_j, and n_e is the number of basis functions in the neighborhood used for storing information for x_j. The method can be implemented by a memory-based structure that uses overlapped rectangular pulses as basis functions. With such basis functions, only W_{ijk}'s need to be memorized and linear memory arrays are enough for this purpose.

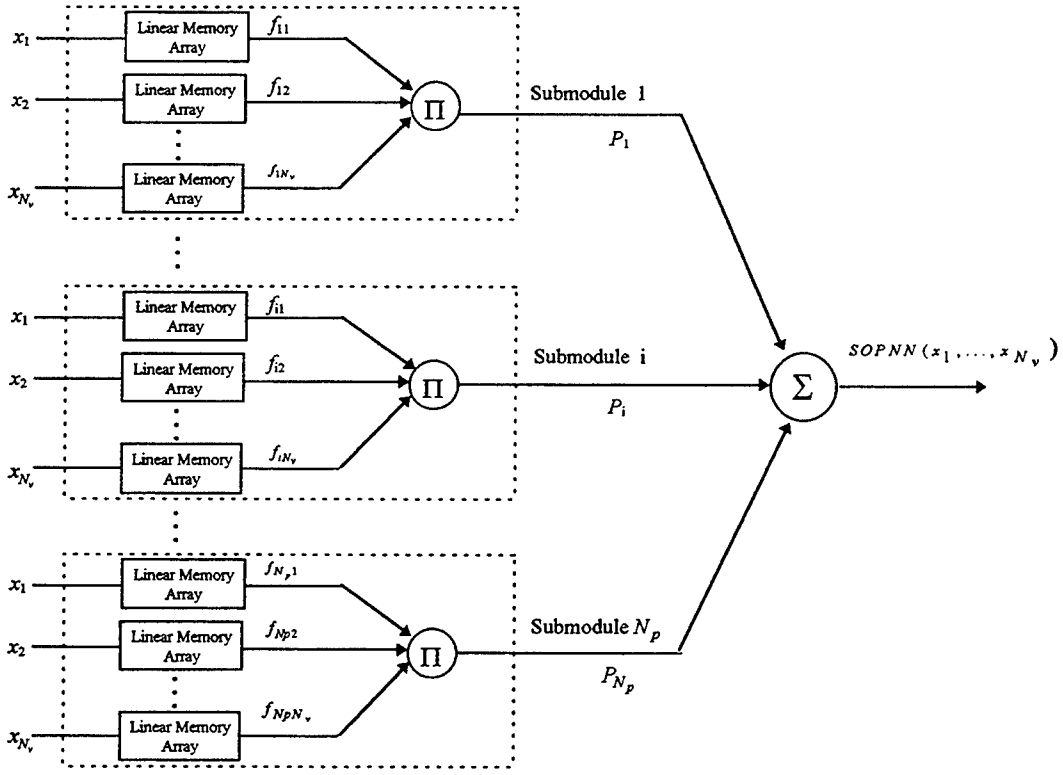
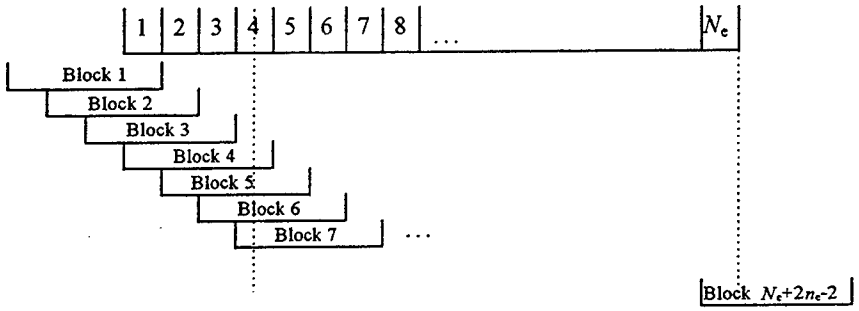


Figure 5. The new neural network structure.

Figure 6 illustrates the arrangement of the overlapped rectangular pulse functions, and the computation of f_{ij} given an input x_j. The range of each input variable (e.g., x_j) is divided into N_e

elements and every n_e neighboring elements are grouped into a block. Each block is assigned a basis function, which may be a bell-shaped function, a cubic spline function, a triangular function or a rectangular pulse function. With the rectangular pulse function, while each element is covered by n_e blocks, the computation of $f_{ij}(x_{sj})$ is simply the addition of weights associated to the n_e blocks covering the specific x_{sj}. The arrangement makes the learning for one element also alter the values in the neighborhood. This creates the generalization capability. As shown in Figure 6, there are N_e+2n_e-2 blocks. Thus the memory size for each variable in each submodule should equal $N_e+2*(n_e-1)$, which is usually small (typically 20 to 200). For N_p submodules and N_v variables, the total memory size required will be $N_p*N_v*(N_e+2n_e-2)$. A small memory size of 20k is required for a structure with 20 submodules, 10 input variables and 100 blocks (i.e., $N_e+2n_e-2 = 100$). The requirement of a small memory makes this scheme easy to implement and very attractive. The ease of implementation, good learning behavior, and the capability of handling problems with a high dimensional input space should make the structure adequate for many PVI and decision aids problems.



$f_{ij}(x_j$ in element 4) = sum of weights for blocks 4, 5, 6 and 7

Figure 6. Blocks for overlapped rectangular pulses and the computation of f_{ij} ($n_e = 4$ in this illustration).

6. Discussion and Conclusion

This report gives an overview on the pilot-vehicle interface technology and discusses some potential areas to which neural network technology can be applied. These areas include the use of neural networks for rule development and possible replacement of part of an expert system, recommending or performing air combat maneuvering, voice input, and problems involving

classification and interpolation. The main advantage of example-based learning is that explicit knowledge is not required. A special neural network structure that can handle problems with higher-dimensional input space is introduced. Due to the capability in handling higher-dimensional problems, we see potential usage of the new structure in the above identified areas.

References

1. Rumelhart, D. E., Hinton, G. E. & Williams, R. J., "Learning Internal Representation by Error Propagation," In Rumelhart, D. E. & McClelland, J. L. (Eds), *Parallel Distributed Processing: Exploration in the Microstructure of Cognition,* Vol. 1, pp.318-362. Cambridge, MA: MIT Press.
2. Hornik, K., Stinchcombe, M. & White, H., "Multi-layer Feedforward Networks are Universal Approximators," *Neural Networks*, Vol. 2, pp. 359-366, 1989.
3. Lee, S. & Kil, R. M., "A Gaussian Potential Function Network with Hierarchically Self-Organizing Learning," *Neural Networks*, Vol. 4, pp. 207-224, 1991.
4. Leonard, J. A., Kramer, M. A. & Ungar, L. H., "Using Radial Basis Functions to Approximate a Function and Its Error Bounds," *IEEE Transactions on Neural Networks*, Vol. 3, pp. 624-627, 1992.
5. Albus, J. S., "A New Approach to Manipulator Control: The Cerebellar Model Articulation Controller (CMAC)," *Journal of Dynamic Systems, Measurement, and Control, Transaction of ASME*, pp. 220-227, 1975.
6. Zhang, Q. & Benveniste, A., "Wavelet Network," *IEEE Transactions on Neural Network.* Vol. 3, pp. 889-898, 1992.
7. Cheng, Y. H. & Lin, C. S., "A Learning Algorithm for Radial Basis Function Networks: with Capability of Adding and Pruning Neurons," *Proceedings of IEEE International Conference on Neural Networks, 1994.*
8. Chiang, C. T. & Lin, C. S., "CMAC with General Basis Functions," to appear on *Neural Networks*.
9. Pao, Y. H., *Adaptive Pattern Recognition and Neural Networks*, Addison Wesley, Reading, MA, 1989.
10. Harrison, L, J. Janowitz, and M. Castronuovo, *The Pilot-Vehicle Interface*, Report DOT/FAA/CT-92/21, U. S. Dept. of Transportation, FAA Technical Center, Atlantic City International Airport, N.J., Nov. 1993.
11. Aeronautics and Space Engineering Board, and National Research Council, *Aeronautical Technologies for the Twenty-First Century*, National Academy Press, Washington, D. C., 1992.
12. Baber C., "Automatic Speech Recognition in Adverse Environments," *Human Factors*, 38(1), pp. 142-155, 1996.
13. Barry, T. P., T. J. Solz, and J. M. Reising, "Connected Speech Study for Cockpit Applications," WL-TR-95-3008, Flight Dynamics Directorate, Wright Laboratory, Feb. 1994.
14. Hammer, J. M. and R. L. Small, "An Intelligent Interface in an Associate System", in *Human/Technology Interaction in Complex Systems*, editor: W. B. Rouse, pp. 1-44. Vol. 7, 1995.
15. Wiener, E. L. and D. C. Nagel (eds), *Human Factors in Aviation*, Academic Press, Inc, San Diego, CA, 1988.

16. Sima, Jiri, "Neural Expert Systems," *Neural Networks*, Vol. 8, No. 2, pp. 261-271, 1995.
17. McMillan, C., M. C. Mozer and P. Smolensky, "Learning Explicit Rules in a Neural Network," *Proceedings of IJCNN*, Seatle, Vol. 2, pp. 83-88, 1991.
18. Samad, T., "Towards connectionist rule-based Systems," *Proceedings of IEEE Conference on Neural Networks*, San Diego, Vol. 2, pp. 525-532, 1988.
19. Yang Q., and V. K. Bhargava, "Building Expert Systems by a Modified Perceptron Network with Rule -transfer Algorithms," *Proceedings of IJCNN*, San Diego, pp.77-82, 1990
20. Stottler, R. H., and A. L. Henke, *Intelligent TDA Systems using Neural Networks*, Report TEC-0007, U. S. Army Corps of Engineers, Topographic Engineering Center, Fort Belvoir, Virginia, May 1992.
21. Schvaneveldt, R. W., A. E. Benson, T. E. Goldsmith, W. L. Wang, *Neural Network Models of Air Combat Maneuvering*, AL-TR-1992-0037, Armstrong Laboratory, 1992.
22. Frasconi P., M. Gori, M. Maggini, and G. Soda, "An Unified Approach for Integrating Explicit Knowledge and Learning by Example in Recurrent Networks," *Proceedings of IJCNN*, Seattle, Vol. 1, pp. 811-816, 1991
23. Handelman, D. A., S. H. Lane and J. J. Gelfand, "Integration of Knowledge-based System and Neural Network Techniques for Autonomous Learning Machines," *Proceedings of IJCNN*, Washington, Vol. 1, pp. 683-688, 1989
24. Hruska, S. I., D. C. Kuncicky and R. C. Lacher, "Hybrid Learning in Expert Networks," *Proceedings of IJCNN*, Seatle, Vol. 2, pp. 117-120, 1991
25. Saito K., and R. Nakano, "Medical Diagnostic Expert System Based on PDP Model," *Proceedings of IEEE Conference on Neural Networks*, San Diego, Vol. 1, pp. 255-261, 1988.
26. Enbutsu, I., K. Baba, and N. Hara, "Fuzzy Rule Extraction from a Multilayer Neural Network," *Proceedings of IJCNN*, Seatle, Vol. 2, pp. 461-465, 1991.
27. Harmon, M. E., and L. C. Baird "Multi-agent Residual Advantage Learning with General Function Approximation," WL-TR-96-1065, Avionics Directorate, Wright Laboratory, April 1996.
28. Harmon, M. E., L. C. Baird, and A. H. Klopf, "Advantage Updating Applied to a Differential Game," In Tesauro, G, et al (eds.), Advances in Neural Information Processing Systems 7, MIT Press, Cambridge MA, 1995.
29. Rabiner, L. R., and B. H. Juang, "Hidden Markov Models for Speech Recognition - Strengths and Limitations," in Laface P., and R. De Mori (eds), *Speech Recognition and Understanding* pp. 3-29, Springer-Verlag, 1992.
30. Bourlard, H. A., and N. Morgan, Connectionist Speech Recognition, Kluwer Academic Publisher, Boston, 1994.
31. Bridle, J. S., "Neural Networks or Hidden Markov Models for Automatic Speech Recognition: Is There a Choice," in Laface P., and R. De Mori (eds), *Speech Recognition and Understanding*, pp. 225-257, Springer-Verlag, 1992.
32. Lin, C. S. and C. K. Li, "A Sum-of-Product Form Neural Network (SOPNN)," submitted to *Neural Networks*.

Control of Linear Systems with Saturating Actuators – with Applications to Flight Control Systems

Zongli Lin
Assistant Professor
Department of Applied Mathematics and Statistics

State University of New York at Stony Brook
Stony Brook, NY 11794-3600

Final Report for
AFOSR Summer Faculty Research Program
Wright Laboratory/Control and Dynamics Directorate
Wright-Patterson Air Force Base, OH 45433-7531

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

and

Wright Laboratory

August 1996

Control of Linear Systems with Saturating Actuators – with Applications to Flight Control Systems

Zongli Lin
Assistant Professor
Department of Applied Mathematics and Statistics
State University of New York at Stony Brook

Abstract

Two problems in the control of linear systems with saturating actuators are treated, resulting in two design techniques.

1. A composite nonlinear feedback design based on a nominal linear feedback is proposed for linear systems with position limited actuators. The design yields nonlinear feedback laws that both increase the speed of closed-loop system response to the command input and reduce the overshoot, while not imperiling the performance achieved by the nominal linear feedback controller in the face of actuator position saturation. A flight control system is used to demonstrate the design.
2. A robust stabilizing feedback design technique for linear systems with rate limited actuators is proposed. The design combines two design techniques recently developed for linear systems with position limited actuators, piecewise-linear LQ control and low-and-high gain feedback. An F-16 class open-loop unstable fighter aircraft model is used to demonstrate the effectiveness of the proposed design method. The combined design takes advantages of both design techniques, while avoiding their disadvantages.

These two design techniques are presented as two separate parts of this final report.

Part 1. Toward Improvement on Tracking Performance – Nonlinear Feedback for Linear Systems

1.1. Introduction and Problem Statement

In classical control theory, performance specifications are often given in terms of step response. In deed, many command inputs are superposition of unit steps. It is well known that the behavior of the step response of a linear system is closely related to the damping ratio of the closed-loop system (see, for example, [9] and [4]). In particular, to ensure a quick approach to the command input, the damping ratio should not be too large. On the other hand, if the damping ratio is too small, there will be rather persistent oscillations with high peaks. A compromise is usually made in choosing the damping ratio. The goal here is to propose a *nonlinear* feedback design method that will take advantages of both the quick response of systems with small damping ratio and the small overshoot of systems with large damping ratio. At the same time, the actuator saturation will be taken into account at the onset of our design. As can be expected, the control design becomes more complicated in the presence of actuator saturation.

More specifically, we consider a second order single input single output linear system,

$$\begin{cases} \dot{x} = Ax + B\text{sat}(u), \quad x \in \mathbb{R}^2 \\ y = Cx \end{cases} \tag{1.1}$$

where sat : $\mathbb{R} \to \mathbb{R}$ represents actuator saturation and without loss of generality is assumed to have a unity saturation level, i.e., $\text{sat}(u) = \text{sign}(u)\min\{1, |u|\}$. We make the following assumptions on the system matrices (A, B, C).

Assumption 1.1.

1. *The pair (A, B) is controllable and hence, without loss of generality, can be assumed to be in the following form*
$$A = \begin{bmatrix} 0 & 1 \\ a_1 & a_2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
2. *The matrix C is given by $C = [c_1 \quad c_2]$, where $c_1 \neq 0$.*

Remark 1.1. *The assumption of controllability (rather than stabilizability) is made to allow for pole placement. The assumption on the matrix C is made to allow for the possibility of the output tracking a constant command input. In fact, it can be easily verified that in the absence of actuator saturation, the DC gain of the closed-loop system under any linear feedback $u = f_1x_1 + f_2x_2 + Gr$, where r is the command input, is $-Gc_1/(a_1 + f_1)$.*

Our goal is to design a nonlinear feedback law that will cause the output to track a step input rapidly without experiencing large overshoot. We propose a composite control law that is composed by combining a linear and a nonlinear feedback law. The design is sequential. First, a linear feedback law is designed that yields a closed-loop system of small damping ratio while at the same time not exceeding the actuator limits for the desired command input levels. (Evidently, under limited actuator power, arbitrarily large commands cannot be expected to be followed.) Then, utilizing an appropriate Lyapunov function for the closed-loop system under this linear feedback law, the nonlinear feedback law is constructed. The role of this nonlinear feedback law is to use larger feedback gains to cause the closed-loop system to be highly damped as the system output approaches the command input and thus reduce overshoot. Special care is taken to ensure that the ability of the linear feedback law to track a certain class of command inputs would not be imperiled when the actuator saturation occurs due to large control input from the nonlinear feedback law.

The remainder of this part is organized as follows. In Section 1.2, the design algorithm is presented. This design algorithm is applied to a flight control system in Section 1.3. The open-loop system in this example is exponentially unstable. Finally we draw a brief conclusion in Section 1.4.

1.2. Main Results

In this section, we provide an algorithm for the construction of composite nonlinear state feedback laws for the plant (1.1). The proposed feedback control law is derived by combining a linear and a nonlinear feedback law. The design is sequential. First a linear feedback law is designed and then, based on this linear feedback law, a nonlinear feedback law is constructed.

Step 1 – Linear Feedback Design. Choose a linear feedback law

$$u_{\text{L}} = Fx + Gr \tag{1.2}$$

where $F = [f_1 \quad f_2]$ is such that $A + BF$ is Hurwitz, the closed-loop system $C(sI - A - BF)^{-1}B$ has a small damping ratio, and $G = -(a_1 + f_1)/c_1$.

Once F is chosen, the level of the command input that can be followed is determined as follows.

Lemma 1.1. *Let $P > 0$ be such that*

$$(A + BF)^{\text{T}}P + P(A + BF) = -Q \tag{1.3}$$

where Q is any positive definite matrix. Such a P exists since $A + BF$ is Hurwitz. For later use, we partition P as

$$P = \begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix} \tag{1.4}$$

Let $c > 0$ be the largest positive number such that

$$x \in \{x : x^{\mathrm{T}} P x \leq c\} \Longrightarrow |Fx| \leq 1 - \Delta \tag{1.5}$$

where $\Delta \in (0, 1)$. Then, the linear feedback law (1.2) will cause the system output to asymptotically track a step command input of amplitude r, as long as the initial state $x(0)$ and r satisfy

$$\left(x_1(0) - \frac{r}{c_1}, x_2(0)\right) \in \{x : x^{\mathrm{T}} P x \leq c\}, \text{ and } \left|\frac{a_1}{c_1} r\right| \leq \Delta \tag{1.6}$$

Proof. Omitted due to space limitation. ∎

We note here that from the zero initial state, any step of amplitude r, $|r| \leq \sqrt{c_1^2 c / p_1}$ and $\left|\frac{a_1}{c_1} r\right| \leq \Delta$, can be asymptotically tracked. By increasing Δ and/or decreasing $|p_1|$ through choice of f_1, f_2 and Q, we can increase the input command amplitude that can be asymptotically tracked. This change of f_1 and f_2 will of course affect the closed-loop damping ratio and hence the response speed.

Step 2 – Nonlinear Feedback Design. The nonlinear state feedback law is given by

$$u_{\mathrm{N}} = -\rho(x, r) B^{\mathrm{T}} P \left(x - \begin{bmatrix} r/c_1 \\ 0 \end{bmatrix}\right) \tag{1.7}$$

where $\rho : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}_+$ is any nonnegative function locally Lipschitz in x, which is chosen to reduce the damping ratio as the output approaches the command input. The choice of such a $\rho(x, r)$ will be discussed shortly.

Step 3 – Composition of Linear and Nonlinear Feedback. The final nonlinear feedback law is then composed as

$$u = u_{\mathrm{L}} + u_{\mathrm{N}} = [F - \rho(x, r) B^{\mathrm{T}} P] x + \frac{-a_1 - f_1 + \rho(x, r) p_2}{c_1} r \tag{1.8}$$

With the nonlinear state feedback law as given by (1.8), we have the following result concerning the step response of the closed-loop system.

Theorem 1.1. *Consider the system given by (1.1). For any nonnegative function $\rho(x, r)$, locally Lipschitz in x, the nonlinear feedback law (1.8) will cause the system output to asymptotically track the step command input of amplitude r from an initial state $x(0)$, provided that $x(0)$ and r satisfy (1.6).*

Proof. Omitted due to space limitation. ∎

Finally, the freedom in choosing the nonlinear function is utilized to increase the "damping ratio" of the closed-loop system as the output approaches the command input. To guide the choice of such a function, we establish the following lemma.

Lemma 1.2. *Consider the single input single output linear system*

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases} \tag{1.9}$$

where

$$A = \begin{bmatrix} 0 & 1 \\ a_1 & a_2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Let A be Hurwitz and $P > 0$ be the unique solution to the Lyapunov equation,

$$A^{\mathrm{T}}P + PA = -Q \tag{1.10}$$

where Q is any positive definite matrix. Let the linear feedback law be $u = -\rho B^{\mathrm{T}}Px$, $\rho \geq 0$. Then, as the constant ρ tends to $+\infty$, one of the two closed-loop poles approaches $-\infty$ and the other a finite location in the open left half plane, and hence the damping ratio of the closed-loop system tends to $+\infty$.

Proof. Omitted due to space limitation. ∎

Guided by the above lemma, to retain the fast response of system with small damping ratio, and at the same time, to reduce the damping ratio as the output approaches the command input, $\rho(x, r)$ can chosen to be a non-increasing function of $|y - r|$ such that,

$$\lim_{|y-r|\to|y(0)|} \rho(x, r) = 0 \text{ and } \lim_{|y-r|\to 0} \rho(x, r) = \kappa$$

where κ is a (large) positive constant. As will be seen in the next section, for moderately large values of function $\rho(x, r)$, the actuator does not saturate since high gain occurs only in small signal operation. However, our design allows for arbitrary choice of $\rho(x, r)$, and when the control signal is larger than 1 because of large value of $\rho(x, r)$, the value of $\rho(x, r)$ is automatically clipped off by the actuator.

1.3. Application to Flight Control

In this section, the proposed nonlinear feedback design is applied to an F-16 aircraft derivative. At the flight condition corresponding to an altitude of 20,000 feet and a Mach number of 0.9, the short period dynamics of this aircraft model are given by (1.1) (see [1]) with

$$A = \begin{bmatrix} 0 & 1.0000 \\ 0.3730 & -2.1840 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = [-32.8781 \quad -24.3282] \tag{1.11}$$

where the output is the pitch rate and the control input is the elevator deflection scaled such that the maximum actuator capacity is unity. The physical actuator control authority for maneuvering flight is $\pm 12^\circ$ of deflection. The pitch rate and the commanded pitch rate r are scaled accordingly.

Following the proposed design procedure, we choose $f_1 = -2.8066 \quad f_2 = 0, \quad G = -0.0740$, which gives a closed-loop system damping ratio of 0.7. Let $Q = -I$ and solve (1.3) for P, yielding

$$P = \begin{bmatrix} 1.2348 & 0.2055 \\ 0.2055 & 0.3230 \end{bmatrix} \tag{1.12}$$

Pick $\Delta = 0.5$. Then the largest $c > 0$ for (1.5) to hold is 0.0350. Hence, by Lemma 1.1, the linear feedback control law

$$u_{\mathrm{L}} = -2.8066x_1 + 0.0740r \tag{1.13}$$

will cause the system to asymptotically track a step command input of amplitude r, as long as the initial state $x(0)$ and r satisfy (1.6), i.e.,

$$\left(x_1(0) + \frac{r}{32.8781}, x_2(0)\right) \in \{x : x^{\mathrm{T}}Px \leq 0.0351\}, \text{ and } |r| \leq 44.0725 \tag{1.14}$$

This implies that from the zero initial state any step command input of amplitude no greater than $\min\{5.5353, 44.0725\} = 5.5353$ can be asymptotically tracked. This translates to a commanded pitch rate of $5.5353 \times 12 \approx 66^{o}/\mathrm{sec}$, which shows that the allocated maneuvering control authority is more than adequate at this flight condition.

Finally, we choose $\rho(x, r)$ as $\rho(x, r) = 5e^{-|y-r|}$. The final nonlinear feedback law is then given by

$$u = -[2.8066 + 0.2055\rho(x, r)]x_1 - 0.3230\rho(x, r)x_2 - [0.0740 + 0.00625\rho(x, r)]r \tag{1.15}$$

Some simulation results of the above flight control system are shown in Figs. 1.1 and 1.2. We see from the simulation results that the overshoot in the response due to the nominal linear feedback is avoided and hence better tracking performance is obtained.

Given the ability of the proposed nonlinear feedback to suppress overshoot of the response due to the nominal linear feedback, one could intentionally design a nominal linear feedback that results in smaller damping ratio to increase the response speed and use the nonlinear feedback to suppress the resulting larger overshoot. To this end, we redesign the linear feedback law with higher gain of $f_1 = -7.8259$, $f_2 = 0$ and $G = -0.2267$. This linear feedback law results in a damping ratio of 0.4.

Let $Q = -I$ and solve (1.3) for P, yielding

$$P = \begin{bmatrix} 2.0817 & 0.0671 \\ 0.0671 & 0.2597 \end{bmatrix} \tag{1.16}$$

Pick $\Delta = 0.5$. Then the largest $c > 0$ for (1.5) to hold is 0.0084. Hence, by Lemma 1.1, the linear feedback control law

$$u_{\mathrm{L}} = -7.8259x_1 + 0.2267r \tag{1.17}$$

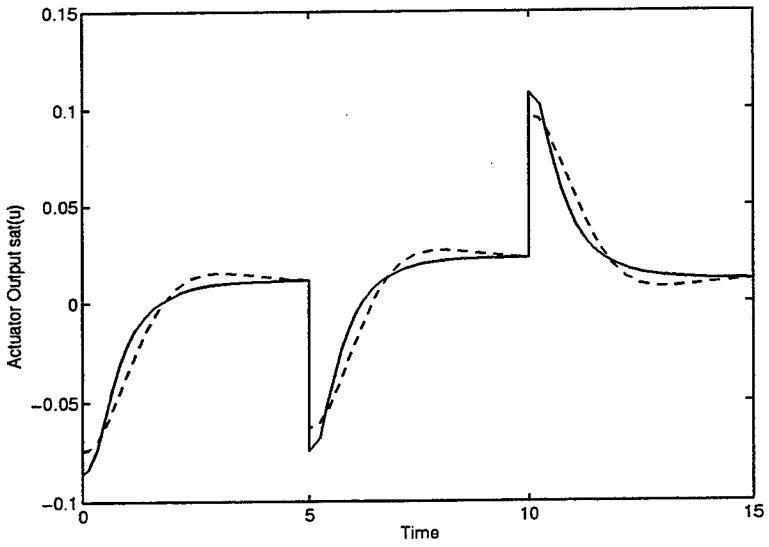


Figure 1.1: Actuator output. The solid line represents the actuator output due to the nonlinear feedback, and the dashed line the nominal linear feedback ($\rho(x, r) = 0$).

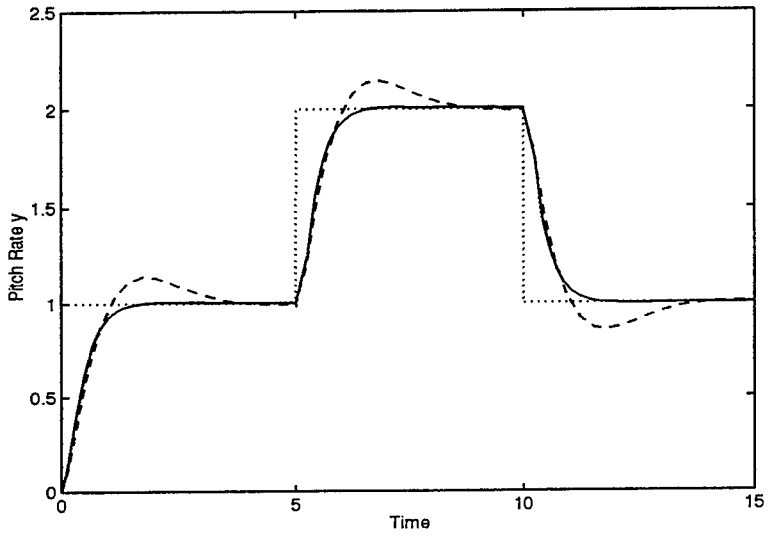


Figure 1.2: Tracking performance. The dotted line is the command input, the solid line represents the response due to the nonlinear feedback, and the dashed line the nominal linear feedback ($\rho(x, r) = 0$).

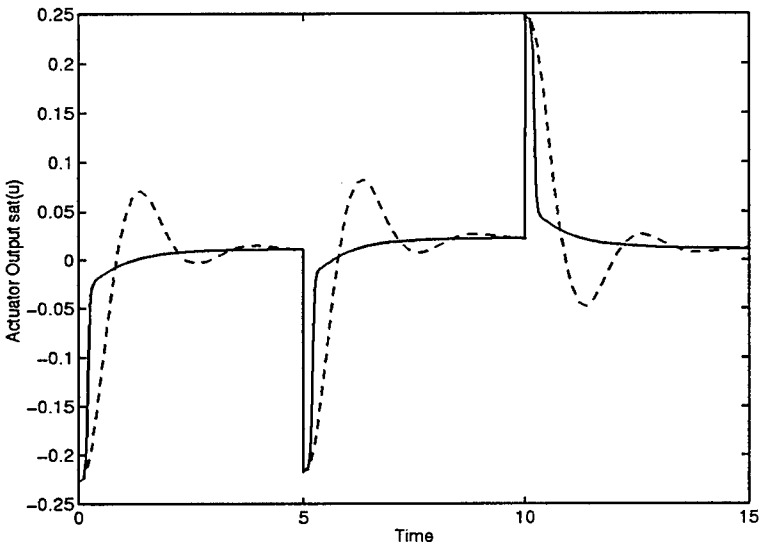


Figure 1.3: Actuator output. The solid line represents the actuator output due to the nonlinear feedback, and the dashed line the nominal linear feedback ($\rho(x,r)=0$).

will cause the system to asymptotically track a step command input of height r, as long as, the initial state $x(0)$ and r satisfy

$$\left(x_1(0)+\frac{r}{32.8781},x_2(0)\right)\in\{x:x^{\mathrm{T}}Px\leq 0.0084\},\text{ and }|r|\leq 44.0725 \tag{1.18}$$

This implies that from the zero initial state any step command input of amplitude no greater than $\min\{2.0885,44.0725\}=2.0885$ can be asymptotically tracked. In other words, as expected, the slewing amplitude is now reduced to $2.0885\times 12\approx 25^o/\text{sec}$, which is still very adequate.

Choosing $\rho(x,r)$ as $\rho(x,r)=30e^{-10|y-r|}$. The final nonlinear feedback law is then given by

$$u=-[7.8259+0.0671\rho(x,r)]x_1-0.2597x_2-[0.2267+0.00204\rho(x,r)]r \tag{1.19}$$

Some simulation results are shown in Figs. 1.3 and 1.4. Comparing Figs. (1.4) and (1.2), it is evident that the closed-loop system performance is further improved in that its response is faster while still avoiding overshoot.

1.4. Conclusion

Nonlinear feedback laws were constructed for linear systems. These nonlinear feedback laws yield fast closed-loop response without inducing overshoot. The actuator saturation was also taken into account at the onset of the control design and hence does not require any *a posteriori* compensation. Note, however, that the proposed control concept is not a saturation avoidance scheme.

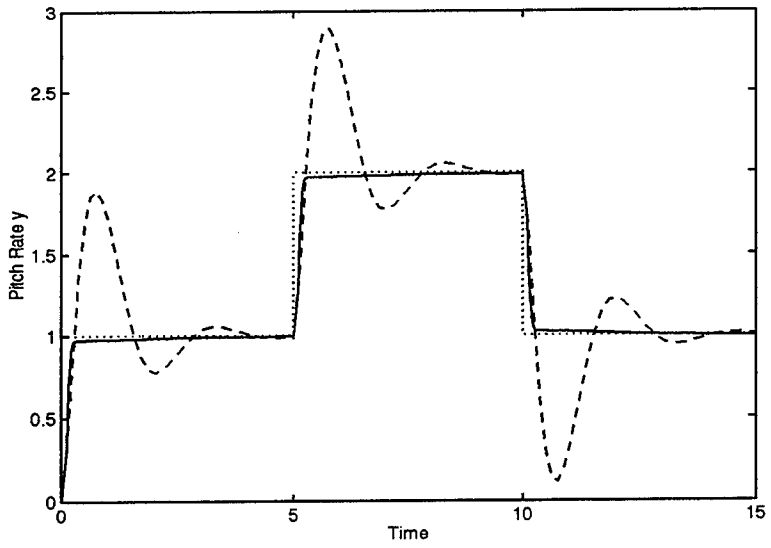


Figure 1.4: Tracking performance. The dotted line is the command input, the solid line represents the response due to the nonlinear feedback, and the dashed the line nominal linear feedback ($\rho(x,r)=0$).

Part 2. Stabilizing Feedback Design for Linear Systems with Rate Limited Actuators

2.1. Introduction

Every physical actuator is subject to constraints. These constraints include both position and rate saturation. In the past few years there has been much interest concerning stabilization of linear systems with position saturating actuators, resulting in several promising design techniques. In this paper, we will recourse to the low-and-high gain (LHG) design technique ([8]) and the piecewise linear LQ control (PLC) design technique ([10]) in this section. Additional related work might be found in these two papers.

In comparison with actuator position saturation, few design techniques are currently available to deal with actuator rate saturation. However, actuator rate saturation often presents a more serious challenge to control engineers, especially flight control engineers. It is known ([2]) that actuator rate saturation could induce a considerable phase-lag. Such phase-lag associated with rate saturation has a destabilizing effect. For example, investigators have identified rate saturation as a contributing factor to the recent mishaps of YF-22 [3] and Gripen [5] prototypes and the first production Gripen [7]. For further discussion on the destabilizing effect of actuator rate saturation, see [2].

The goal here is to propose a method of designing stabilizing feedback control laws for linear systems taking into account the effect of actuator rate saturation. The proposed design method

views the problem of stabilization with rate saturating actuators as a problem of robust stabilization with position saturating actuators in the presence of input additive uncertainties. The state feedback law is then designed for the stabilization with position saturating actuators. Our state feedback design combines the two recently developed design techniques, the PLC and the LHG design techniques. It inherits the advantages of the both design techniques, while avoiding their disadvantages. Thus, the exact knowledge of the dynamics of the actuators will not be needed and the actuator disturbances can be rejected. In particular, in the LHG design, a low gain feedback law is first designed in such a way that the actuator does not saturate in position and the the closed-loop system remains linear. The gain is chosen low to enlarge the region in which the closed-loop system remains linear and hence enlarge the domain of attraction of the closed-loop system. Then, utilizing an appropriate Lyapunov function for the closed-loop system under this low gain feedback control law, a linear high gain feedback control law is constructed and added to the low gain feedback control to form the final LHG feedback control law. Such a linear low-and-high gain feedback control law speeds up the transient response for the states in a certain subspace of the state space and is capable of stabilizing the system in the presence of input-additive plant uncertainties and rejecting arbitrarily large bounded input-additive disturbances. The disadvantage of this control law is that the transient speed for the states outside that certain subspace of the state space remains that of the low gain feedback, which is typically sluggish due to low feedback gain for a large domain of attraction. On the other hand, the aim of the PLC scheme is to increase the feedback gain piecewisely while adhering to the input bound as the trajectories converge toward the origin. Such a design results in fast transient speed for all states. However, it lacks robustness to large uncertainties and the ability of rejecting arbitrarily large bounded disturbances.

The remainder of this part is organized as follows. In Section 2.2, we precisely formulate our problems. Section 2.3 provides the design algorithm and proves that the proposed algorithm results in feedback laws that solve the problems formulated in Section 2.2. In Section 2.4, an F-16 class fighter aircraft model is used to demonstrate the effectiveness of the proposed design algorithm. A brief concluding remark is given in Section 2.5.

2.2. Problem Formulation

Consider the linear dynamical system

$$\begin{cases} \dot{x} = Ax + Bv, \;\; x(0) \in \mathcal{X} \subset \mathbb{R}^n \\ \dot{v} = \mathrm{sat}_\Delta(-T_1 v + T_2 u + d), \;\; v(0) \in \mathcal{V} \subset \mathbb{R}^m \end{cases} \tag{2.1}$$

where $x \in \mathbb{R}^n$ is the plant state, $v \in \mathbb{R}^m$ is the actuator state and input to the plant, $u \in \mathbb{R}^m$ is the control input to the actuators, for $\Delta = (\Delta_1, \Delta_2, \cdots, \Delta_m)$, $\Delta_i > 0$, the function $\mathrm{sat}_\Delta : \mathbb{R}^m \to \mathbb{R}^m$ is the standard saturation function that represents actuator rate saturation, i.e.,

$\text{sat}_\Delta(v) = [\text{sat}_{\Delta_1}(v_1), \text{sat}_{\Delta_2}(v_2), \cdots, \text{sat}_{\Delta_m}(v_m)]$, $\text{sat}_{\Delta_i}(v_i) = \text{sign}(v_i)\min\{\Delta_i, |v_i|\}$, the positive definite diagonal matrices $T_1 = \text{diag}(\tau_{11}, \tau_{12}, \cdots, \tau_{1m})$ and $T_2 = \text{diag}(\tau_{21}, \tau_{22}, \cdots, \tau_{2m})$ represent the "time constants" of the actuators and are not precisely known, and finally $d : \mathbb{R}_+ \to \mathbb{R}^m$ ate the disturbances appearing at the input of the actuators.

We also make the following necessary assumptions on the system.

Assumption 2.1. *The pair (A, B) is stabilizable;*

Assumption 2.2. *The nominal value of T_1 and T_2 is known and is given by $T^* = diag(\tau_1^*, \tau_2^*, \cdots, \tau_m^*)$ with $\tau_i > 0$. There exist known matrices $\mu_1 = diag(\mu_{11}, \mu_{12}, \cdots, \mu_{1m})$, $\mu_2 = diag(\mu_{21}, \mu_{22}, \cdots, \mu_{2m})$, $|\mu_{ij}| \leq 1$, and $\nu_1 = diag(\nu_{11}, \nu_{12}, \cdots, \nu_{1m})$, $\nu_2 = diag(\nu_{21}, \nu_{22}, \cdots, \nu_{2m})$, $|\nu_{ij}| \geq 1$, such that, $\mu_1 T^* \leq T_1 \leq \nu_1 T^*$ and $\mu_2 T^* \leq T_2 \leq \nu_2 T^*$.*

Assumption 2.3. *The disturbance is uniformly bounded by a known (arbitrarily large) constant D, i.e., $\|d(t)\| \leq D, \forall t \geq 0$.*

Before stating our problem, we make the following preliminary assumption.

Theorem 2.1. *The data $(D, \mathcal{W}_0)$ is said to be* admissible for state feedback *if D is a nonnegative real number, and $\mathcal{W}_0$ is a subset of $\mathbb{R}^{n+m}$ which contains the origin as an interior point.*

The goal here is to provide a design algorithm for solving the following problem.

Problem 2.1. *Given the data $(D, \mathcal{W}_0)$, admissible for state feedback, find, if possible, a feedback law $u = F(x, v)$, such that the closed-loop system satisfies,*

1. *if $D = 0$, the point $x = 0$ is locally asymptotically stable and $\mathcal{X} \times \mathcal{V}$ is contained in its basin of attraction;*
2. *if $D > 0$, every trajectory starting from $\mathcal{X} \times \mathcal{V}$ enters and remains in $\mathcal{W}_0$ after some finite time.*

2.3. A Combined PLC/LHG Design Algorithm

As stated earlier, the proposed design algorithm is a combination of the piecewise linear LQ control [10] and the low-and-high gain feedback [8]. Naturally we organize this section as follows. Subsections 2.3.1 and 2.3.2 respectively recapitulate the PLC and the LHG design techniques. Subsection 2.3.3 presents the proposed combined PLC/LHG design algorithm. Finally, in Subsection 2.3.4, the proposed design algorithm is shown to solve Problem 2.1.

2.3.1. Piecewise Linear LQ Control Design

Consider the linear dynamical system subject to actuator position saturation,

$$\dot{x} = Ax + B\text{sat}_\Delta(u), \;\; x(0) \in \mathcal{X} \subset \mathbb{R}^n, \; u \in \mathbb{R}^m \tag{2.2}$$

where the saturation function $\text{sat}_\Delta : \mathbb{R}^m \to \mathbb{R}^m$ is as defined in Section 2.2, and the pair (A, B) is assumed to be stabilizable.

The PLC design is based on the following LQ algebraic Riccati equation,

$$A'P + PA - PBR^{-1}B'P + I = 0 \tag{2.3}$$

where $R = \text{diag}(\epsilon) = \text{diag}(\epsilon_1, \epsilon_2, \cdots, \epsilon_m)$, $\epsilon_i > 0$, are the design parameters to be chosen later.

Key to the PLC scheme is the notion of invariant sets. A nonempty subset of ε in $\mathbb{R}^n$ is positively invariant if for a dynamical system and for any initial condition $x(0) \in \varepsilon$, $x(t) \in \varepsilon$ for all $t \geq 0$. For the closed-loop system comprising of the system (2.2) and the LQ control $u = -R^{-1}B'Px$, simple Lyapunov analysis shows that the Lyapunov level set

$$\varepsilon(P, \rho) = \{x : x'Px \leq \rho\}, \; \forall \rho > 0$$

is an invariant set, provided that saturation does not occur for all $x \in \varepsilon(P, \rho)$. To avoid the saturation from occurring, while fully utilizing the available control capacity, for a given ρ, $\epsilon = (\epsilon_1, \epsilon_2, \cdots, \epsilon_m)$ will be chosen to be the largest such that

$$|u_i| = \left|\frac{1}{\varepsilon_i} B_i'Px\right| \leq \Delta_i, \;\; \forall x \in \varepsilon(P, \rho)$$

where B_i is the ith column of matrix B and u_i is the ith element of u. The existence and uniqueness of such an ϵ are established in [10] and an algorithm for computing such an ϵ is also given in [10]. More specifically, it is shown through the existence of a unique fixed point that the following iteration converges from any initial value to the desired value of ϵ,

$$\epsilon_{n+1} = \sqrt{\rho}\Phi(\epsilon_n) \tag{2.4}$$

where $\Phi(\epsilon) = [\phi_1(\epsilon), \phi_2(\epsilon), \cdots, \phi_m(\epsilon)]'$ and for each $i = 1$ to m,

$$\phi_i(\epsilon) = \frac{1}{\Delta_i}\sqrt{B_i'P(\epsilon)B_i}$$

The aim of the PLC scheme is to increase the feedback gain piecewisely while adhering to actuator bounds as the trajectories converge towards the origin. This is achieved by constructing nested level sets, $\varepsilon_0, \varepsilon_1, \cdots, \varepsilon_N$, in such a way that as the trajectories traverse successively the surface of each ε_i and the control law is switched to higher and higher gains as each surface is crossed.

The procedure in designing a PLC law is as follows. Given the set of initial conditions $\mathcal{X} \subset \mathbb{R}^n$, choose an initial level set ε_0 as,

$$\varepsilon_0 = \inf_{\rho}\{\varepsilon(P,\rho):\ \mathcal{X} \subset \varepsilon(P,\rho)\} \tag{2.5}$$

We denote the value of ρ associated with ε_0 as ρ_0, and the corresponding values of ϵ, R and P as ϵ_0, R_0 and P_0 respectively. A simple approach to determining ε_0 and ρ_0 can also be found in [10]. More specifically, it is shown that the size of ε_0 grows monotonically as the parameter ρ grows. Hence, ε_0 and ρ_0 can be determined by a simple iteration procedure. Here we would like to note that, as explained in [10], increasing ρ indefinitely for exponentially unstable A will not result in an ε_0 that grows without bound.

To determine the inner level sets ε_i's, choose successivefully smaller ρ_i where $\rho_{i+1} < \rho_i$ for each $i = 1, 2, \cdots, N$. A simple choice of such ρ_i's is the geometric sequence of the form

$$\rho_i = \rho_0(\Delta\rho)^i,\ \ i = 0, 1, 2, \cdots, N$$

where the ρ-reduction' factor $\Delta\rho \in (0,1)$. (Consequently, the values of ϵ, R and P associated with each of these ρ_i's are denoted as ϵ_i, R_i and P_i respectively.) For a discussion of the choice of N and $\Delta\rho$, see [10].

As shown in [10], a critical property of such a sequence of level sets ε_i is that they are nested in the sense that $\varepsilon_i \subset \varepsilon_{i+1}$ for each $i = 0$ to $N-1$.

2.3.2. Low-and-High Gain Feedback Design

Consider the linear dynamical system subject to actuator position saturation, input additive disturbances and uncertainties,

$$\dot{x} = Ax + B\text{sat}_\Delta(u + f(x) + d),\ \ x(0) \in \mathcal{X} \subset \mathbb{R}^n, u \in \mathbb{R}^m \tag{2.6}$$

where the saturation function $\text{sat}_\Delta : \mathbb{R}^m \to \mathbb{R}^m$ is as defined in Section 2.2, the locally Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^m$ represents the input additive plant uncertainties and d the input-additive disturbance. The LHG feedback design for this system is given as follows. First, the level set ε_0 is determined as in the PLC design. Correspondingly, a state feedback law with a possibly low feedback gain is determined as,

$$u_{\text{L}} = -R_0^{-1}B'P_0x$$

A high gain state feedback is then constructed as,

$$u_{\text{H}} = -kR_0^{-1}B'P_0x,\ \ k \geq 0$$

The final low-and-high gain state feedback is then given by a simple addition of the low and high gain feedbacks $u_{\rm L}$ and $u_{\rm H}$, viz.,

$$u = -(1+k)R_0^{-1}B'P_0x, \;\; k \geq 0 \tag{2.7}$$

Here the design parameter k is referred to as the high gain parameter. As demonstrated in [8] the freedom in choosing the value of this high gain parameter can be utilized to achieve robust stabilization in the presence of input additive plant uncertainties $f(x)$ and input-additive disturbance rejection. Moreover, the transient speed for the states not in the range space of $B'P_0$ will increase as the value of k increases. To see this, let us consider the following Lyapunov function, $V_0(x) = x'P_0x$. The evaluation of $\dot{V}$ along the trajectories of the closed-loop system in the absence of uncertainties and disturbances gives,

$$\begin{aligned}\dot{V} =& -x'x - x'P_0BR_0^{-1}B'P_0x + 2x'P_0B[\mathrm{sat}_\Delta(-(k+1)R_0^{-1}B'P_0x) + R_0^{-1}B'P_0x] \\ =& -x'x - x'P_0BR_0^{-1}B'P_0x - 2\sum_{i=1}^{m} v_i[\mathrm{sat}_{\Delta_i}((k+1)v_i) - v_i]\end{aligned}$$

where we have denoted the ith element of $v = -R_0^{-1}B'P_0x$ as v_i. By the choice of P_0, it is clear that $|v_i| \leq \Delta_i$ and hence $-v_i[\mathrm{sat}_{\Delta_i}((k+1)v_i) - v_i] \leq 0$, for each $i = 1$ to m. For any x not in the range space of $B'P_0$, that is $B'P_0x \neq 0$, then, for any i such that $v_i \neq 0$, $-v_i[\mathrm{sat}_{\Delta_i}((k_2+1)v_i) - v_i] < -v_i[\mathrm{sat}_{\Delta_i}((k_1+1)v_i) - v_i]$ if $k_2 > k_1$. However, for any x in the null space of $B'P_0$, $-v_i[\mathrm{sat}_{\Delta_i}((k+1)v_i) - v_i] = 0$ for all i.

2.3.3. Combined PLC/LHG Design

In this subsection, we present the proposed combined PLC/LHG state feedback design for linear systems subject to actuator rate saturation (2.1).

The feedback control law design is carried out in the following three steps.

Step 1. Choose a pre-feedback

$$u = v + \bar{u} \tag{2.8}$$

Let $\tilde{x} = [x, v]'$. Then the system (2.1) under the above pre-feedback is given by,

$$\dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{B}\mathrm{sat}_\Delta(T_2u + f(\tilde{x}) + d), \;\; \tilde{x}(0) \in \mathcal{X} \times \mathcal{V} \subset \mathbb{R}^{n+m} \tag{2.9}$$

where $f(\tilde{x}) = [\,0 \quad -T_1 + T_2\,]\,\tilde{x}$ and,

$$\tilde{A} = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}, \;\; \tilde{B} = \begin{bmatrix} 0 \\ I \end{bmatrix}$$

Assumption 2.1, i.e., the pair (A, B) is stabilizable, implies that $(\tilde{A}, \tilde{B})$ is stabilizable.

Step 2. Apply the PLC design to the system (2.9), and obtain a sequence of nested level sets $\varepsilon_0, \varepsilon_1, \cdots, \varepsilon_N$ (and correspondingly, the parameters $\epsilon_0, \epsilon_2, \cdots, \epsilon_N$) and a piecewise linear feedback law,

$$\bar{u} = \begin{cases} \bar{u}_i = -(\mu_2 T^*)^{-1} \tilde{R}_i^{-1} \tilde{B}' \tilde{P}_i \tilde{x} & \text{for } \tilde{x} \in \varepsilon_i \setminus \varepsilon_{i+1},\ i = 0, 1, \cdots, N-1 \\ \bar{u}_N = -(\mu_2 T^*)^{-1} \tilde{R}_N^{-1} \tilde{B}' \tilde{P}_N \tilde{x} & \text{for } \tilde{x} \in \varepsilon_N \end{cases} \tag{2.10}$$

Step 3. Design the LHG state feedback based on the PLC feedback law (2.10) and obtain the following combined final PLC/LHG feedback law,

$$u = \begin{cases} u_i = v - (k+1)(\mu_2 T^*)^{-1} \tilde{R}_i^{-1} \tilde{B}' \tilde{P}_i \tilde{x} & \text{for } \tilde{x} \in \varepsilon_i \setminus \varepsilon_{i+1},\ i = 0, 1, \cdots, N-1 \\ u_N = v - (k+1)(\mu_2 T^*)^{-1} \tilde{R}_N^{-1} \tilde{B}' \tilde{P}_N \tilde{x} & \text{for } \tilde{x} \in \varepsilon_N \end{cases} \tag{2.11}$$

where $k \geq 0$ is a design parameter to be specified later.

2.3.4. Proof

In what follows, we will show as a theorem that the combined PLC/LHG state feedback law (2.11) solves Problem 2.1. The effectiveness of this feedback law in comparison with both the PLC and the LHG feedback laws will be demonstrated in Section 2.4.

Theorem 2.1. *Let Assumption 2.1 hold. Given the admissible data $(D, \mathcal{W}_0)$, there exists a $k^*(D, \mathcal{W}_0) \geq 0$ such that, for all $k \geq k^*$, the combined PLC/LHG state feedback law (2.11) solves Problem 2.1. Moreover, if $D = 0$, k^* is independent of $\mathcal{W}_0$.*

Proof. Omitted due to space limitation. ∎

2.4. Stabilizing Feedback Design for an F-16 Aircraft

In this section, the applicability and effectiveness of the proposed combined PLC/LHG design algorithm is demonstrated with an F-16 fighter aircraft derivative. At the flight condition corresponding to an altitude of 10,000 feet and a Mach number of 0.7, the second order pitch plane dynamics (short period) of this aircraft is given by ([6]),

$$\begin{cases} \begin{bmatrix} \dot{\alpha} \\ \dot{q} \end{bmatrix} = \begin{bmatrix} -1.1500 & 0.9937 \\ 3.7240 & -1.2600 \end{bmatrix} \begin{bmatrix} \alpha \\ q \end{bmatrix} + \begin{bmatrix} -0.1770 \\ -19.5000 \end{bmatrix} \delta, \quad |\alpha(0)| \leq 0.1\text{rad}, \ |q(0)| \leq 1\text{rad/sec} \\ \dot{\delta} = \text{sat}_1(-T_1 \delta + T_2 u + d), \quad |\delta(0)| \leq 0.1\text{rad} \end{cases}$$

where α, q and δ are respectively the trim value of the angle of attack, pitch rate and stabilizer deflection. Here the actuator is rate limited to ± 1.0 rad/sec. The constants T_1 and T_2 have a same nominal value of 20 1/sec, which corresponds to a nominal actuator bandwidth of 20 rad/sec. However, the actual values of T_1 and T_2 vary and lie within $\pm 5\%$ of the nominal value, i.e., $\mu_1 = \mu_2 = 0.95$ and $\nu_1 = \nu_2 = 1.05$.

It is straightforward to see that the above system is in the form of (2.1). Our goal here is to design an effective robust stabilizing feedback law that will also be able to reject the actuator disturbance d. The feedback law we are to design is the combined PLC/LHG feedback law proposed in the previous section.

Following the design procedure given in Section 2.3.3, we first find,

$$\rho_0 = 1.2335,\ \epsilon_0 = 5.0805,\ P_0 = \begin{bmatrix} 1.0456 & 0.4868 & -3.3592 \\ 0.4868 & 0.4185 & -2.1537 \\ -3.3592 & -2.1537 & 20.9250 \end{bmatrix}$$

Choosing $N = 5$ and $\Delta\rho = 0.6$, we obtain a feedback law of the form (2.11) with,

$$\rho_1 = 0.7401,\ \epsilon_1 = 3.3505,\ P_1 = \begin{bmatrix} 0.9130 & 0.4107 & -2.5622 \\ 0.4107 & 0.3737 & -1.7118 \\ -2.5622 & -1.7118 & 15.1680 \end{bmatrix}$$

$$\rho_2 = 0.4441,\ \epsilon_2 = 2.2150,\ P_2 = \begin{bmatrix} 0.8083 & 0.3486 & -1.9619 \\ 0.3486 & 0.3357 & -1.3696 \\ -1.9619 & -1.3696 & 11.0484 \end{bmatrix}$$

$$\rho_3 = 0.2664,\ \epsilon_3 = 1.4675,\ P_3 = \begin{bmatrix} 0.7252 & 0.2975 & -1.5072 \\ 0.2975 & 0.3030 & -1.1021 \\ -1.5072 & -1.1021 & 8.0825 \end{bmatrix}$$

$$\rho_4 = 0.1599,\ \epsilon_4 = 0.9741,\ P_4 = \begin{bmatrix} 0.6592 & 0.2552 & -1.1613 \\ 0.2552 & 0.2746 & -0.8912 \\ -1.1613 & -0.8912 & 5.9355 \end{bmatrix}$$

$$\rho_5 = 0.0959,\ \epsilon_5 = 0.6477,\ P_5 = \begin{bmatrix} 0.6066 & 0.2199 & -0.8972 \\ 0.2199 & 0.2496 & -0.7235 \\ -0.8972 & -0.7235 & 4.3736 \end{bmatrix}$$

We note here again that this control law reduces to an LHG feedback control law if $N = 0$ and to a PLC law if $k = 0$ and $\mu_1 = \mu_2 = \nu_1 = \nu_2 = 1$.

Figs. 2.1 and 2.2 are some simulation results. In these simulations, $T_1 = 21$, $T_2 = 19$, $d = 5\sin(10t + 3)$. These simulations show good utilization of the available actuator rate capacity, and the degree of disturbance rejection increases as the value of the parameter k increases. Higher inaccuracy in the actuator parameter can be tolerated by further increasing the value of the design parameter k.

We next compare the performance of the combined PLC/LHG design with both the PLC and the LHG design. To compare with the PLC laws, we set $\mu_2 = 1$ and $k = 0$ in (2.11). Simulation (Fig. 2.3) shows that even in the absence of a disturbance, the aircraft goes unstable due to the presence

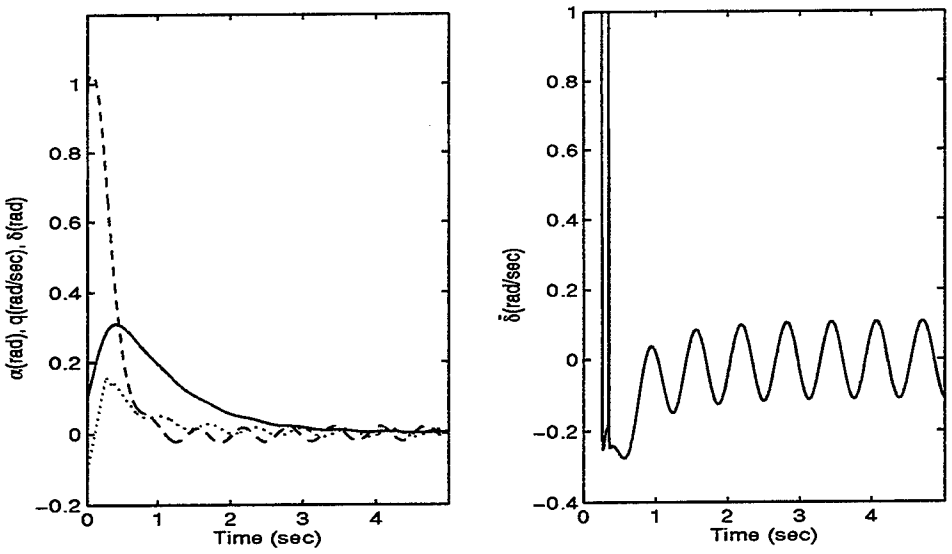


Figure 2.1: Combined PLC/LHG design ($N = 5$, $k = 64$). $T_1 = 21$, $T_2 = 19$, $d = 5\sin(10t + 3)$. The left plot: states of the plant and the actuator, the solid line is the angle of attack α, the dashed line the pitch rate q, and the dotted line the deflection δ; the right plot: the actuator rate $\dot{\delta}$.

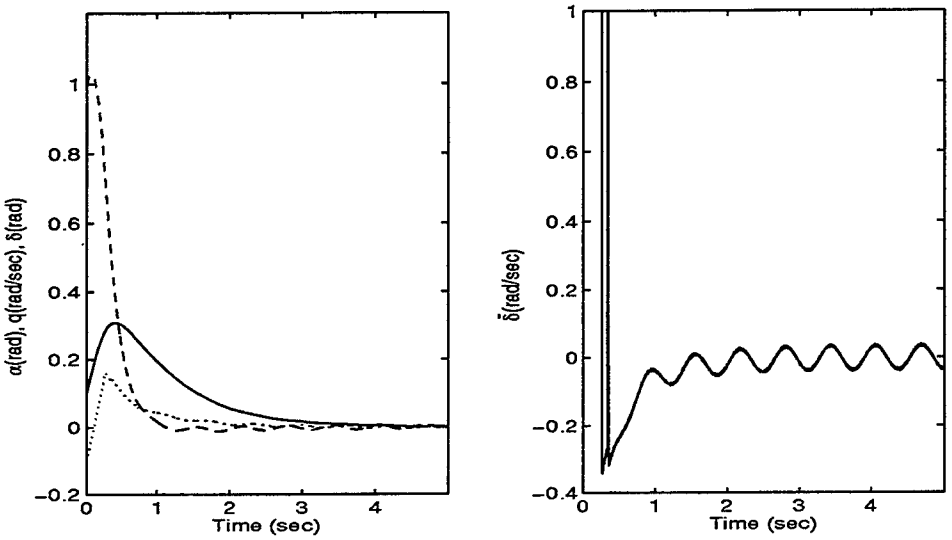


Figure 2.2: Combined PLC/LHG design ($N = 5$, $k = 200$). $T_1 = 21$, $T_2 = 19$, $d = 5\sin(10t + 3)$. The left plot: states of the plant and the actuator, the solid line is the angle of attack α, the dashed line the pitch rate q, and the dotted line the deflection δ; the right plot: the actuator rate $\dot{\delta}$.

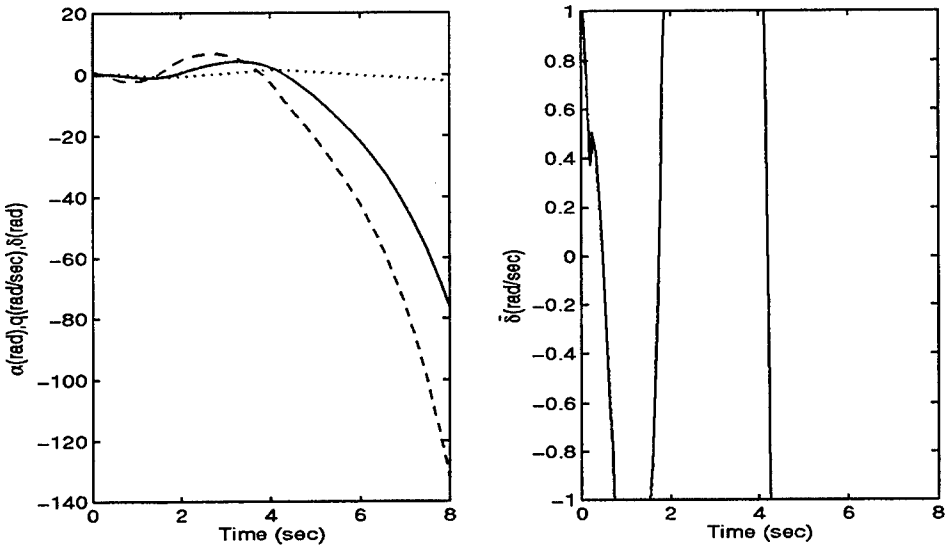


Figure 2.3: PLC Design ($k = 0$). $T_1 = 17$, $T_2 = 23$, $d = 0$. The left plot: states of the plant and the actuator, the solid line is the angle of attack α, the dashed line the pitch rate q, and the dotted line the deflection δ; the right plot: the actuator rate $\dot{\delta}$.

of the 15% inaccuracy in the actuator parameters T_1 and T_2. To compare with the LHG feedback law, we set $N = 0$ in (2.11), resulting a feedback law with constant gain $-(\mu_2 T^*)^{-1}\tilde{R}^{-1}\tilde{B}'\tilde{P}_0$. Simulations (Fig. 2.4) shows that the combined PLC/LHG feedback law results in much better transient performance. By increasing the value of N, the transience performance can be further improved.

2.5. Conclusions

A stabilizing feedback control law design method for linear systems with rate-limited actuators is proposed. This design method is based on two design techniques recently developed for linear systems with position limited actuators and takes advantages of the both design techniques, while avoiding their disadvantages. Application of the proposed design to an F-16 fighter aircraft demonstrates the applicability, robustness and effectiveness of the novel control law.

Acknowledgement

I would like to thank my focus point Dr Siva Banda for his valuable guidance throughout this summer research program, and Dr Meir Pachter for many technical discussions. I would also like to thank members of the Control Analysis Section at Wright Laboratory for providing me with various conveniences.

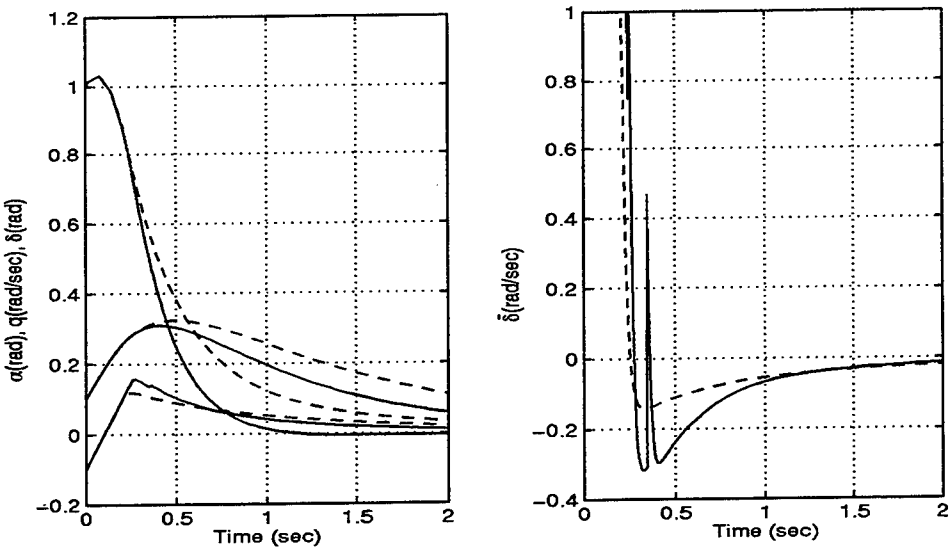


Figure 2.4: Transience performance: PLC/LHG ($N = 5, k = 10$, solid lines) vs. LHG ($N = 0, k = 10$, dashed lines). $T_1 = 21$, $T_2 = 19$, $d = 0$.

References

[1] J. H. Blakelock, *Automatic Control of Aircraft and Missiles*, John Wiley & Sons, Inc., New York, 1990.

[2] J.M. Berg, K.D. Hammett, C.A. Schwartz, and S.S. Banda, "An analysis of the destabilizing effect of daisy chained rate-limited actuators," *IEEE Trans. Contr. Sys. Tech.*, Vol. 4, 1996.

[3] M.A. Dornheim, "Report pinpoints factors leading to YF-22 crash," *Aviation Week Space Technol.*, pp. 53-54, Nov. 9, 1992.

[4] G.F. Franklin, J.D. Powell and A. Emami-Naeini, *Feedback Control of Dynamic Systems*, Addison Wesley, New York, 1991.

[5] J.M. Lenorovitz, "Gripen control problems resolved through in-flight, ground simulations," *Aviation Week Space Technol.*, pp. 74-75, June 18, 1990.

[6] M. Pachter, P.R. Chandler and M. Mears, "Reconfigurable tracking control with saturation," submitted for publication.

[7] C.A. Shifrin, "Gripen likely to fly again soon," *Aviation Week Space Technol.*, pp. 72, Aug. 23, 1993.

[8] A. Saberi, Z. Lin and A.R. Teel, "Control of linear systems subject to input saturation," *IEEE Trans. AC*, Vol. 41, No.3, pp. 368-378, 1996.

[9] H.S. Tsien, *Engineering Cybernetics*, McGraw-Hill Book Company, Inc., New York, 1954.

[10] G.F. Wredenhagen and P.R. Belanger, "Piecewise-linear LQ control for systems with input constraints," *Automatica*, vol. 30, pp. 403-416, 1994.

STUDY ON DEAD RECKONING TRANSLATION IN HIGH LEVEL ARCHITECHETURE

Kuo-Chi Lin
Associate Professor
Institute for Simulation and Training
University of Central Florida
3280 Progress Drive
Orlando, FL 32826-0544

Final Report for:
Summer Faculty Research Program
Wright Laboratory

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

and

Wright Laboratory

August 1996

STUDY ON DEAD RECKONING TRANSLATION IN HIGH LEVEL ARCHITECHETURE

Kuo-Chi Lin
Associate Professor
Institute for Simulation and Training
University of Central Florida

Abstract

In HLA, the concept of dead reckoning (DR) is extended to attribute extrapolation. The federation can use any formula that is agreed upon by the participating federates. If a non-DIS-standard DR algorithm is chosen, a DIS-compatible simulator has to modify its software to join the federation. The cost can be significant for this kind of modification. This paper suggest the use of dead-reckoning translators to solve this problem. To assist the study, a software tool Dead Reckoning Translator Simulation Program (DRTSP) was designed and developed. It is a MATLAB based program using the Graphic User Interface (GUI) to provide a user friendly environment. A sample flight trajectory recorded from a flight simulator is used to test the algorithms. The output of the program are the number of PDUs (updates) and the errors of each dead reckoning algorithms. Conclusions are drawn from the results of the numerical experiments.

STUDY ON DEAD RECKONING TRANSLATION IN HIGH LEVEL ARCHITECHETURE

Kuo-Chi Lin

Introduction

In response to the DoD Modeling and Simulation Master Plan, the Defense Modeling and Simulation Office (DMSO) is establishing a common technical framework to facilitate the interoperability of all types of models and simulations among themselves and with C4I systems, as well as to facilitate the reuse of modeling and simulation components. This Common Technical Framework, and specifically the High Level Architecture (HLA), represents the highest priority effort within the DoD modeling and simulation community [1]. HLA uses a different architecture from the Distributed Interactive Simulation (DIS). In the DIS architecture, a simulation joins a distributed exercise by connecting to a communications network where the other exercise members reside. All computing in DIS exercises is done in the member simulations, and members interoperate by sending and receiving data in the form of Protocol Data Units (PDUs). In contrast, in the HLA, each simulation operates within the virtual world via a common software module known as the HLA Runtime Infrastructure (RTI). While each "federate" (exercise member) has only one connection to the RTI, the RTI can be distributed throughout the "federation" (exercise), or located centrally. To let the legacy simulators participate in the HLA exercise, various devices are developed to serve as the interface between the simulation and RTI. These devices may have one or more functions, and bear different names such as "middleware", "gateway", etc.

Dead Reckoning (DR) is an extrapolation methodology used in Simulator Networking (SIMNET) and DIS to reduce the network traffic [2][3]. SIMNET, which is the predecessor of DIS, only uses one DR algorithm. In DIS, entities can use various DR algorithms. There is a list of formulas in Annex B of the DIS standard (see Appendix), which includes the SIMNET DR formula as formula #2. When SIMNET simulator participate in a DIS exercise, a "translator" is used to convert the PDUs to DIS compatible format. However, it does not convert the dead reckoning algorithms. All other participants can only use the SIMNET DR formula, or accept the high error cause by using a different DR formula.

In HLA, the concept of dead reckoning is extended to attribute extrapolation. The federation can use any formula that is agreed upon by the participating federates. HLA no longer dictates the set of formula that can be used in the exercise. Therefore, if a non-DIS-standard DR algorithm is chosen, a DIS-compatible simulator has to modify its software to join the federation. The cost can be significant for this kind of modification.

This report suggest the use of dead-reckoning translators to solve this problem. The variables sent or received through RTI are based on the DR algorithm designated by the federation. The DR translator handles the local PDUs to and from the DIS-compatible simulators.

Dead Reckoning Translation

Figure 1 shows the block diagram of the dead reckoning translation. The DIS-compatible simulator sends or receives DIS PDUs which contain the standard DIS dead reckoning information. The translator converts the PDUs from the simulator into the dead-reckoning variables subscribed to by other federates and send them to RTI. The translator also converts the dead-reckoning variables received from RTI to the DIS PDU for the simulator. RTI publishes and subscribes the read reckoning variables under the HLA rules.

In this setup, if the DR algorithms used in the federation are all DIS standard DR algorithms, the translation become a rather trivial format conversion. This paper, however, studies the situations when there are non-DIS-standard DR algorithms are used, or if other federates do not support all DIS standard DR algorithms.

The following principles are used for dead reckoning translation:

Case 1. The algorithm of incoming DR information is of higher order than the algorithm to be converted to. The translator takes the incoming DR information and generates a DR trajectory. Based on this trajectory, the DR variables of the lower algorithm are generated as the output. An example converting DIS second-order DR PDU into SIMNET DR variables will be shown in the later section.

Case 2. The algorithm of incoming DR information is of lower (or equivalent) order than the algorithm to be converted into. The translator generates the output DR variables from the incoming information and puts in zeros if necessary. For example, to convert the first-order DR variables to the second-order DR variables, the translator takes the incoming velocity information and sets the acceleration to zero.

The non-DIS-standard DR variables may have different forms. Some computation may be necessary before apply the above principles. There is an example in the later section.

Dead Reckoning Translator Simulation Program

To assist the research on the dead reckoning translation, a MATLAB based software tool, Dead Reckoning Translator Simulation Program (DRTSP), was developed. The Graphic User Interface (GUI) program window is shown in Figure 2. The program reads in a sample trajectory file and simulates the dead reckoning of the federate and the dead reckoning translation. Though the GUI program window, users can select the DR algorithms on both side of the translator. The algorithms include all DIS standard methods, SIMNET, and two non-DIS-standard formulas which will be discussed in the later section. Other parameters such as the thresholds can also be changed by the user. The simulation results are shown in the Output Window of the GUI program window. The accuracy is represented by the root-mean-square (RMS) error and the maximum (MAX) error of the dead reckoning compared with the true trajectory. The number of PDUs shows the possible increase in the number of data packages.

In the following sections, two examples are used to demonstrate the concept of dead reckoning translation and the results obtained using DRTSP.

Translation between SIMNET and DIS

For the exercise involving both SIMNET and DIS-compatible simulators, as mentioned in the early section, the selection of dead reckoning algorithms may cause problem. SIMNET only supports one DR formula given by:

$$P = P_0 + V_0 \Delta t \qquad (1)$$

where P is the position at time Δt after the update time, P_0 is the position at update time, and V_0 is the velocity at update time. The variables (P_0, V_0) are the DR variables passed through RTI. This formula is compatible with DIS (the #2 formula in the DIS standard). Therefore, all DIS-compatible simulators have no problem receiving the SIMNET DR information. However, unless the DR algorithm used by the federation is limited to Equation (1), SIMNET simulator will have problem handling DR information from others.

In this example, let us assume that all the DIS-compatible simulators use the DIS #5 formula in the standard, given by:

$$P = P_0 + V_0 \Delta t + A_0 \Delta t^2 / 2 \qquad (2)$$

where and A_0 is the acceleration at update time. The variables (P_0, V_0, A_0) are the DR variables passed through RTI. Since SIMNET simulator does not support Equation (2), it will use only part of the information, namely (P_0, V_0), and Equation (1) to do dead reckoning. Figure 3 shows the dead reckoning of the sample trajectory. The errors are in general greater than the original one which uses (P_0, V_0, A_0) and Equation (5).

Now put the translator discussed in the previous section is put between RTI and the SIMNET simulator. The translator will send a new set of (P_0, V_0) based on "the dead reckoning of the dead reckoning trajectory" to the SIMNET simulator. The simulator then uses Equation (1) to do dead reckoning. The results are shown in Figure 4. The comparison of the data in Table 1 and Table 2 shows that the error is much smaller when the translator is used. The increases in PDUs only have effects on the local network between the translator and SIMNET simulator. The global network traffic will be reduced because of the use of a high order formula represented by Equation (2).

Non-DIS-standard Dead Reckoning Formulas

In DIS standard, the Entity State PDU contains the position, velocity, and acceleration of the vehicle no matter what DR algorithm is used. Therefore, it is logical to take the advantage of all the available information and the second-order formulas such as Equation (2). However, in HLA, the federations are allowed to use non-DIS-standard algorithms. This paper uses the following two formulas as examples:

$$P = \left(1 + \frac{\Delta t}{\Delta t_{-1}}\right) P_0 - \frac{\Delta t}{\Delta t_{-1}} P_{-1} \tag{3}$$

$$\begin{aligned} P = & \left(1 + \frac{(2\Delta t_{-1} + \Delta t_{-2})\Delta t}{\Delta t_{-1}(\Delta t_{-1} + \Delta t_{-2})} + \frac{\Delta t^2}{\Delta t_{-1}(\Delta t_{-1} + \Delta t_{-2})}\right) P_0 \\ & + \left(-\frac{(\Delta t_{-1} + \Delta t_{-2})\Delta t}{\Delta t_{-1}\Delta t_{-2}} - \frac{\Delta t^2}{\Delta t_{-1}\Delta t_{-2}}\right) P_{-1} \\ & + \left(\frac{\Delta t_{-1}\Delta t}{\Delta t_{-2}(\Delta t_{-1} + \Delta t_{-2})} + \frac{\Delta t^2}{\Delta t_{-2}(\Delta t_{-1} + \Delta t_{-2})}\right) P_{-2} \end{aligned} \tag{4}$$

The definition of all variables are illustrated in Figure 5. In these two algorithms, only the position, represented by P_0, is sent at the update time t_0. The previous updates P_{-1} and P_{-2} sent at times t_{-1} and t_{-2}, respectively, have been stored for Equations (3) and (4).

There are two advantages of using these algorithms. First, the update package is smaller--only the position is sent, but the order of accuracy is higher. Second, for some attributes such as temperature, the rate of change of the attribute may not be available. In this situation, simply send the current value and let other federates to extrapolate the data.

To convert the above non-DIS-standard update information into DIS-standard PDU information, the translator needs to evaluate the velocity and acceleration at the update time. For Equation (3), the corresponding velocity is given by

$$V_0 = \frac{P_0 - P_{-1}}{\Delta t_{-1}} \tag{5}$$

and acceleration $A_0 = 0$. For Equation (4), the corresponding velocity and acceleration are given by

$$V_0 = \frac{(2\Delta t_{-1} + \Delta t_{-2})P_0}{\Delta t_{-1}(\Delta t_{-1} + \Delta t_{-2})} - \frac{(\Delta t_{-1} + \Delta t_{-2})P_{-1}}{\Delta t_{-1}\Delta t_{-2}} + \frac{\Delta t_{-1}P_{-2}}{\Delta t_{-2}(\Delta t_{-1} + \Delta t_{-2})} \tag{6}$$

$$A_0 = \frac{2P_0}{\Delta t_{-1}(\Delta t_{-1} + \Delta t_{-2})} - \frac{2P_{-1}}{\Delta t_{-1}\Delta t_{-2}} + \frac{2P_{-2}}{\Delta t_{-2}(\Delta t_{-1} + \Delta t_{-2})} \tag{7}$$

respectively. If the conversion is between two algorithms of different orders of accuracy, the principles discussed in the previous sections can apply.

Conclusion

From the above discussions and examples, the idea of dead reckoning translator does lead to better results. In Table 1, with no translator, the maximum errors can exceed the threshold from 40% up to 600%. Using the translator, the maximum errors only exceed the threshold modestly, between 4% and 84%, as shown in Table 2. In this study, the threshold used in the translator is the same as the threshold used in the original dead reckoning. If the maximum error after the translator is still not acceptable, it can be reduced by lower the threshold used in the translator.

Before all new systems can be made HLA compliant, using the translator on the legacy systems is a very attractive idea.

Acknowledgment

This research is sponsored by Summer Faculty Research program, Air Force Scientific Research Office and Wright Laboratory.

References

[1] Defense Modeling & Simulation Office, "Interim Definition of the DoD High Level Architecture for Modeling and Simulation", HLA Briefing, April 1, 1996.

[2] Pope, A. R., "The SIMNET Network and Protocols", Report No. 7627, BBN Systems and Technologies, Cambridge, MA, June, 1991.

[3] IST, "Standard for Distributed Interactive Simulation--Application Protocols, Version 2.0," IST Report No IST-CR-94-50, March, 1994.

Table 1. Without translation, A: DIS simulator, B: SIMNET

threshold	# PDU A	# PDU B	RMS error (ft)		MAX error (ft)	
(ft)	(901 data points)		A	B	A	B
1	34	34	0.34	0.87	1.00	7.04
2	25	25	0.67	1.70	1.99	12.6
3	21	21	0.91	2.08	2.99	13.7
4	18	18	1.29	2.73	3.98	12.8
10	13	13	2.90	5.14	9.99	24.9
20	10	10	5.42	7.72	20.0	43.5
30	8	8	7.59	10.9	29.8	59.4
40	6	6	11.1	12.7	39.8	56.6

Table 2. With translation, A: DIS simulator, B: SIMNET

threshold	# PDU A	# PDU B	RMS error (ft)		MAX error (ft)	
(ft)	(901 data points)		A	B	A	B
1	34	83	0.34	0.38	1.00	1.84
2	25	63	0.67	0.77	1.99	3.61
3	21	48	0.91	1.12	2.99	5.39
4	18	42	1.29	1.55	3.98	7.04
10	13	26	2.90	3.19	9.99	11.7
20	10	18	5.52	5.34	20.0	20.6
30	8	14	7.59	7.72	29.8	33.7
40	6	11	11.1	11.0	39.8	41.6

Figure 1. Block diagram of dead reckoning translation.

Figure 2. Program window of DRTSP.

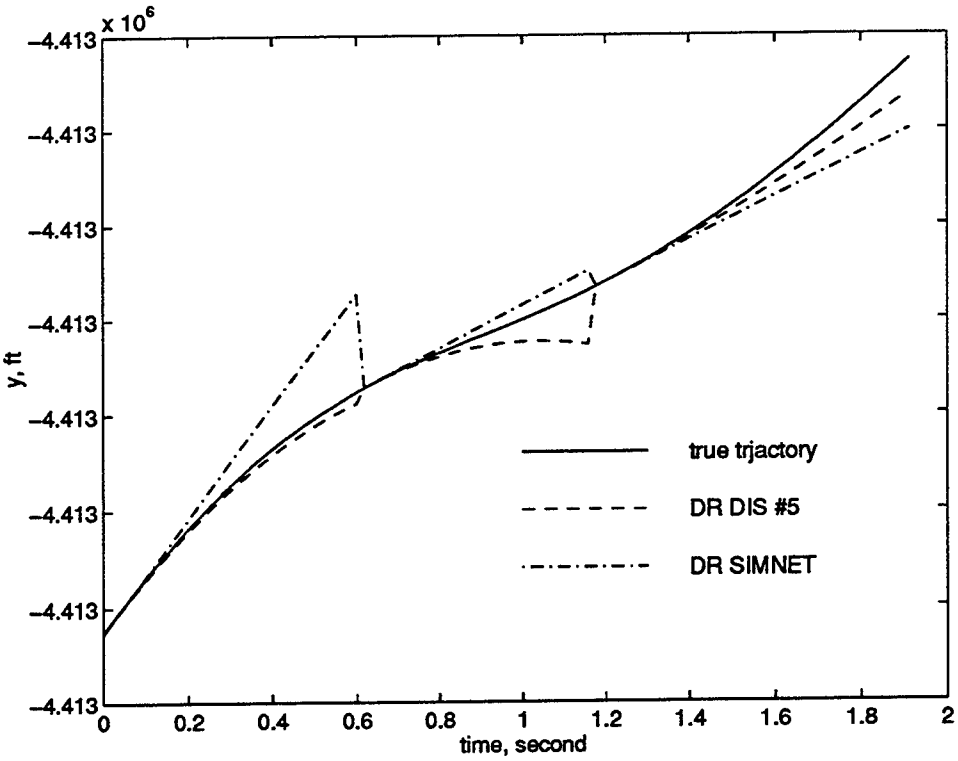


Figure 3. SIMNET uses DIS #5 DR variables without translation.

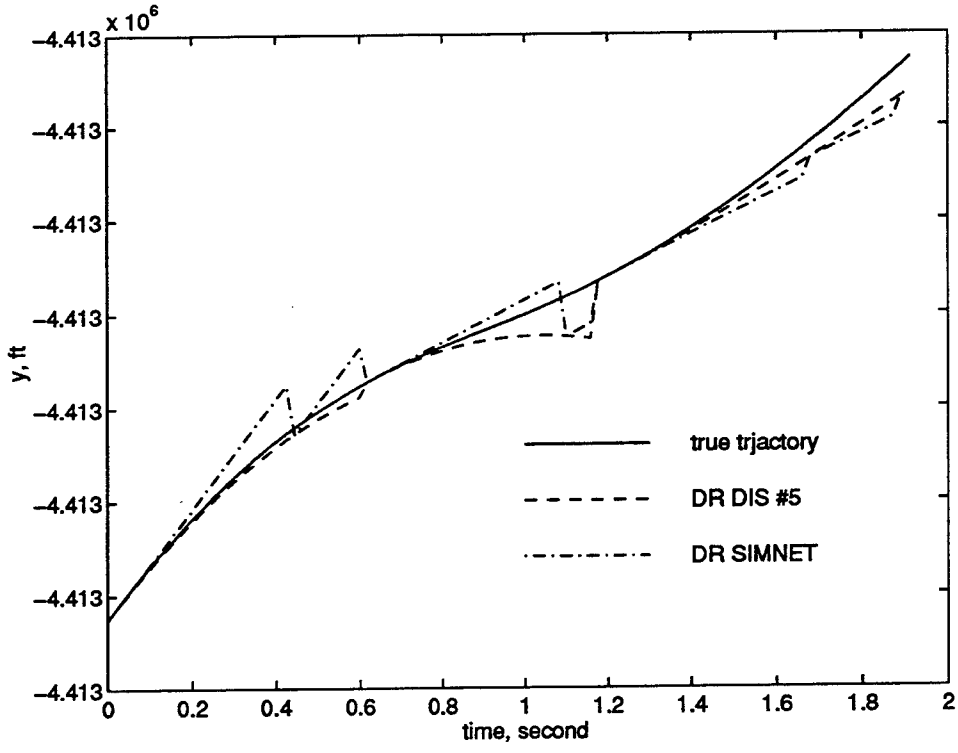


Figure 4. SIMNET uses DIS #5 DR variables with translation

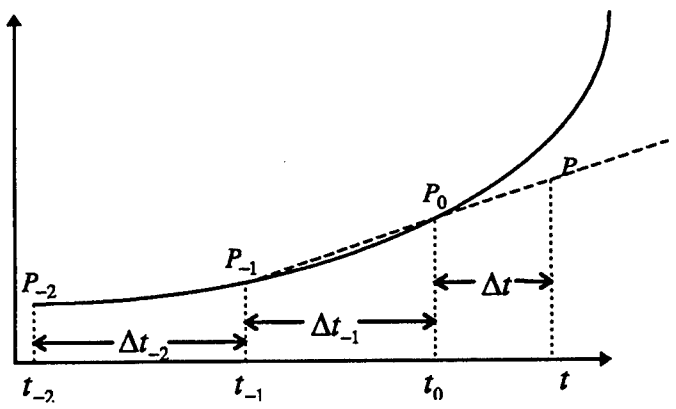


Figure 5. Variables used in the non-DIS-standard algorithms.

Appendix

Table A1 is Table 1 in Annex B of the DIS standard [3].

Table A1. DEAD RECKONING FORMULAS

FIELD	MODEL		FORMULA
1	STATIC		N/A
2	DRM (FPW)	(1)	$P = P_0 + V_0 \Delta t$
3	DRM (RPW)	(1)	$P = P_0 + V_0 \Delta t$
		(2)	$[R]_{w \to b} = [DR][R_0]_{w \to b}$
4	DRM (RVW)	(1)	$P = P_0 + V_0 \Delta t + A_0 \Delta t^2 / 2$
		(2)	$[R]_{w \to b} = [DR][R_0]_{w \to b}$
5	DRM (FVW)	(1)	$P = P_0 + V_0 \Delta t + A_0 \Delta t^2 / 2$
6	DRM (FPW)	(1)	$P = P_0 + [R_0]^{-1}_{w \to b} (R_1 V_b)$
7	DRM (FPB)	(1)	$P = P_0 + [R_0]^{-1}_{w \to b} (R_1 V_b)$
		(2)	$[R]_{w \to b} = [DR][R_0]_{w \to b}$
8	DRM (RVB)	(1)	$P = P_0 + [R_0]^{-1}_{w \to b} (R_1 V_b + R_2 V)$
		(2)	$[R]_{w \to b} = [DR][R_0]_{w \to b}$
9	DRM (FVB)	(1)	$P = P_0 + [R_0]^{-1}_{w \to b} (R_1 V_b + R_2 V)$

A CONCEPTUAL MODEL FOR HOLOGRAPHIC RECONSTRUCTION AND MINIMIZING ABERRATIONS DURING RECONSTRUCTION OF CYLINDRICAL HOLOGRAMS.

James S. Marsh
Professor
Department of Physics

The University of West Florida
Pensacola, FL 32514

Final Report for:
Summer Faculty Research Program
Wright Laboratory
Eglin AFB

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

and

Wright Laboratory
Eglin AFB

August, 1996

A conceptual model for holographic reconstruction and minimizing aberrations during reconstruction of cylindrical holograms.

James S. Marsh
Professor
Department of Physics
The University of West Florida

Abstract

A simple model reproduces most, though not all, the features of holographic image formation. In looking at a holographic image, it is as though you are looking at the original object through the hologram surface, but the holographic volume is filled with a medium with index of refraction $\mu = \lambda/\lambda_0$, where λ is the reconstruction and λ_0 the reference wavelength.

I calculate the wave aberration for reconstruction of a point on the axis of an 18 in. diameter cylindrical hologram, with the reference beam (λ_0 = 694 nm) located 62 in. along the axis from the point, for different locations of the reconstruction beam at wavelengths 688 and 632 nm. For the two reconstruction wavelengths, the region in which to place the reconstruction beam in order to minimize the aberrations is found. These regions are consistent with the rule to place the reconstruction beam where it will satisfy the Bragg condition at the pupil.

A conceptual model for holographic reconstruction and minimizing aberrations during reconstruction of cylindrical holograms.

James S. Marsh

0. Introduction.

This report contains two distinct sections which, however, are related in that the first section motivates the second. In my last two reports I explored the equations for image formation in holographic reconstructions. The goal was to eliminate inherent distortion in the reconstructed image field compared to the original object field. In addition, I attempted to find a way around the inconveniently strong dependence of the image field distortion on the holographic pupils through which the image field was viewed.

Based on the work of the earlier reports, I was able to develop a simple model for the mechanism of holographic reconstruction. This model is the subject of the first section of this report. This allowed me to grasp in a conceptual way the difficulties in the way of achieving the goals of zero distortion and pupil indepencence of the image field. On this basis I concluded that those goals were not to be achieved in any practical way.

Because of this I switched the emphasis to reconstruction with maximum image quality. It had already become clear with experience that greater weight should be given to this should it conflict with minimum distortion and pupil independence. In the second section I explore the wave aberrations for the reconstruction of a point object at the center of an 18 inch diameter cylindrical hologram as a function of placement of the reconstruction beam. Based on this I find the region in which the reconstruction beam can be place in order to minimize aberrations and thus maximize the theoretical resolution of the reconstruction. The results of this investigation are at the end of this report in the section entitled "Summary".

1. A conceptual model for holographic reconstruction.

1.1 Introduction.

Because we must reconstruct holograms under non-ideal conditions, there are inevitably distortions introduced into the image field. That is to say, the measured position of a holographic image does not coincide, in general, with the position of the corresponding object within the holographic volume. This is because we are making the hologram in ruby laser light, 694 nm, and reconstructing in laser diode, 688nm, or He-Ne laser light, 632 nm.

Because of these distortions, questions arise as to what measurements of holographic image positions actually mean in terms of the actual object positions. Suppose sets of stereo pairs of the holographic images are taken and used to "locate" objects in 3 dimensions. Or a set of holographic images is used in the catscan algorithm to "locate" objects. What is the relationship of these locations to actual object locations? Questions such as this become even more acute if reconstructions are done in red (632 nm) or green (532 nm) light instead of laser diode (688 nm) light.

A complete theory of these distortions exists in the equations of holographic image formation. But these equations are not very helpful in giving us a *conceptual* feeling for the nature of the distortion. What we need here is a simple conceptual model that can be conveyed to those concerned with image processing so that they will be aware of the nature of the distortion problem, or at least the main features of the problem.

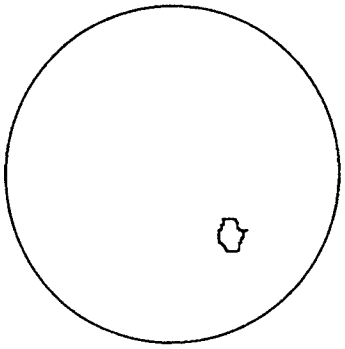
This section presents such a model. This model is based on the geometric interpretation of the holographic image forming equations I developed in my last report[1]. This model does not do *everything* that the image forming equations say happens, but it does a major portion of what happens and what it does includes the major difficulty in relating image position to object position within the hologram. The purpose of this model is to help *visualise* the distortion problem and thus, one hopes, suggest solutions.

---

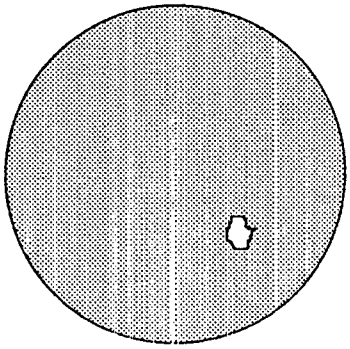
[1] J. S. Marsh, "Geometric interpretation of holographic image equations", Final report for the Summer Research Extension Program, Eglin AFB, (May 1966)

1.2 The fish bowl model

We imagine making a cylindrical hologram with wavelength λ_o = 694 nm and reconstructing with wavelength λ.



Hologram is made



Hologram is reconstructed.

In the figure, the left side shows a cylindrical hologram, as seen from the top, with an object located at its actual position in the holographic volume. On the right hand side the situation at reconstruction is shown. During reconstruction with wavelength λ it is as though we are looking at the hologram with the object in its actual position within the holographic volume, but the holographic volume itself is filled with a medium whose index of refraction is $\mu=\lambda/\lambda_0$.

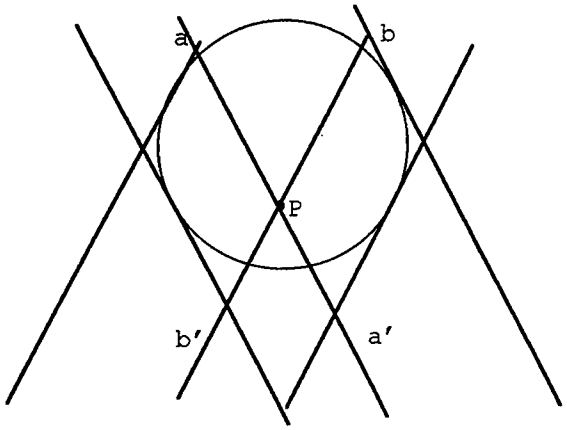
This model accurately accounts for the apparent refraction of the ray from the object at the surface of the hologram, and reasonably well for the shift of image distance from the pupil relative to object distance. It completely ignores the shearing effect. All these effects were discussed in the memo "Geometric interpretation of holographic image equations."

So. Looking at the holographic image is like looking at the object imbedded in a medium with index of refraction $\mu=\lambda/\lambda_0$ which fills the holographic volume. The problem is to find the actual position of the object, given that we are only observing it embedded in the medium.

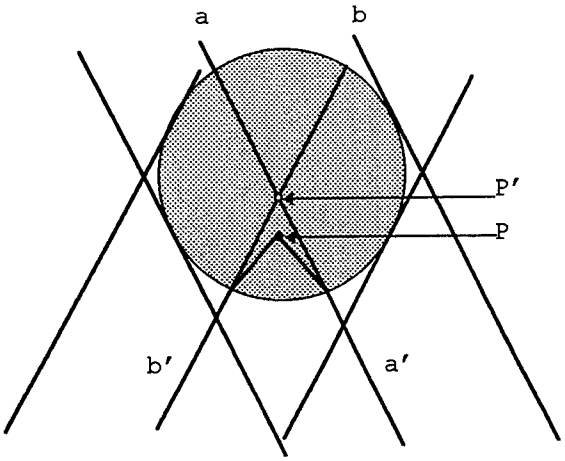
In terms of DeVoe's rods, it would like be mounting the rods and plate in a cylinder, then filling the cylinder with water or some other medium, and then locating the actual positions of the rods from the stereographic pictures of the rods imbedded in the water or what ever medium is used to fill the cylinder. It is like locating the actual position of a fish in a fish bowl by looking at the fish through the sides of the bowl.

1.3 Stereo object location.

This conceptual model allow us, for example, to get a picture of what is happening when stereo pairs of images are used to locate the image.



This figure depicts a stereo pair of images of an object at its actual position inside the holographic volume. One stereoscopic view is along the line parallel to a′-a and the other is along the line parallel to b′-b. Because the object lies on the line a′-a in one view and on the line b′-b in the other, we know that the object is at the point P.



This figure depicts a stereo pair of images on the holographic reconstruction. The object is again at its actual point in the holographic volume, but is is imbedded in a medium of index of refraction μ where $\mu<1$ in our case. Then rays from the object will refract as shown at the surface of the hologram, and the object will *appear* to be along the lines a'-a and b'-b in the two stereo views. The position of the object will be judged to be at the intersection of a'-a and b'-b, point P', which is different from the actual location of the object, point P.

Having located point P', we must now deduce the location of P. This illustrates the crux of the problem of mapping the objected field from the distorted image field. Although this conceptual model does not include all the features of the actual distortion in the holographic image field, if we can find an efficient way to correct for the distortion in this conceptual model, I am confident that we can find a way to extend this to the actual model.

2. Minimizing aberrations during reconstruction of cylindrical holograms.

2.1 Introduction.

Dr. Roland Anderson[2] has shown us that for holographic reconstructions with minimum aberrations in the image, placement of the reconstruction beam can be quite critical. He was working with plane holograms, but the same conclusion is true for cylindrical holograms. It is desirable to minimize aberrations in the reconstruction in order to improve the resolution and to decrease the scale of the speckel. Speckle is a very annoying noise source.

In this exercise I will explore the aberration field for several cases of interest of holographic reconstruction for cylindrical holograms. The main conclusion is that positioning of the reconstruction beam transverse to the axis of the hologram is fairly critical while positioning the beam along the axis is much less critical. The detailed conclusions are in the Summary section at the end of this report. They tell where to place the reconstruction beam for reconstructions with minimum aberration for a 9 in. radius hologram with reconstruction at 688 nm and 632 nm.

2.2 Wave aberration.

The basic equations for aberrations come from Champagne[3] and were further refined by Rebordao[4] and Peng and Frankena[5]. In the pupil of the hologram we have a wave of the form

$$\exp(i\Delta\phi) = \exp(ik'(-r_i+r_c+\mu(r_o-r_r))), \qquad (1)$$

where r_i, r_c, r_o, and r_r are distances from the image, reconstruction, object, and reference point respectively. $\mu = \lambda/\lambda_0$ where λ is the reconstruction wavelength and λ_0 is the reference beam wavelength.

---

[2] Private Communication.

[3] Edwin B. Champagne, J. Opt. Soc. Am., 57, 51-55 (1967)

[4] J. M. Rebordao, J. Opt. Soc. Am. A, 1, 788-790 (1984)

[5] Ke-Ou Peng and Hans J. Frankena, Applied Optics, 25, 1319-1326 (1986)

The wave aberration is calculated by calculating $PHASE=-r_i+r_c+\mu(r_o-r_r)$ over a square pupil in the hologram centered by the pupil point. The radius at which this quantity exceeds $\lambda/2$ is the radius of the pupil aperture which, if used, will give the best resolution at the image point in question. This is the Rayleigh criterion for diffraction limited operation of an optical instrument. The resolution is computed by Abbe's criterion for a microscope[6]. The results are generated by a routine written for Mathcad 6.0.

I have assumed an 18 in. diameter hologram, made with the ruby laser reference beam at 694 nm and reconstructed with laser diode light at 688 nm or HeNe laser light at 632 nm. I have assumed that we are reconstructing an object at the center of the hologram on the axis. With the coordinate system origin at the pupil[7], as shown in the figure, this puts the object at $y_o=-9$ in. and $x_o=z_o=0$.

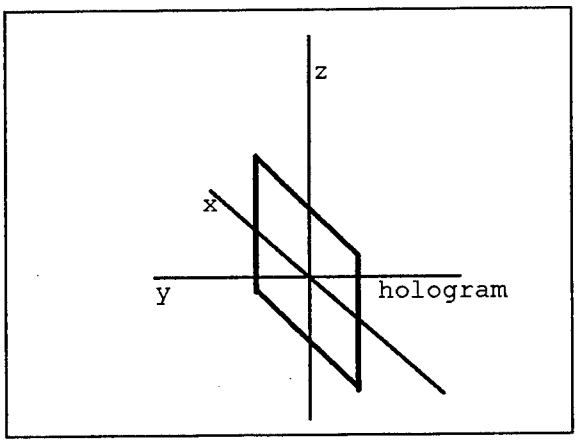


Figure 1.

I take the reference beam to be on axis, 57 in. from the end of the hologram or 62 in. from the center of the hologram. This makes the coordinates of the reference beam be $x_r=0$, $y_r=-9$ in., $z_r=62$ in.

Though the hologram is shown as flat, in general it will be curved. The axis of the hologram will be parallel to the z axis at y=-9 in.

2.3 Flat hologram.

Though the hologram is a cylinder, it is nevertheless interesting to start with the hologram flat, in order to make contact with Dr.

---

[6] This is described in my first report, J. S. Marsh, "Themeasurementof image positions in cylindrical holograms", Final report for the Summer Faculty Research Program, Eglin AFB (August 1995)

[7] This is different from previous reports, where I put the origin of the coordinate system at the center of the hologram on the axis.

Anderson's work and to see, in the next section, what specific effect the curvature of the hologram has on the aberrations.

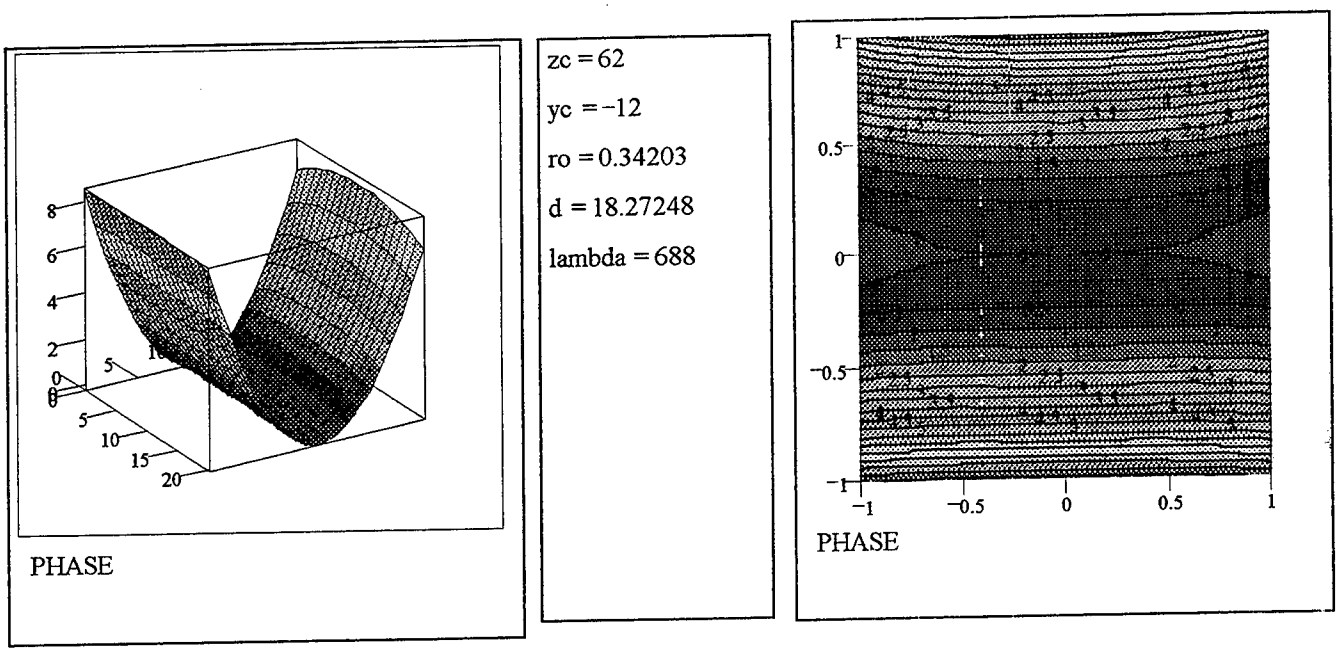


Figure 2

On the left we see a 3-D depiction of the wave aberration over a pupil square 2 in. on a side. The figure on the right is a contour map of the same thing. This kind of wave aberration is mostly symptomatic of astigmatism, which comes about because of the large angles made by the reference and reconstruction beams relative to the pupil normal. This is the sort of thing seen by Anderson.

2.4 Curved hologram.

In the next figure we make the hologram into a cylindrical hologram with a 9 in. radius, leaving everything else the same as in the previous section.

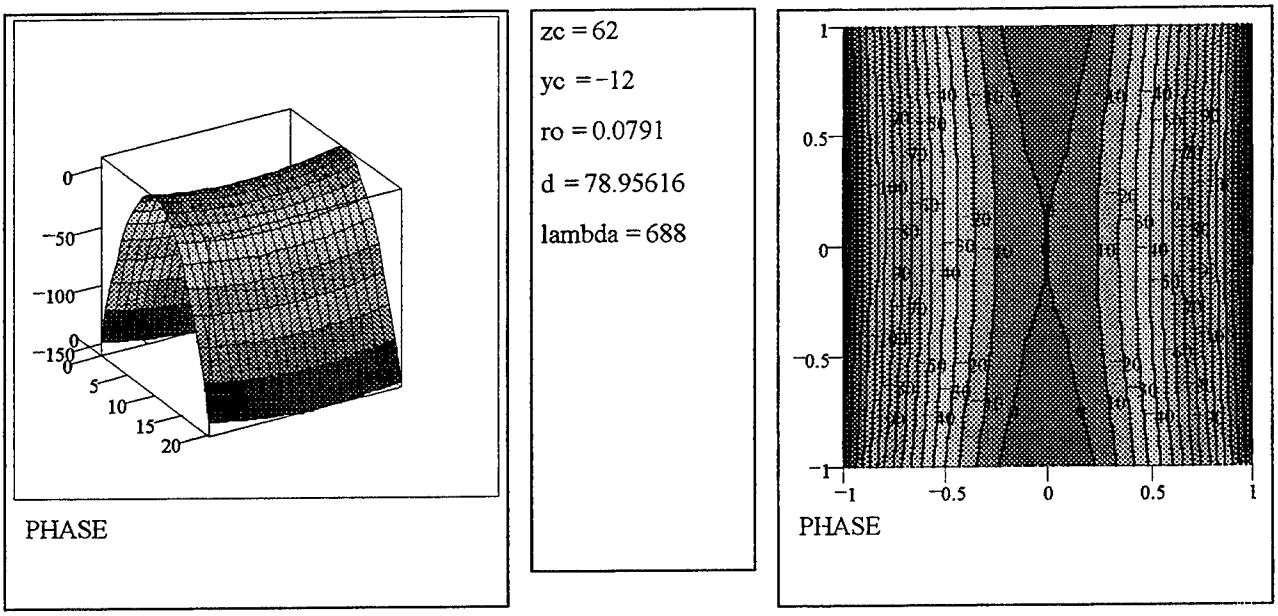


Figure 3.

Here we see that the curvature of the hologram introduces its own source of astigmatism. In this case is an order of magnitude larger and in the opposite sense compared that that of the flat hologram.

This observation makes it plausible that the fish bowl model for reconstructions in a cylindrical hologram should indeed have the astigmatism so strongly implied by that model.

2.5 Region of minimal aberration, 688 nm

Now I take this case and explore the effect that different placements of the reconstruction beam have on the aberrations. We take λ=688 nm, x_r=0, y_r=-9, z_r=62 in. The object point is taken at the center of the hologram, x_o=z_o=0, y_o=-9 in. The image point is calulated from the image forming equations. The reconstruction point varies over the locations x_c=0, y_c=-9,(.1),-10 in., z_c=60.5,(.5),63.5 in.

For each point of x_c, y_c, z_c, the resolution is calculated in microns. A small number for the resolution indicates a fine resolution resulting from small aberrations. The results are shown in the next figure.

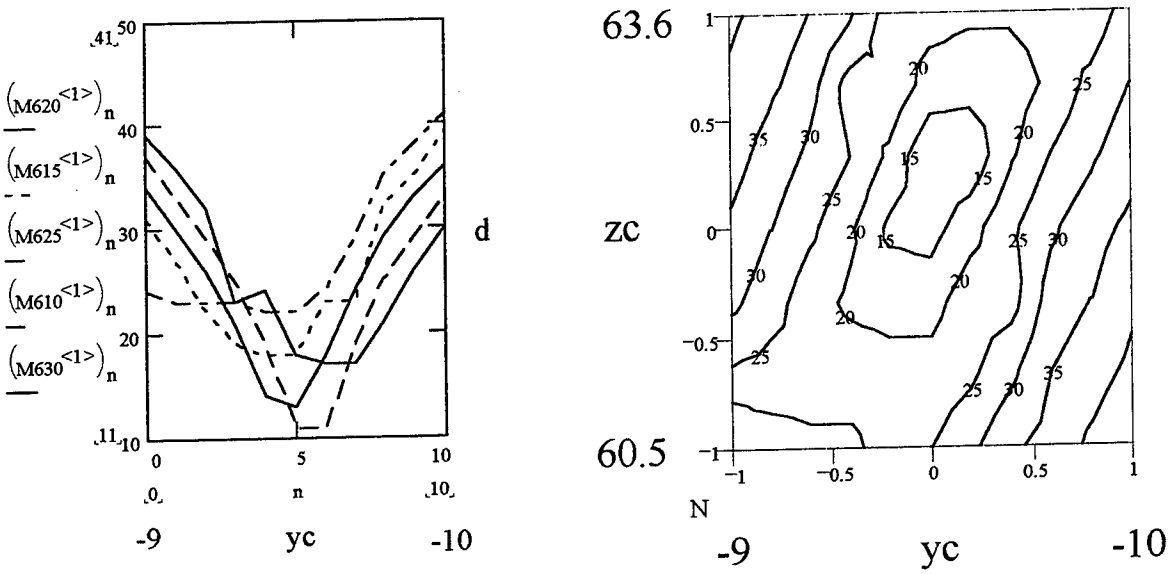


Figure 4.

On the right is a contour map of d = resolution in microns. On the left are some sections throught the contour map. M615, for instance, is the section for z_c=61.5 in. On the right we see the region of minimal aberration enclosed within the contour d = 15. If the reconstruction beam is placed within this region the aberrations will be such that resolution down to 15 microns or better can be theoretically achieved. The best resolution attainable within this region can be seen from the figure on the left, where we see that resolution down to about 11 microns can be achieved.

The region of minimal aberration is relatively small, roughly between $-9.4<y_c<-9.6$ and $62.0<z_c<62.5$ in. The contours show us how much we lose by straying outside this region.

However, this situation can be improved by taking field curvature into account. This has so far been ignored.

2.6 Curvature of field.

The next figure is generated with paramater values chosen to

emphasize the effect. On the left is a wavefront aberration with contours on the right.

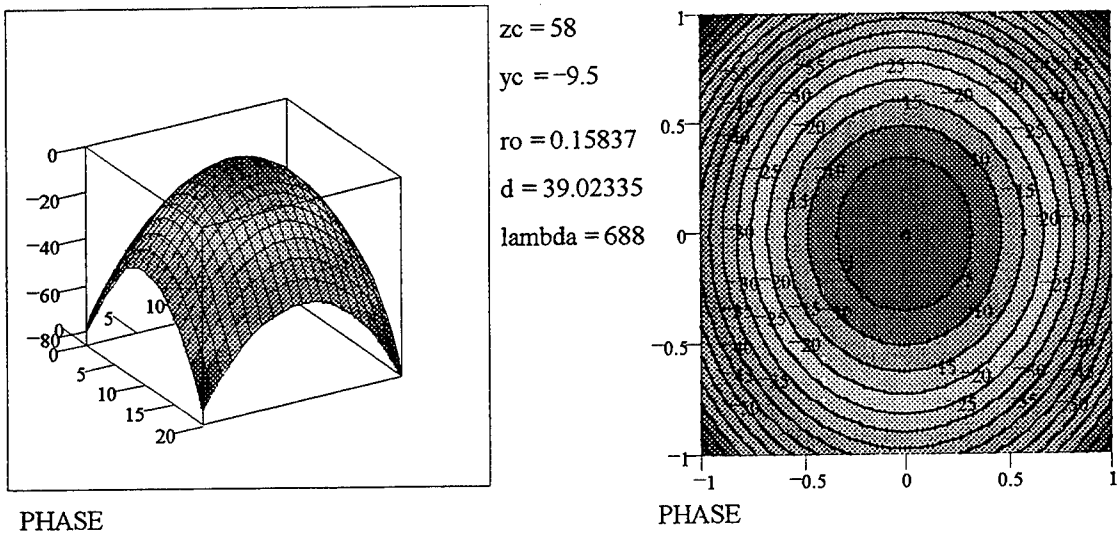


Figure 5.

The wavefront aberration just looks like a spherical wave coming out of the aperture, and so it is. This is because the position of the image point is chosen incorrectly, or rather, the aberration has shifted the position of the image point. This aberration is called curvature of field. It is closely connected to astigmatism.

If the camera is imaging the plane containing the image point, with the image point location calculated from the image forming equations, the point will appear to be out of focus. To find out where the image point is, we must correct the image point distance until we find where it is. This is done by adjusting the image point distance until the curvature of field wave aberration is eliminated. Another criterion to use is to adjust the image point distance until the resolution is the best possible. The next figure shows this accomplished for the case in the preceding figure.

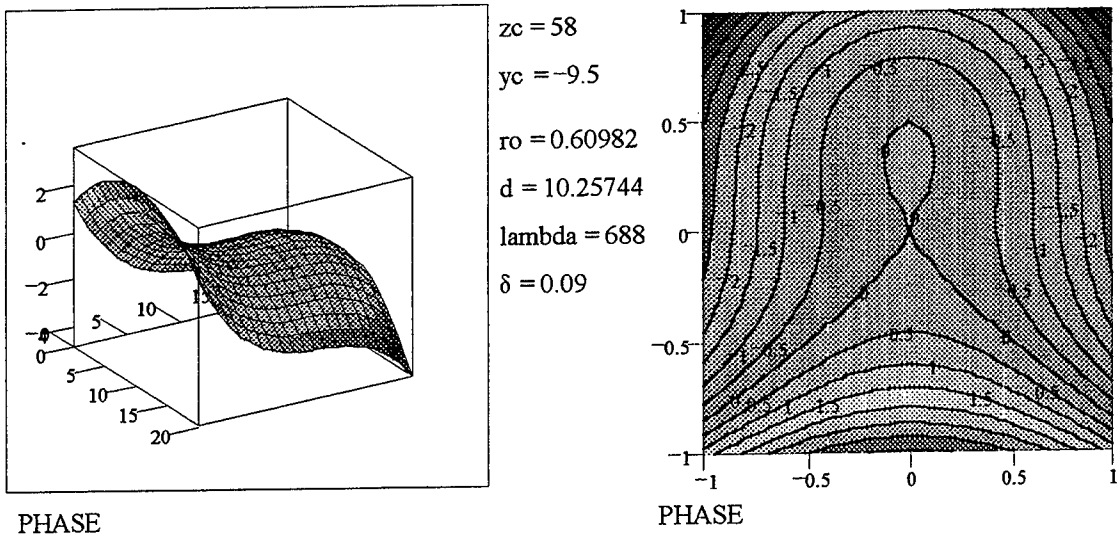


Figure 6.

Yet another criterion for recognizing curvature of field is that you have a situation in which the *average* wave aberrration is non-zero. Correcting for curvature of field means placing the image position so that the average wave aberration is zero. This is the method used by Rebordao[4] to eliminate curvature of field.

The parameter δ is the correction to the image position, in inches, necessary to correct for curvature of field. We see that the resolution, d, has gone from d=79 microns, for no correction, to d =10 microns, with the correction. We see from the value of δ that the image point is .09 inches further away from the pupil than predicted by the image forming equations.

The camera would image this corrected image position to get the sharpest picture. The eye would naturally accomodate to this plane, because there the object would appear sharpest. Looking at the image in this plane is looking at the circle of least confusion.

2.7 Region of minimal aberration, 688 nm.
Corrected for curvature of field.

From the preceding, it is evident that a more realistic picture of image forming performance will come about if curvature of field is taken into account. The next figure shows the result.

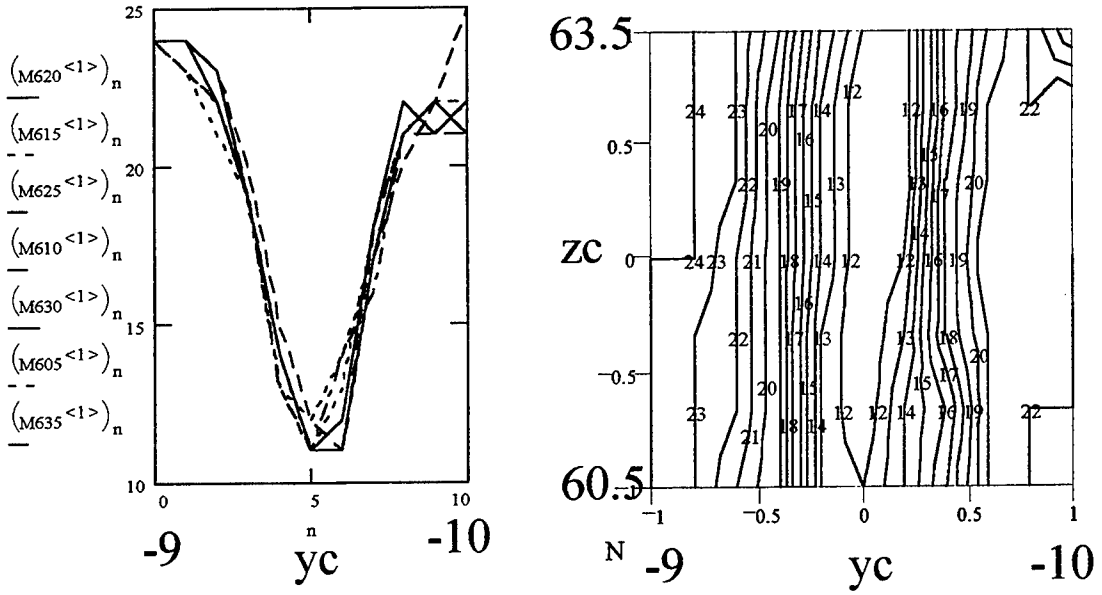


Figure 7.

This is to be compared with figure 4. Compared to figure 4, we see an improvement in resolution in the region of minimum aberration. Also, we see that the region of minimum aberration is more parallel to the hologram axis. This is convenient. It means that position of the reconstruction beam along the z axis is not very critical. From the diagram we conclude that the region of minimum aberration is $60<z_c<64$ inches and $-9.4<Y_c<-9.6$ inches. Notice that the restrictions on z_c are much less stringent than they were before we corrected for curvature of field.

2.8 Region of minimal aberration, 632 nm.

We repeat the above for a 632 nm, He-Ne reconstruction beam. For comparison purposes we first show the results without correction for curvature of field.

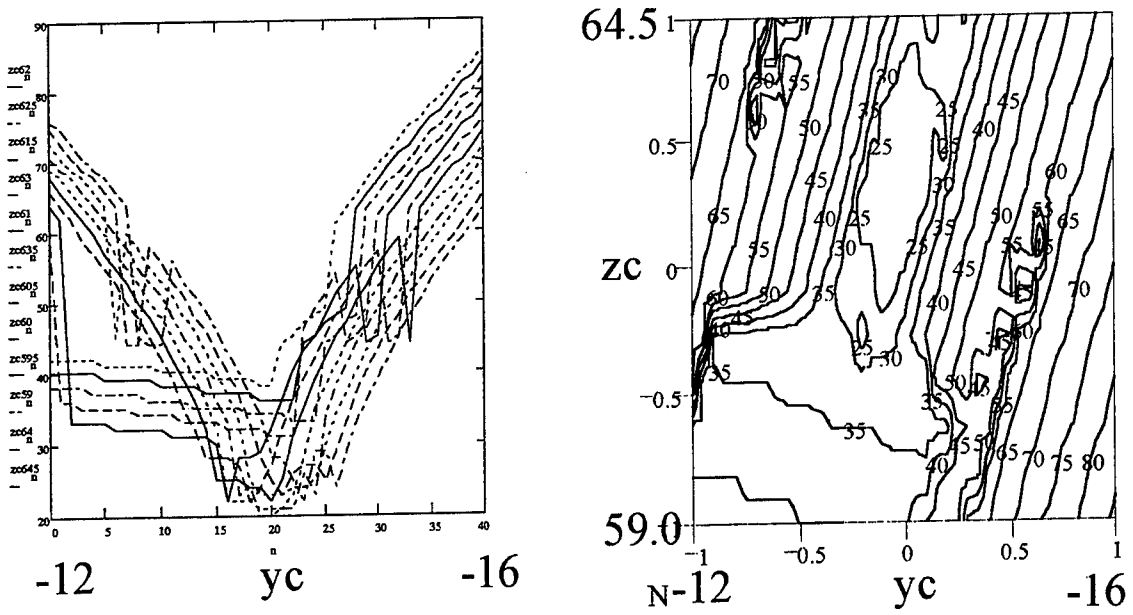


Figure 8.

This should be compared with figure 4. We see from this diagram that the region of minimal aberration is at y_c near -14 in. for 632 nm. This compares to Y_c near -9.5 in. for 688 nm from figure 4.

The region of minimum aberration seems to be near where we want to put the reconstruction beam to satisfy the Bragg condition. This is a lucky circumstance. It means that we will get our brightest image at the same place that minimizes aberrations. However this is a numerical observation. I have no concrete theoretical reason why the region of minimum aberration should coincide with the position that satisfies the Bragg condition.

Now we look at the same thing that has been corrected for curvature field.

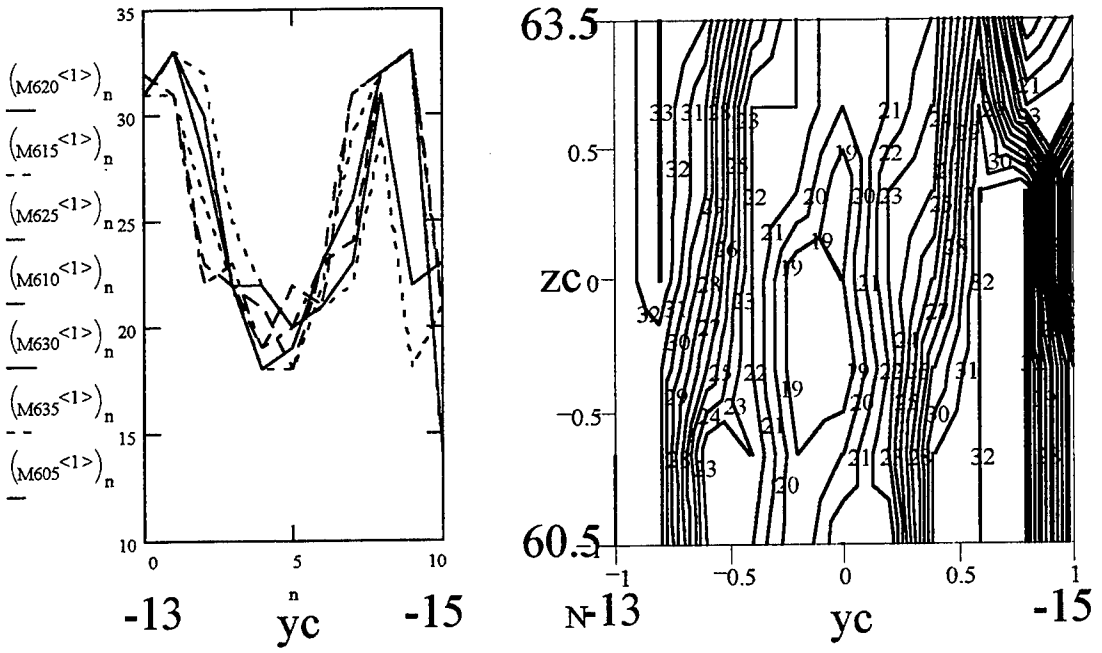


Figure 9.

This figure should be compared with figure 7. In comparion with figure 8, we see that the region of minimal aberration has become more parallel to the hologram axis after correction for field curvature. An additional region of minimal aberration seems to appear at the extreme right hand side of the figure near y_c=-15 in. However I m not prepared at this time to regard this as real, due to numerical anomalies noticed during the calculations. At some future time I will track down the source of the anomalies and reexamine this region.

From figure 9 we conclude that the region of minimal aberration for this hologram at 632 nm reconstruction is $61<z_c<63$ in. and $-13.6>y_c>-14.2$ in.

2.9 Summary.

For a 9 in. radius hologram made with a 694 nm reference beam positioned on the axis of the hologram, 62 in. from the center of the hologram[8] we have the following results.

For a 688 nm reconstruction beam, the reconstruction of a point at the center of the hologram will occur with minimal aberration if the

[8] For a hologram 10 in wide, this puts the reference beam 57 in. from the end of the hologram.

reconstruction beam is place anwhere from 60 to 64 in. along the axis from the center of the hologram, and from .4 to .6 inches behind the axis of the hologram[9]. The best theoretical resolution in this region is from 11 to 15 microns.

For a 632 nm reconstruction beam, the reconstruction of a point at the center of the hologram will occur with minimal aberration if the reconstruction beam is place anwhere from 61 to 63 in. along the axis from the center of the hologram, and from 4.6 to 5.2 inches behind the axis of the hologram. The best theoretical resolution in this region is from 18 to 22 microns.

2.10 Changing the parameters.

If the parameters of the problem change, for example if the position of the reference beam is changed or we go to a hologram of different diameter, what happens to these results? In principal these calculations have to be done all over again in order to find out. However a rule of thumb that works for the two cases considered in this work is that the reconstruction beam should be placed roughly the same distance along the axis from the hologram as the reference beam and displaced perpendicular to the axis to a position that satisfies the Bragg condition at the pupil in question. The so far limited experience implies that this will be near a position of minimum aberration.

[9] Now I am giving measurements relative to the axis of the hologram rather than the face.

COMPUTATIONAL STUDIES OF THE REACTIONS OF CH_3I WITH H AND OH

Paul Marshall

Associate Professor

Department of Chemistry

University of North Texas

PO Box 5068, Denton, Texas 76203-0068

Final Report for:

Summer Faculty Research Program

Wright Laboratory, Materials Directorate, Wright-Patterson AFB, Ohio 45433

Sponsored by:

Air Force Office of Scientific Research

Bolling AFB, Washington, D.C.

and

Wright Laboratory

September 1996

COMPUTATIONAL STUDIES OF THE REACTIONS OF CH_3I WITH H AND OH

Paul Marshall
Associate Professor
Department of Chemistry
University of North Texas
PO Box 5068, Denton, Texas 76203-0068

Abstract

Transition states for attack by H and OH radicals at the C-H and C-I bonds of CH_3I have been characterized at the Gaussian-2 level of theory. The results are employed in a transition state theory analysis to obtain *ab initio* rate constants and product branching ratios. For the H reaction the major pathway is predicted to be I-atom abstraction, while for OH attack H-atom abstraction is faster than HOI formation. The calculated rate constants agree well with available kinetic measurements at around room temperature. The results yield rate constant expressions applicable to combustion conditions and which are employed in a model of a stoichiometric CH_4/air flame to assess the contributions of the two title reactions to the flame chemistry of CH_3I. H-atom attack is the fastest reaction, followed by OH attack. Above 900 K unimolecular decomposition of CH_3I is predicted to be the next fastest removal pathway, followed by O-atom attack.

COMPUTATIONAL STUDIES OF THE REACTIONS OF CH_3I WITH H AND OH

Paul Marshall

Introduction

Interest has recently focused on the reaction kinetics of iodine-containing molecules in two contexts. Solomon et al. [1] proposed that tropical storms rapidly transport CH_3I to the lower stratosphere and argued for a role for iodine in catalytic ozone destruction at altitudes around 20 km in mid-latitudes. The major sources of CH_3I are suggested to be natural emissions from marine plankton and burning of terrestrial biomass [1]. Iodine chemistry is also relevant to combustion. Halon production is banned under the Montreal Protocols on Substances that Deplete the Ozone Layer and therefore there is a search for substitutes for the halon fire extinguishing agents CF_3Br and CF_2ClBr [2]. Potential candidates for some applications are compounds such as CF_3I, whose ozone depletion and global warming potentials are small because it possesses a short atmospheric lifetime [3]. A barrier to detailed understanding of the chemistry of iodine-mediated flame suppression is a lack of kinetic and product branching information for iodine reactions, especially at elevated temperatures. Here we focus on CH_3I, which is the simplest prototype compound containing a C-I bond, and may be formed in flames via $CH_3 + I$ recombination. The present work describes the first *ab initio* analysis of the kinetics of the reactions

$$CH_3I + H \rightarrow CH_2I + H_2 \quad (1a)$$

$$\rightarrow CH_3 + HI \quad (1b)$$

$$CH_3I + OH \rightarrow CH_2I + H_2O \quad (2a)$$

$$\rightarrow CH_3 + HOI \quad (2b)$$

The transition state for 1b has been characterized previously by Schiesser et al. [4]. Our results are employed to derive high-temperature rate constants via transition state theory (TST) and branching ratios for H vs I abstraction. The data help clarify the contributions of channels 2a and 2b to the kinetic measurements of Brown et al. [5]. The potential energy surfaces (PES) for reactions 1 and 2 were analyzed with the Gaussian-2 theory of Pople and coworkers [6], which has been extended to iodine compounds by Glukhovtsev et al. [7]. The present kinetic results show good accord with

the available data for reactions 1 and 2 near room temperature, and are employed to assess the relative contributions of possible pathways for CH_3I destruction in a stoichiometric CH_4/air flame.

Methodology

The general quantum chemistry methods are outlined elsewhere [8][9], and have been implemented with the GAUSSIAN 92 and 94 codes [10][11]. Computations were carried out on a SGI Power Challenge computer. The geometries and vibrational frequencies (scaled by a standard factor of 0.893 [12]) of the stationary points on the PES were initially obtained at the self-consistent field (Hartree-Fock) level using the 6-31G(d) atomic basis set. Minima and transition states were distinguished by the number of imaginary frequencies, 0 or 1, respectively. The stability of the wavefunction with respect to relaxation of internal constraints [13] was verified for each stationary point. Next, the geometries were refined with a partial correction for the effects of electron correlation by means of MP2/6-31G(d) theory. The moments of inertia and frequencies were employed to calculate the rate constant k via canonical TST as implemented in the POLYRATE program [14]:

$$\mathbf{k} = \Gamma \frac{k_B T}{h} \frac{Q_{TS}}{Q_{reactants}} \exp\left(-\frac{E_0^{\ddagger}}{RT}\right) \qquad (4)$$

Γ is a one-dimensional zero-curvature tunneling correction, the Qs are the partition functions, and $E_0^{\ddagger}$ is the enthalpy of the transition state (TS) relative to the reactants at 0 K. The latter term is derived from the G2 energies, which approximate QCISD(T)/6-311+G(3df,2p) results by a series of additive approximations [6] based on the MP4/6-311G(d,p) energy. These high-level energy calculations were carried out for iodine-containing species as detailed by Glukhovtsev et al. [7], and are all-electron frozen-core computations without effective core potentials.

Results and discussion

CH_3I + H

As a check on the accuracy of the G2 methodology we have computed $\Delta_f H_0$ for the C-H and

C-I bond breaking reactions 1a and 1b. The results are -9.0 and -61.7 kJ mol^{-1}, in excellent accord with the experimental values of -5.3 ± 6.7 and -62.2 ± 0.9 kJ mol^{-1} [15], and thus the thermochemistry of the C-H and C-I bonds is well-described at the G2 level.

Figure 1 illustrates the computed geometries of the TSs for reactions 1a and 1b at the HF/6-31G(d) and MP2/6-31G(d) levels of theory. The parameters for TS 1b are similar to those obtained by Schiesser et al. [4] who employed effective core potentials, except that they obtained a bent H-I-C group upon inclusion of electron correlation. Table 1 lists the TS frequencies and Table 2 lists the components of the G2 energies and the energy barriers for the two reaction pathways. TST rate constants have been computed in two ways, first using HF frequencies and second using MP2 frequencies. Both calculations employ the same fundamental classical barrier and MP2 geometries, but different vibrational partition functions and zero-point vibrational energies. The rate constants are plotted in Arrhenius form in Fig. 2. The results are consistent with the known bond strengths DH_{298} [15], where the weaker C-I bond (DH_{298} = 239 kJ mol^{-1}) is much more reactive than the C-H bond (DH_{298} = 431 kJ mol^{-1}). This is in accord with previous experimental studies which concluded or assumed that HI was the dominant product [16][17][18][19]. Figure 2 reveals good agreement between the *ab initio* and experimental k_{1b} values at room temperature, where the HF and MP2 rate constants bracket the experimental values. Most of the difference between the two sets of results arises from the different frequencies computed for the doubly-degenerate loose C-I-H bending mode in the TS. The geometric mean of the two TST approaches yields a rate expression which fits the experimental data well. A key piece of information for combustion modeling is this first TST extrapolation of k_{1b} to elevated temperatures, which is summarized in Table 3 along with the estimated branching ratio for the minor CH_2I production channel.

CH_3I + OH

The TS structures for attack by the OH radical at the H and I atoms are shown in Fig. 1. In the latter case no barrier beyond the endothermicity was discernable on the HF/6-31G(d) PES and a TS could be located only at the MP2/6-31G(d) level of theory. G2 analysis revealed that the energy of this apparent TS in fact lies *between* the energies of the reagents and products, which is consistent with the idea that there is no barrier to reaction 2b beyond the endothermicity. This is confirmed by

the QCISD(T)/6-311G(d,p) energies calculated at selected points along the MP2/6-31G(d) intrinsic reaction coordinate (IRC), and plotted in Fig. 3. The energy is below that of HOI + CH_3 at all points. A distinct TS therefore cannot be localized and canonical TST cannot be applied. For this channel k_{2b} was roughly estimated via the relation $k = A \exp(-E_a/RT)$ by assuming a preexponential factor $A = 1 \times 10^{-11}$ cm^3 molecule^{-1} s^{-1}, equal to that measured for CF_3I + OH [20], and an activation energy $E_a = 23$ kJ mol^{-1}. This is equal to the endothermicity at 298 K based on DH_{298}(HO-I) = 216 kJ mol^{-1} [20]. Strictly this is an upper bound and any variational effects on the fairly flat IRC of Fig. 3 would tend to lower k_{2b}.

Visualization of the normal modes of the H-abstraction TS, 2a, showed that the lowest frequency mode is predominantly internal rotation of the OH group. Analysis by the method of Truhlar [21] using HF and MP2 data yielded reduced moments of inertia of 4.33 and 9.77×10^{-46} kg m^2, respectively, which were employed in the TST calculations. Figure 4 shows that, unlike in the previous system, the more reactive bond is the stronger C-H bond and that $CH_2I + H_2O$ formation, the more exothermic channel, is expected to dominate at all temperatures. A check on the reliability of the kinetic calculations is the comparison with the measurements of the total removal rate constant $k_2 = k_{2a} + k_{2b}$ by Brown et al. [5]. They are seen to be in excellent agreement, as is the room temperature k_2 value obtained by Gilles et al. [19]. Again, the geometric mean of the HF and MP2-based TST results gives the best accord with experiment and provides a TST extrapolation of these measurements to combustion conditions (Table 3). This extrapolation at 2000 K is about 25 times greater than the value obtained from a simple linear Arrhenius extrapolation of the data of Brown et al. [5], which reflects the significant curvature of the TST Arrhenius plot. This predicted curvature is similar to that measured for the analogous reaction of OH with CH_4 [22]. The TST analysis also indicates that HOI formation by OH attack on CH_3I is of only minor importance in both flames and the atmosphere.

Combustion implications

Comparison of the rate constants for reactions 1, 2 and

$$O + CH_3I \rightarrow IO \text{ and other products} \qquad (3)$$

for which $k_3 = 1.0 \times 10^{-11} \exp(+170/T)$ cm^3 molecule^{-1} s^{-1} [19] shows that, even in low-temperature

regions of a flame, attack by H and O atoms on CH_3I is rapid with rate constants that are close to gas kinetic. The major pathway for both processes involves C-I bond cleavage. At higher temperatures OH attack also becomes rapid.

In order to assess the importance of these processes more quantitatively we compare the effective first-order loss rates of CH_3I (the reciprocal of the lifetimes) via reactions 1-3 and unimolecular decomposition in the low-pressure limit

$$CH_3I + Ar \rightarrow CH_3 + I + Ar \qquad (4)$$

for which $k_4 = 4.2 \times 10^{-9} \exp(-21420/T)$ cm^3 molecule^{-1} s^{-1} [23]. An adiabatic premixed stoichiometric CH_4/air flame at atmospheric pressure was modeled with CHEMKIN [24] using the GRI-Mech mechanism [25]. The results apply to trace quantities of CH_3I (i.e. the C/H/O chemistry was assumed to be unaffected). Of course at fixed pressure the rate of unimolecular decomposition depends only on the temperature, while the rates of the bimolecular steps depend both on the temperature and the radical concentrations. As may be seen from Fig. 5, the dominant removal pathway is via H-atom attack, because [H] > [O]. The next most important removal pathway, about an order of magnitude less effective than reaction 1, is attack by OH. Thus CH_2I will be a minor product of CH_3I decay, with HI as the dominant product. Above 900 K, unimolecular decay and O-atom attack are of comparable importance, although both are slower than reaction 2.

Conclusions

Ab initio data have been employed to derive rate expressions for the reactions of CH_3I with H and OH radicals, and to show that major products of H-atom attack are HI + CH_3, whereas OH attack yields predominantly H_2O + CH_2I. Where the rate expressions can be compared with experimental results, theory is in good accord and provides a means to extrapolate limited-temperature information over wider temperature ranges. This shows that *ab initio* calculation of rate constants is sufficiently accurate to survey a wide range of possible steps in combustion mechanisms, in order to help limit the number of processes to be included in combustion models and to aid in the identification of key reactions that merit more detailed laboratory study. For example, in a stoichiometric CH_4/air flame the fastest path for CH_3I destruction is shown to be via H-atom attack, followed by OH attack, with lesser roles for O-atom attack and unimolecular decomposition.

References

1. S. Solomon, R. R. Garcia, and A. R. Ravishankara, *J. Geophys. Res.* 99 (1994) 20,491.

2. W. Tsang and A. W. Miziolek, eds, *Halon Replacements: Technology and Science* (ACS Symposium Series 611, 1995).

3. S. Solomon, J. B. Burkholder, A. R. Ravishankara, R. R. Garcia, *J. Geophys. Res.* 99 (1994) 20,929.

4. C. H. Schiesser, B. A. Smart, and T.-A. Tran, *Tetrahedron* 51 (1995) 3327.

5. A. C. Brown, C. E. Canosa-Mas, and R. P. Wayne, *Atmos. Environ.* 24 (1990) 361.

6. L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, *J. Chem. Phys.* 94 (1991) 7221.

7. M. N. Glukhovtsev, A. Pross, M. P. McGrath, and L. Radom, L. *J. Chem. Phys.* 103 (1995) 1878.

8. W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley, New York, 1986).

9. J. B. Foresman and Æ. Frisch, *Exploring Chemistry with Electronic Structure Methods* (Gaussian, Pittsburgh, 1993).

10. M. J. Frisch, G. W. Trucks, M. Head-Gordon, P. M. W. Gill, M. W. Wong, J. B. Foresman, B. G. Johnson, H. B. Schlegel, M. A. Robb, E. S. Replogle, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzalez, R. L. Martin, D. J. Fox, D. J. DeFrees, J. Baker, J. J. P. Stewart, and J. A. Pople, *GAUSSIAN 92* (Gaussian, Pittsburgh, PA, 1992).

11. M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, *GAUSSIAN 94* (Gaussian, Pittsburgh, PA, 1995).

12. J. A. Pople, A. P. Scott, M. W. Wong, and L. Radom, *Israel J. Chem.* 33 (1993) 345.

13. R. Seeger and J. A. Pople, *J. Chem. Phys.* 66 (1977) 3045.

14. R. Steckler, W.-P. Hu, Y.-P., Liu, G. C. Lynch, B. C. Garrett, A. D. Isaacson, D.-h. Lu, V. S. Melissas, T. N. Truong, S. N. Rai, G. C. Hancock, J. G. Lauderdale, T. Joseph, and D. G. Truhlar, *POLYRATE* version 6.5 (University of Minnesota, Minneapolis, 1995).

15. D. F. McMillen and D. M. Golden, *Ann. Rev. Phys. Chem.* 33 (1982) 493. Data extrapolated to 0 K using *ab initio* H_{298}-H_0 values.

16. I. O. Leipunskii, I. I. Morozov, and V. L. Tal'roze, *Dokl. Phys. Chem.* 198 (1971) 547. Russ. orig. p. 136.

17. M. R. Levy and J. P. Simons, *J. Chem. Soc. Faraday Trans. 2* 71 (1975) 561.

18. A. Sillesen, E. Ratajczak, and P. Pagsberg, *Chem. Phys. Lett.* 201 (1993) 171.

19. M. K. Gilles, A. A. Turnipseed, R. K. Talukdar, Y. Rudich, P. W. Villalta, L. G. Huey, J. B. Burkholder and A. R. Ravishankara, *J. Phys. Chem.* 100 (1996) 14005.

20. R. J. Berry, W.-J. Yuan, A. Misra, and P. Marshall, *to be submitted.*

21. D. G. Truhlar, *J. Comp. Chem.* 12 (1991) 266.

22. S. Madronich and W. Felder, *20th Symp. (Int.) Combust.* (The Combustion Institute, Pittsburgh, 1984) 703.

23. K. Saito, H.Tahara, O. Kondo, T. Yokubo, T. Higashihara, I. Murakami, *Bull. Chem. Soc. Jpn.* 53 (1980) 1335.

24. R. J. Kee, F. M. Rupley, and J. A. Miller, *Chemkin-II: A Fortran Chemical Kinetics Package for the Analysis of Gas-Phase Chemical Kinetics* (Sandia National Laboratories Report No. SAND89-8009B, 1991).

25. M. Frenklach, H. Wang, C.-L. Yu, M. Goldenberg, C. T. Bowman, R. K. Hanson, D. F. Davidson, E. J. Chang, G. P. Smith, D. M. Golden, W. C. Gardiner, and V. Lissianski, *GRI-Mech 1.2*, http://www.gri.org.

Table 1. Transition state frequencies, in cm^{-1}.[a]

TS for reaction 1a		TS for reaction 1b		TS for reaction 2a		TS for reaction 2b
HF	MP2	HF	MP2	HF	MP2	MP2[b]
2100 i	1822 i	790 i	916 i	3011 i	2348 i	268 i
287	293	196	67	59	49	40
523	544	196	67	97	130	142
526	568	405	608	316	306	173
889	930	738	777	454	514	575
922	951	738	777	555	633	593
1169	1195	1144	1201	688	796	714
1181	1196	1414	1453	858	941	984
1336	1393	1414	1453	1046	1036	1069
1384	1773	2933	3060	1092	1257	1425
2967	3091	3062	3203	1356	1402	1443
3069	3208	3062	3203	1429	1449	3067
				2969	3089	3222
				3070	3190	3244
				3605	3591	3614

[a]HF/6-31G(d) values scaled by 0.8929 and MP2=FULL/6-31G(d) values scaled by 0.9646.

[b]No TS at the HF/6-31G(d) level of theory.

Table 2. Absolute G2 energies of transition states for CH_3I reactions.

Reaction	MP4/6-311G(d,p)[a]	ΔE(+)[b]	ΔE(2df)[b]	ΔE(QCI)[b]	ΔE(ZPE)[b,c]	Δ[b]	E_0(G2)[a]	$E_0^{\ddagger}$[d]
1a	-6957.21128	-1.42	-65.06	-3.05	+32.47	-27.39	-6957.31092	50.1
1b	-6957.22601	-1.56	-66.99	-5.46	+34.85	-28.13	-6957.32847	4.0
2a	-7032.30952	-9.64	-102.74	-4.27	+40.08	-31.72	-7032.46799	15.5
2b	-7032.29596	-11.89	-114.23	-5.93	+46.26[e]	-32.21	-7032.46415	25.6

[a]In au. 1 au ≈ 2625 kJ mol^{-1}. The G2 energy is $E_0(G2) = E(MP4/6\text{-}311G(d,p)) + \Delta E(+) + \Delta E(2df) + \Delta E(QCI) + \Delta E(HLC) + \Delta E(ZPE) + \Delta + 0.00114\ n_{pair}$ au, where the individual terms are defined in ref. 6.

[b]Component of G2 energy in 10^{-3} au.

[c]Based on HF/6-31G(d) frequencies.

[d]G2 enthalpy relative to reagents at 0 K, in kJ mol^{-1}.

[e]MP2/6-31G(d) frequencies used.

Table 3: Computed rate constants and product branching ratios for CH_3I reactions.

Reactants	**Major products**	**Minor products**	**Overall rate constant, cm^3 $molecule^{-1}$ s^{-1} [a]**			**Minor product branching ratio**		
			A	n	B	A	n	B
$H + CH_3I$	$CH_3 + HI$	$CH_2I + H_2$	2.14e-15	1.66	300	5.86e-02	0.24	5210[b]
$OH + CH_3I$	$CH_2I + H_2O$	$CH_3 + HOI$	2.72e-24	3.97	-450	3.68e+12	-3.97	3220

[a]In the form $A\ T^n \exp(-B/T)$ for T = 250-2500 K.

[b]For T = 400-2500 K.

Fig. 1 Geometries of transition states computed at the HF/6-31G(d) and in parentheses MP2/6-31G(d) levels of theory. Distances in Å and angles in degrees. (1a) $CH_3I + H \rightarrow CH_2I + H_2$; (1b) $CH_3I + H \rightarrow CH_3 + HI$; (2a) $CH_3I + OH \rightarrow CH_2I + H_2O$; (2b) $CH_3I + OH \rightarrow CH_3 + HOI$.

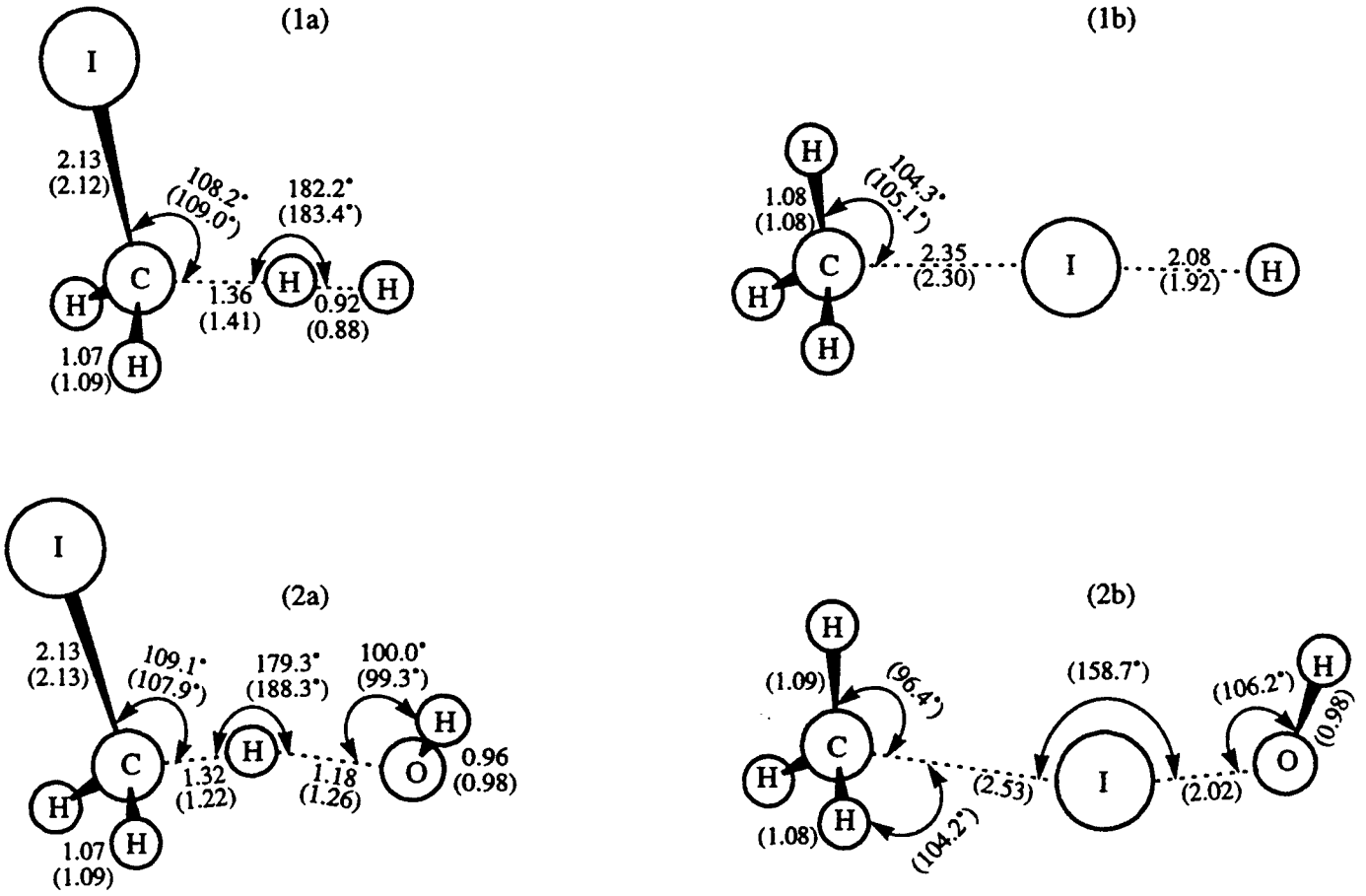


Fig. 2 Comparison of rate constants for H + CH_3I. For the two product channels TST with MP2 frequencies yields the solid line and HF frequencies yield the dashed line. The points are experimental data: (▲) Leipunskii et al. [16], (●) Levy et al. [17], (□) Sillesen et al. [18] and (○) Gilles et al. [19].

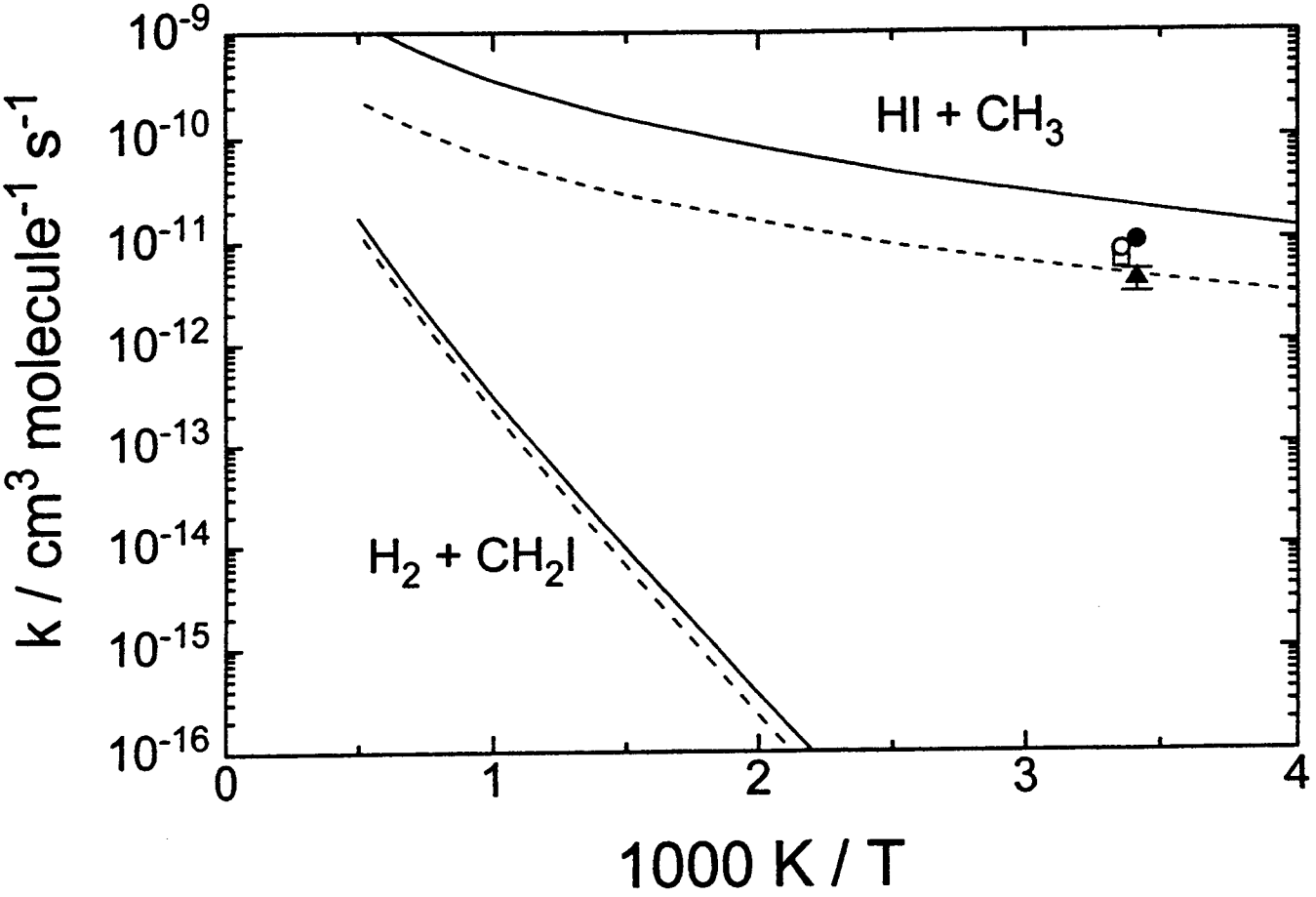


Fig. 3 Evaluation of QCISD(T)/6-311G(d,p) energies along the MP2/6-31G(d) reaction coordinate for OH + CH_3I → CH_3 + HOI.

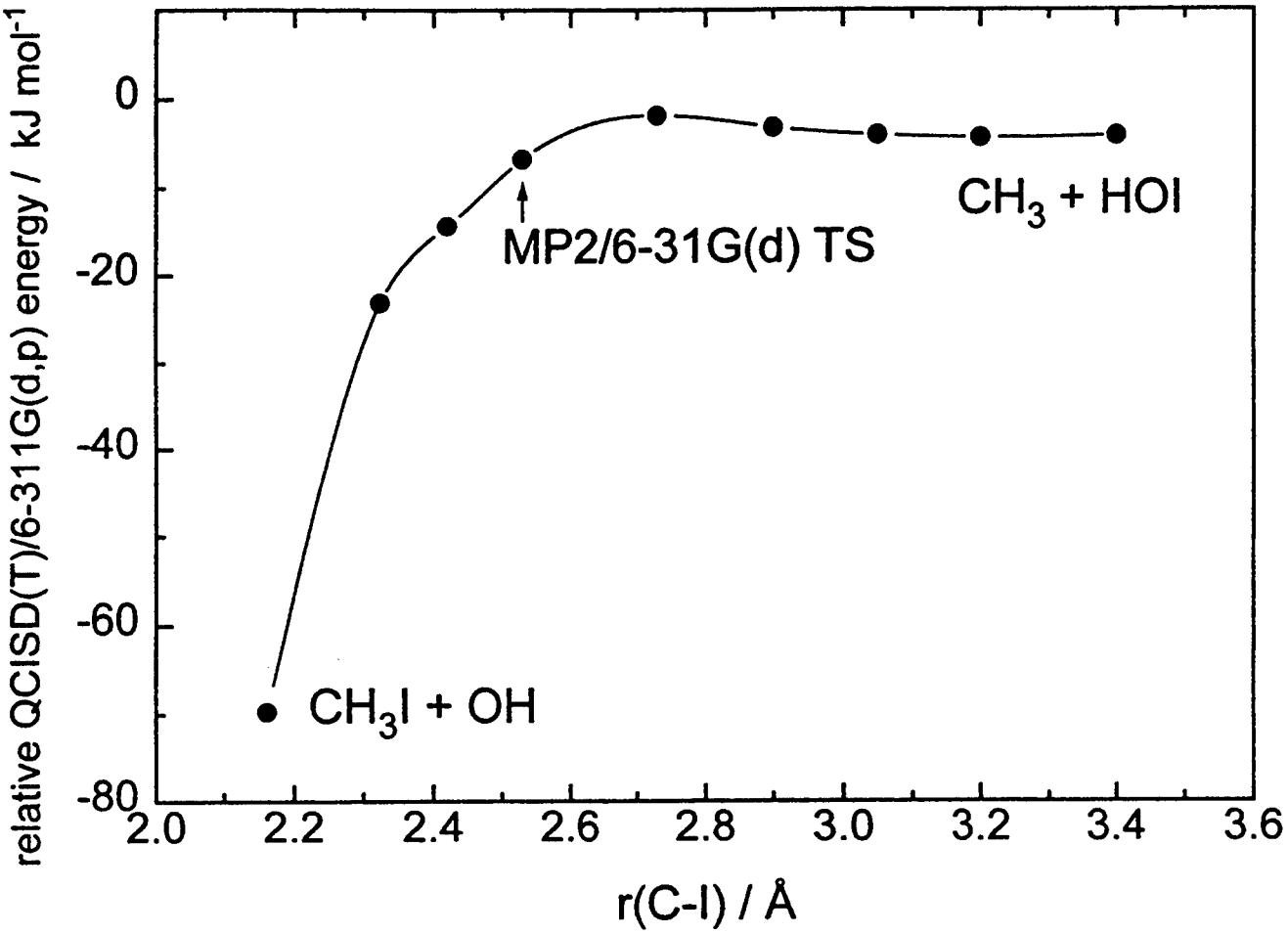


Fig. 4 Comparison of rate constants for OH + CH_3I. For $CH_2I + H_2O$ formation TST with MP2 frequencies yields the solid line and HF frequencies yield the dashed line. The solid line for $CH_3 + HOI$ shows an empirical estimate. The points are measurements of total OH loss: (■) Brown et al. [5] and (●) Gilles et al. [19].

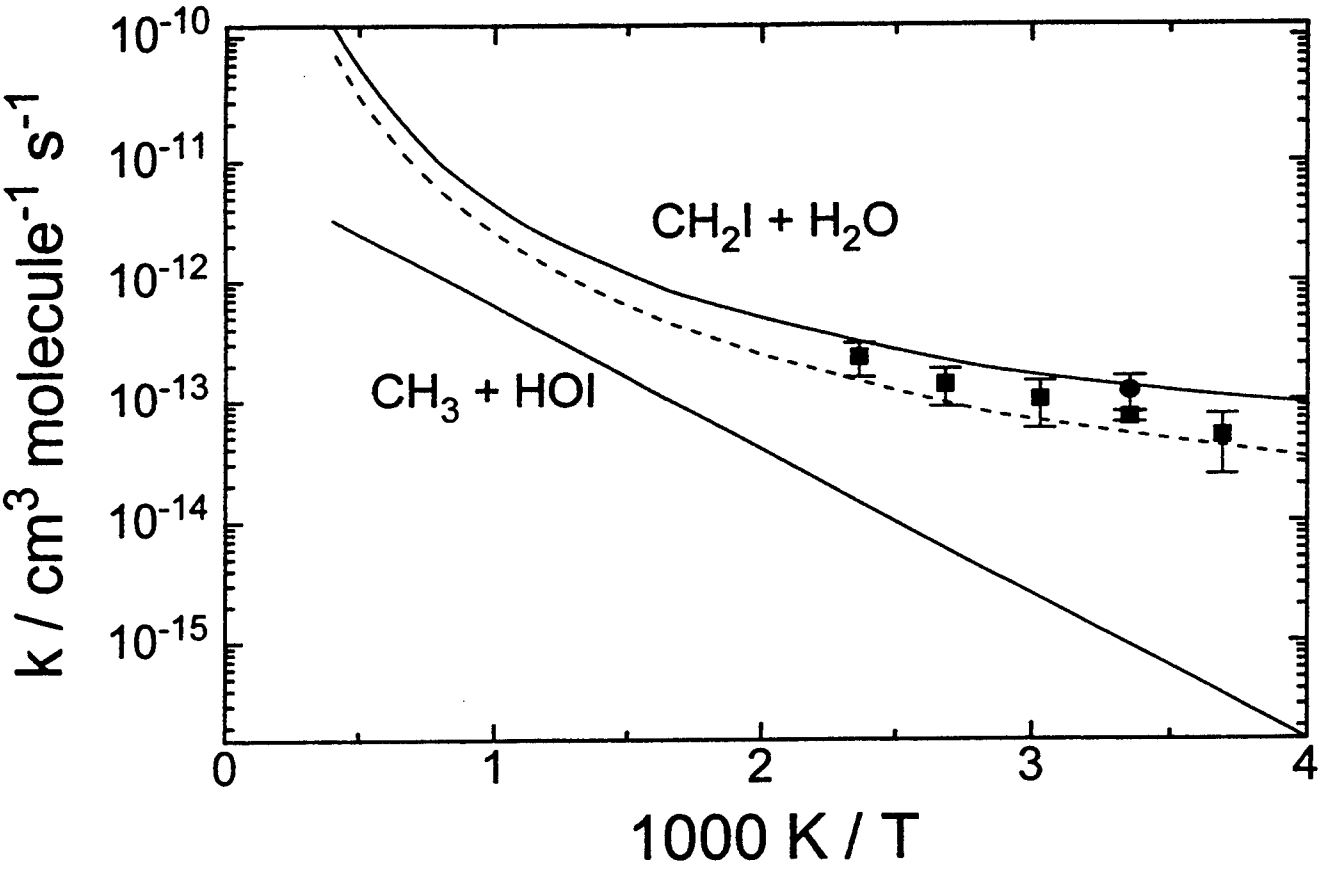
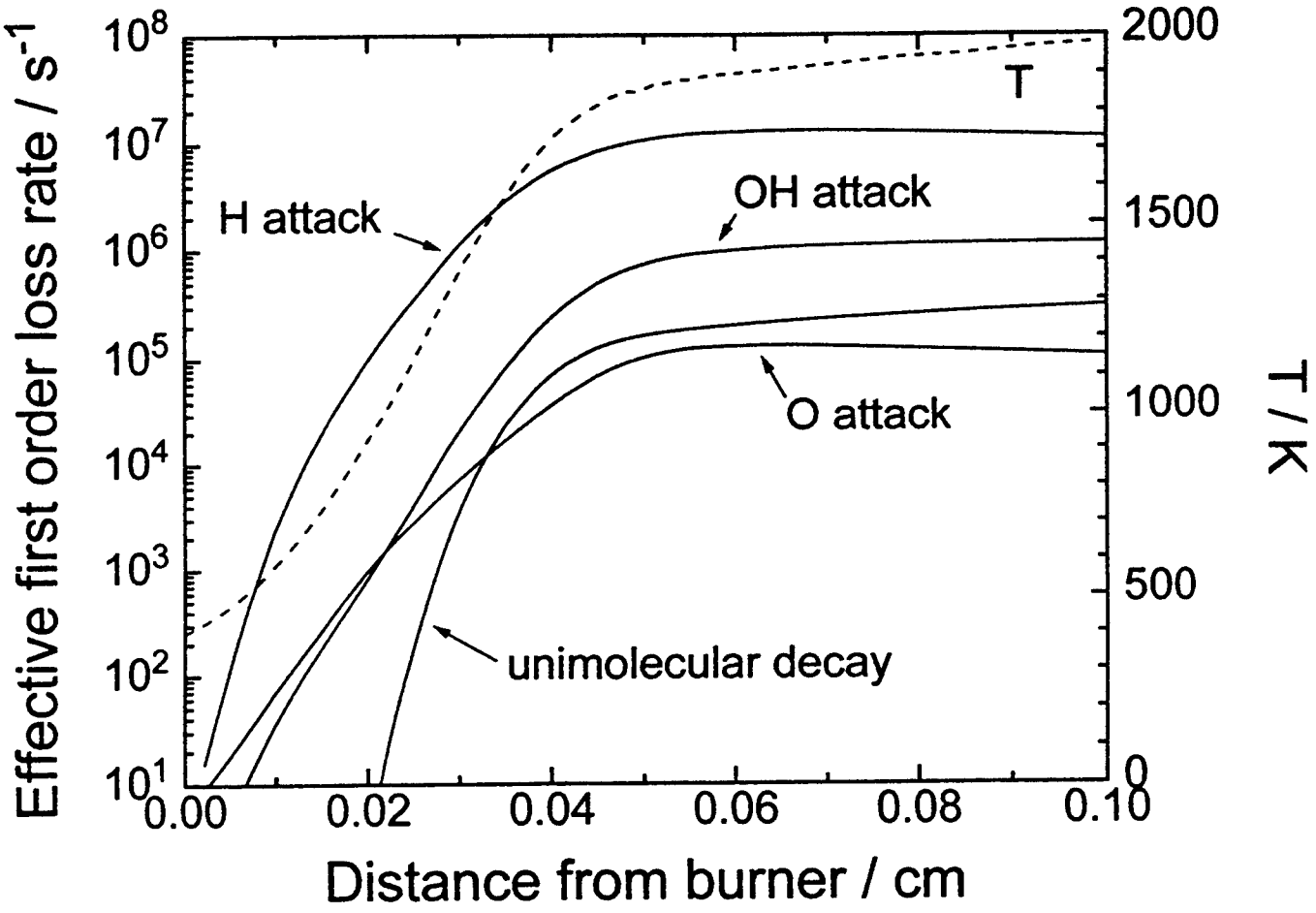


Fig. 5 Reciprocal lifetimes for traces of CH_3I in an adiabatic premixed stoichiometric CH_4/air flame at atmospheric pressure with respect to unimolecular decomposition and attack by H, O and OH (solid lines, left axis), and the temperature profile (dashed line, right axis).



INVESTIGATION OF HOLOGRAPHIC PIV AND HOLOGRAPHIC VISUALIZATION TECHNIQUES FOR FLUID FLOWS AND FLAMES

Hui Meng

Assistant Professor

Department of Mechanical Engineering

Kansas State University

Manhattan, KS 66506

Final Report for:

Summer Faculty Research program

Wright Laboratory

Sponsored by:

Air Force Office of Scientific Research

Bolling Air Force Base, DC

and

Wright Laboratory

August 1996

INVESTIGATION OF HOLOGRAPHIC PIV AND HOLOGRAPHIC VISUALIZATION TECHNIQUES FOR FLUID FLOWS AND FLAMES

Hui Meng
Assistant Professor
Department of Mechanical Engineering
Kansas State University

ABSTRACT

The use of holography in diagnostics of fluid flows and flames is investigated in several flow configurations for two purposes: as a visualization tool and as a quantitative full-field velocity measurement tool. Holography, having the capability of three-dimensional (3D) representation of spatial objects, particle ensembles and multiphase flows, holds great promise as a diagnostics tool — both qualitative and quantitative — for spatially and temporally evolving complex flow structures. Full-field velocity fields can be measured by Holographic Particle Image Velocimetry (HPIV or Holographic PIV), which potentially can fulfill the need for detailed velocity and vorticity data in various turbulent and complex flows encountered in aircraft engine combustion and other Air Force applications.

For holographic visualization, a snap shot of a transient fluid flow is recorded on a hologram which can reconstruct its 3D structure at that instant, while a series of snap shots can record the dynamical evolution of the 3D structure. The 3D images of the flow reconstructed from the hologram can be viewed from different angles, focused at different distances and with different magnifications, allowing analysis of different scales of the flow. This is a very powerful technique for capturing the instantaneous 3D (volumetric) information of turbulent flow phenomena, comparing to standard flow visualization techniques which are mostly restricted to two-dimensional (2D) images. In Holographic PIV, the purpose is to map out the *quantitative* instantaneous volumetric 3D velocity fields with high spatial and temporal resolution. Double exposures in rapid succession are made on a hologram to record millions of flow-tracing particles at two instants, and the hologram then reconstructs the particle images for velocity (displacement) extraction and provides spatially resolved 3D velocity field data. This is an innovative technique; its development is highly challenging. Promising results have been demonstrated in several university research labs as well as in Wright Laboratory in this Summer Research project.

We have developed an innovative holographic flow visualization method based on In-line Recording Off-axis Viewing (IROV) technique, and a basic double-exposure HPIV system which can be conveniently changed between IROV and Off-Axis configurations, at the Wright Laboratory. The IROV configuration enjoys the simplicity of optical geometry and operation while reducing speckle noise by avoiding the superposition of speckle-generating light waves during reconstruction. The Off-Axis configuration allows even higher seeding density by reducing the recording of speckle-generating light waves and by spatial separation of components in reconstruction. It nonetheless involves more complexity in geometry. Our future work include improvement of data processing and implementation of HPIV for vortex-flame interaction studies.

INTRODUCTION

Although it is widely known that fluid dynamics plays a critical role in combustion, the complex interaction between fluid dynamics and combustion is still not well understood. This poor understanding can largely be attributed to the lack of detailed data in the combusting, mostly turbulent flows. Combustion diagnostics techniques have become sophisticated enough to tackle velocity, temperature, species concentrations, and even pressure, albeit limited to single-point or planar measurement. On the other hand, even the simplest jet diffusion flame involves highly transient three-dimensional (3D) turbulent flow structures. To solve the complex problem of measuring volumetric 3D flow velocity (and thus vorticity) fields, it is necessary to resort to new, 3D diagnostics techniques.

Holography, having the capability of three-dimensional (3D) representation of spatial objects, particle ensembles and multiphase flows, holds great promise as a 3D diagnostics tool — both qualitative and quantitative — for spatially and temporally evolving complex flow structures. Holography has been used for diagnostics of particles or aerosols in different flows and combustion (Lee and Kim, 1986). Most recently, it is also developed, as Holographic Particle Image Velocimetry (HPIV or Holographic PIV), to measure full-field 3D velocity fields (Meng and Hussain, 1991; Barnhart et al., 1994; Meng and Hussain 1995 a, b). This leading-edge technique could fulfill the need for detailed velocity and vorticity data in various turbulent and complex flows encountered in aircraft engine combustion and other Air Force applications.

Although Holographic PIV has been demonstrated by at least two research groups to be a promising technique for 3D velocity field measurement (Barnhart et al. 1994, Meng and Hussain 1995a, b). its development has not been widely pursued mainly due to its complexity and stringent hardware and software requirements. Unlike 2D PIV, which makes use of conventional 2D imaging technique (photography or videography), holography is a coherent optical process requiring spatial and temporal coherence of laser beam. Moreover, the holographic imaging and data processing of a volume of a large number of particles (as required by resolving turbulence structures) present challenges both optically and digitally, and the operation of a holographic PIV system currently still appears cumbersome. Considering all these factors, we first develop a basic HPIV system based on a relatively simple and user-friendly holographic technique, In-line Recording Off-axis Viewing (IROV) (Meng and Hussain 1995a,b) and then extend the system to other configurations, based on Off-Axis holography.

On the other hand, holographic flow visualization should be relatively simple; however, its power (3D capability) has not been fully realized in the past. Meulen and Raterink (1977) applied holography to boundary layer and cavitation studies based on the refractive index difference between an injected dye solution and the fluid flow, but their pictures appear to be photographs of 2D planes taken from 3D reconstructed images. Timko (1980) used in-line and off-axis configurations to analyze particle sizes and distributions. The pictures, which also appear to be photographs of 2D planes taken from 3D reconstructed images, in which individual drops of fuel-oil atomized from a nozzle are indiscernible, thus offering a visualization of the overall field instead. Ruff et al. (1990) have used high speed in-line holocinematography techniques to record the 3D time evolution of multiphase flows, such as the 3D time dependent tracks of glass beads (~500 mm) striking a flat surface. They showed 2D top views of the evolution of the

initially spherical drop as it bounces off the metal surface. The present study takes fuller advantage of the three-dimensionality of the holographic technique by showing many views of the 3D reconstructed image (different angles, focuses and magnifications).

EXPERIMENTAL METHODS

Holographic imaging systems

The general recording configuration of the holographic imaging system is shown schematically in Figure 1(a) for both IROV and Off-Axis approaches. Using only one beam, BC2, and placing the flow test section at FIL (stands for flow in-line), the system works as an IROV recording system. With both BC1 and BC2 present and the test section placed at FOA (stands for flow off-axis), the setup becomes an Off-Axis based HPIV. The IROV configuration enjoys the simplicity of optical geometry and operation while reducing speckle noise by avoiding the superposition of speckle-generating light waves during reconstruction. The Off-Axis configuration allows even higher seeding density by reducing the recording of speckle-generating light waves and by spatial separation of components in reconstruction. It nonetheless involves more complexity in geometry.

The two frequency-doubled Nd:YAG pulsed lasers (10 and 30 Hz) use injection seeding to provide a coherence length > 1 m. The laser beam (532 nm, duration ~ 10 ns) from each laser is directed to a pinhole (diamond, 100 microns) for spatial filtering. The pinhole is located at the focal point of the positive lens L1 (f = 1 m). The beam, which now has a clean gaussian profile, is passed through a positive lens (250 mm) and narrows before it joins the second beam, from the other laser, which has undergone similar filtering. The combined beam is then split into two with a beam splitter (BS). For the In-line recording configuration, the combined beam BC1 is dumped, while the combined beam BC2 is used simultaneously as both the object-illumination beam and the reference beam. The collimator expands the beams to a diameter of 75 mm that illuminate a cylindrical volume in the test section, marked as FIL. The scattered light from the particles in FIL configuration produces the object wave by forward scattering. The interference between the reference wave (light that passes through the flow unscattered) and the object wave are recorded in the holographic plate (HP), placed behind the test section. In the Off-Axis configuration, BC1 is used as the object-illumination beam and BC2 as the reference beam. The flow facility is now moved to FOA. Only the illumination beam crosses FOA. The object beam is formed by side scattering light from the particles and reaching HP. The unscattered light from BC2 does not hit HP. A diverging lens and a converging lens (DL and L3) produce a slowly diverging beam used for strong illumination (to produce strongest object beam) of the flow particles. The Off-Axis angle is optimized so that it produces maximum intensity object beam and minimum external noise. In all cases, the two lasers are synchronized and the time delay between the two lasers is monitored by a photo-diode. For the present experiments (low velocities), the time delay range was set from 50 ms to 150 ms. The shutters (S) open to allow only one pulse from each laser (in the general double-exposure case).

The reconstruction of the holograms was performed using a Millennia CW Nd:YVO_4 laser, which produces coherent light of the same wavelength (532 nm) as the recording Nd:YAG lasers. The reconstruction schematic is shown in Figure 1 (b). The beam is directed to a microscope lens (typically x10 times magnification) and a pinhole to

produce a slowly diverging beam. This beam is collimated into a plane wave that reproduces the reference beam used during the recording. The developed (with D-19 Kodak developer) holographic plate is placed normal to the beam so that the real and the virtual 3D images of the object recorded can be seen in each side of the plate, respectively.

The test section can accommodate any of the flows studied. It is equipped with a 3D traverse system that allows for 3D motion of the flow facilities. A suction on top of the test section allows the withdrawal of particles and gases from the flows and flames. The different flow facilities and seeding particles are described below. Most of the air facilities were chosen to generate reproducible 3D vortical flows in unconfined environment and with well-known features. This last property was desirable so that measurements obtained by holography could be verified. In addition, the results should prove the superiority of the holographic technique by giving additional insight in the behavior of those flows that are still not fully understood.

In all cases, the flow was kept as close as possible to the holographic plate for optimal recording of the scattered light. For the flame case, the distance was larger than others (several cm) to protect the holographic plate from the light and heat generated by the flame that could damage the emulsion. The AGFA 8E50 plates have the desirable high resolution (>5000 lines/mm) and sensitivity (insensitive to infrared). Particles on the order of magnitude of 1 microns can be resolved.

Water experiments

Initial tests of the techniques were conducted in water for its simplicity for flow visualization, our previous experience in the HPIV technique applied to water flows (Meng and Hussain, 1995a, b), and easier seeding density control. Several seeding materials were used for different purposes. The holographic visualization was tested with large (typical diameter of 3.5 mm) drops of milk and drops of aqueous suspensions of polymer microspheres (neutrally buoyant with an average diameter of 14.9 μm) injected into a water tank (7.62x10.16x27.94 cm^3). The drops were manually injected into the water using a drop counter; they were let freely fall (with gravitational momentum) from 1 cm above the water surface. These preliminary experiments helped in the determination of the effect of seeding densities and particle sizes in the holographic visualization and HPIV.

Vortex ring generator

The vortex ring flow is a simple 3D flow with well-known features and will be useful in planned vortex-flame interactions studies. The generator, shown schematically in Figure 2, consists of a cylinder 15 cm long and 8.9 cm in diameter, large enough to accumulate enough haze for an experiment run. The bottom end accommodates a 8 W speaker driven by a function generator (HP 3310B) and a power amplifier (HP 467A). Exhaust ports are located between the speaker and the chamber to release backlashed haze. The top is a circular disk, 0.5 cm thick, and with an orifice of 3.17 cm in diameter in the center. The mushroom shape, the wake, and the characteristics of the vortex ring itself (strength, diameter, core diameter, etc.) can be controlled through the frequency and amplitude of the driving functions (sinusoidal, square, ramps, etc.). This allows to generate vortices of different shapes at frequencies multiples of that of the pulses of the lasers, so that synchronized and phase-locked measurements are possible. The seeding

particles were glycerin of < 1 μm of diameter generated with a G300-haze generator, and could be easily introduced in the main cylindrical chamber through the ends.

Shear Layer facility

A single air stream from a compressed air supply at 45 psia was used to generate the shear layer. The facility, shown schematically in Figure 2, is a two-stream mixing layer system. It consists of a diffuser (35 cm long, with small and long diameters of 16 cm and 25 cm respectively), a set of smoothing devices (screens and honeycombs), and a contraction (18 cm long) that converts the circular diffuser into a square end (12.7 cm side) to accommodate the plane mixing layer. The end of the splitter wall that divides the facility into two identical chambers is the slowly converging splitter plate (4 cm long) into a knife edge. In addition to the mixing layer, the facility provides two shear layers at the sharp edges useful for studying unconfined flows. Although the mixing layer can be studied without solid confinements, the shear layers on the sides would obstruct the scattering. For this reason, the shear layer is better; in addition, it can be located closer to the holographic plate. At the free-stream velocity of 1.3 m/s the shear layers had a natural roll up frequency (Ho and Huerre, 1984) of 20 Hz, a multiple of Nd:YAG lasers firing frequency. A speaker was used to acoustically make the rolled up vortices very periodic and to allow phase-locked measurements. The seeding particles were Al_2O_3 (5 and 17.5 μm). Experiments using denser seeding techniques (e.g., $TiCl_4$ mixed with water vapor) are in progress for more effective visualizations.

Flame Facilities

A simple Bunsen burner was used initially to test the feasibility of holography for flame studies. It would provide with a qualitative measure of the effects of heat and light on the holographic emulsion, of the possible distortion of the reference beam by temperature and density gradients, and, since the density is inversely proportional to temperature, of how much increase in the seeding density is necessary. The seeding particles were TiO_2 and Al_2O_3 for visualization and Al_2O_3 (5 and 17 μm) for HPIV.

After holography was proven to work for this simple flame, more complex flames were set up. First, a cylindrical jet with acoustic modulation (through a chamber with a speaker to perturb the propane line at a known frequency) was set up to provide a reproducible vortex-flame system (Figure 2). The driving frequency was again chosen a multiple of the lower Nd:YAG laser frequency, 10 Hz. Another study of more complex flame system involving a jet diffusion flame with co-flowing air is under study. The flame facilities were also run without burning the gas (propane) to study and compare the cold jet flow with the flame.

Display techniques

As it will be explained in the next section, the holographic visualization is based on the scattering of light from particles in highly seeded flows. In some cases, the particle density is so high that the visualization is equivalent to visualizing solid-objects via holographic imaging; in other cases, with less seeding density, it is the speckle generated from the particles what facilitates the visualization. The flow visualization images could be seen with naked eyes in most cases. The images are the 3D replica of the real flow at the time of exposure; hence they can be viewed

from different sides and angles, top and bottom, inside and outside, etc. To facilitate and enhance the in-situ visualization, a set of mirrors and magnification mirrors can be set surrounding the image. To show this feature in this report, photographs taken with a CCD camera from different angles, using different magnifications, and focusing in different parts of the 3D images are shown. The images viewed by a video camera were also displayed on a monitor, showing the 2D plane in focus and normal to the camera, specially used in the particle pairs analysis in HPIV.

RESULTS AND DISCUSSION

Holographic Flow Visualization

In the present study, the IROV technique is used for holographic flow visualization, based on off-axis viewing of the virtual or the real image. Imaging of some structures is made possible by injecting a dense suspension into another fluid, e.g. milk or a polymer microsphere suspension into water. The particles are used to generate speckle (superposition of waves from scattered light from different particles, also responsible for the intrinsic speckle-noise in in-line recordings), which is used constructively to visualize the flow. Typically, the particles will not be uniformly dispersed in the flow (e.g., inside vortex cores less particles are available), so the speckle will not be uniform over the entire flow region. Usually a full spectrum from completely dark areas (no particles) to a very light areas (lots of scattering), is visible through different degrees of shadows. Using the camera and the monitor, individual particles can be seen in the darkest areas, whereas only speckle noise appears in the brightest areas. This visualization technique is very helpful in the qualitative observation of the 3D flows.

In what follows, we discuss experiments of single-exposure holographic visualization of a falling drop of fluid in water, a vortex ring in air, a cold propane circular jet, and a flame. Since our objective is to explore how holography can be used for 3D flow visualization, only the qualitative characterization of these phenomena is studied.

A drop of fluid into water can behave like an impulsive jet flow, which is unstable and rolls up into a ring. The ring also leaves a wave in the surface of the water and a wake behind the ring. Beyond the ring formation, the flow undergoes more complex evolution characterized by cascade. The ring is susceptible to 3D instabilities and breaks down into smaller drops (droplets), which again behave like minijets which become smaller rings ('ringlets'). The cascade follows on into smaller and smaller scales (or droplets and 'ringlets' of secondary generations) until molecular diffusion mixes both fluids. The characteristics of these different features depend on several factors, such as the initial momentum of the falling drop, the viscosity and density of the fluid of the drop, the surface tension, the capillary effects, etc. In fact, depending on the aforementioned characteristics, the initial drop can also behave like a drop in a uniform flow stream, which is susceptible of instabilities such as the 'bag' instability, whereby a nearly spherical drop changes its shape into a surface resembling a bag. The bag develops concave or convex shape depending on the Weber number. This 'bag' then breaks up into droplets by the rim due to secondary instabilities. The bag shape was more often observed for the milk drop than for the polymer-sphere solution drop.

Figures 3 show photographs of the holographic image from two IROV holograms corresponding to two stages of the falling drop flow: the initial drop which takes a donut shape and leaves behind a wake and a surface wave (Fig.

3, top), and the drop cascading into several droplets (Fig. 3, center and bottom). The first picture was taken for a drop of polymer-sphere suspension in water and the a milk drop in water. The two initial drops had a diameter of < 4 mm. The photographs show different aspects of the same 3D holographic image viewed from different angles or focused at different regions. Among other features, it can be seen how the main droplet develops a localized defect which appears to be a preferred site to start the cascade. It is expected that only ideally a symmetrical breakdown (like a 'crown' around the ring) will occur. In practice, some asymmetries will develop due to upstream perturbations (e.g., defect in the injector nozzle). It is noteworthy how each droplet has a trailing 'wakelet', shared by the neighboring droplets. These 'wakelets' remain connected to each other giving the cascade phenomenon a 3D arcade structure, readily seen in the experiments and in these holographic visualizations.

Figure 4 shows two photographs of the holographic image from the IROV hologram of an air vortex ring. The speckle generated by the dense glycerin smoke enlightens the mushroom shape of the ring. This ring is laminar and has not yet developed any 3D axisymmetric instability. The core of vortex can be readily distinguished by the rolled up vortex sheet, the concentric circles around the core center.

Figures 5 (d) show a photograph of the reconstructed image from an IROV hologram of the turbulent cold propane jet flow. The flow was heavily seeded with Al_2O_3 particles of maximum diameter 17.5 μm. The nozzle exit, 1 cm in diameter, can be identified in the image field. Although only one picture is shown here, the 3D holographic image displayed the typical turbulent structures of a circular jet, with many small scales that could be focused, viewed from different angles, etc. Also shown in Figure 5 are photographs of the hologram plates for the cold propane jet (b), the flame (c), and the shear layer (a); these images reveal the footprints of the flows and can be viewed as shadowgraph patterns. . Holographic visualizations of the shear layer and the flame are not shown here because they were too dim. Experiments using denser seeding (TiO_2) are under way to improve these images.

Holographic PIV

While the 3D flow visualization technique provides qualitative characterizations of different flows, flow dynamics cannot be fully understood unless this is complemented by *quantitative* velocity field measurement, which can be provided by HPIV as described below.

Double-exposure HPIV was applied to several of the aforementioned flows for measurement of displacements and velocities, now with lower seeding density. The particles pairs were captured using IROV and Off-Axis configurations. Since the data processing (e.g., particle tracking and 3D correlation) is still under development, no velocity vectors will be shown in this paper. The main objective is to capture clear particle pairs with minimal aberrations. In all cases the particle pairs will be easy to distinguish. Time delays between the exposures ranged from 50 ms to 300 ms. The displacements can be directly measured from the hologram pictures. The window height and time delay is given in each figure caption so that velocities (displacements divided by time delays) can be estimated for each particle pair. The velocities were verified against Pitot tube readings and PIV measurements in several flow areas.

The particles pairs obtained with the IROV technique are shown for cold propane flow from the Bunsen burner in Figure 6, and for the shear layer in Figure 7. The figures show different planes using different magnifications. The holograms were carefully interrogated using a camera mounted on a 3D traverse that was moved through the area of the real image. The figures selected include pictures of the flows near their exit, i.e. the uniform or free-stream flow area. There, the particle pairs are parallel and separated by similar distance as shown in Figure 6(a) and the top right picture of Figure 7. The other pictures in these figures show areas with more interesting features, such as velocity gradients including direction changes. When the particles in the pairs are closer to each other, the area is close to a vortex. The spinning motion generates changes in direction and velocity magnitude. Particles not laying exactly in the focused plane appear out of focus and similar to noise (Fig. 7, top-left). By moving the traverse they can be easily brought into focus while the previously in-focus ones now become blurry. This is typical of browsing through a 3D image from a hologram, . Through interrogation, vortical motions can be tracked and the size or scale of the corresponding eddies calculated (since the size of the areas interrogated is known). An example of vortical motion is shown by the two planes of the shear layer shown in Figure 7 (center and bottom). They display the flow in two areas where the flow has different directions. Since many pairs are in focus in the same plane, the area shown corresponds to a quasi-planar spanwise roller.

The flame holograms obtained with IROV are full of speckle noise; only using big magnifications (e.g., 10:1) can some particles be identified. The speckle are attributed to the destruction of the reference beam by the flame, since the reference beam is also the illumination beam passing through the temperature and density gradients of the flame. In this case, Off-Axis is preferred. The particle pairs for the flame obtained with the Off-Axis configuration are not included in this report.

In data processing to reconstruct the full 3D velocity and vorticity fields from these holograms, further development is necessary. Currently available technique for IROV data processing is a combination of particle tracking and correlation (Meng and Hussain, 1995a), with manual intervention, based on stereoscopic interrogation of the hologram. Off-Axis data processing can be done through 2D-PIV analysis provided sufficient seeding density. An innovative Off-Axis data processing technique is under development.

CONCLUSIONS

A Holographic PIV and holographic flow visualization system has been set up. Preliminary tests have been conducted to study the feasibility of holography in the diagnosis of complex flows and flames. Well-known flows were visualized and measured using IROV and Off-Axis holography configurations. In all cases qualitative visualization and quantitative HPIV gave very good agreement with the expected flow and flame characteristics. The technique is very powerful and promising for studies of flows and flames with many scales, such as turbulence and combustion.

REFERENCES

Barnhart, D.H., Adrian, R.J., and Papen, G.C., 1994, "Phase-conjugate holographic system for high-resolution particle-image velocimetry", Appl. Opt. 33, 30, 7159-7170

Ho, C.M. and Huerre, P., 1984, "Perturbed Free-Shear Layers", Ann. Rev. Fl. Mech., 16, 365-424

Lee, Y.J., and Kim, J.H., 1986, "A Review of Holography Applications in Multiphase Flow Visualization Study", J. of Fl. Eng., 108, 279-287

Meng, H. and Hussain, F., 1991, "Holographic Particle Velocimetry, a 3D measurement technique for vortex interactions, coherent structures and turbulence", Fluid Dyn. Res., 8, 33-52

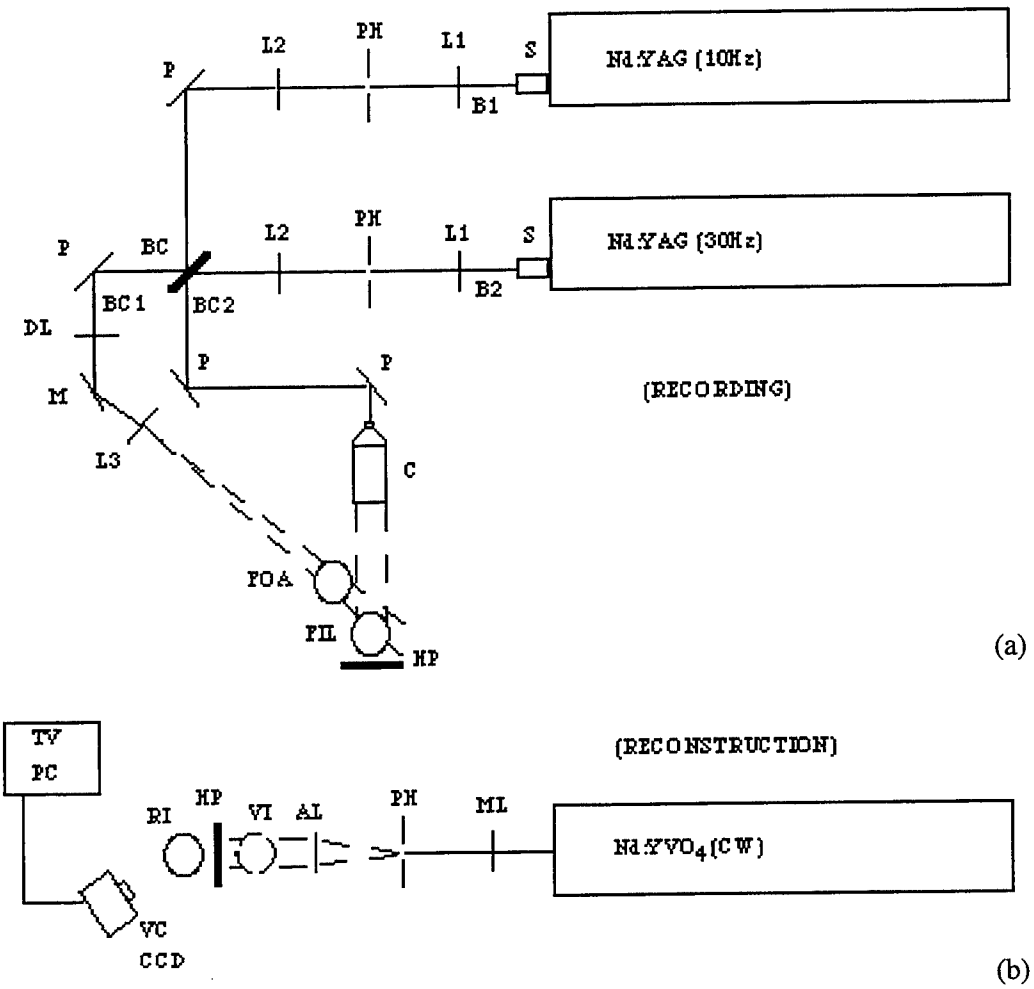
Meng, H. and Hussain, F., 1995a, "In-line Recording and Off-axis Viewing technique for holographic particle velocimetry", Applied Optics, 34, 1827-1840

Meng, H. and Hussain, F., 1995b, "Instantaneous flow field in an unstable vortex ring measured by holographic particle velocimetry", Physics of Fluids, 7, 9-11

Ruff, G.A., Bernal, L.P., and Faeth, G.M., 1990, "High speed in-line holocinematography for dispersed-phase dynamics", Appl. Opt., vol. 29, 31, 4544-4546

Timko, J., 1980, "The Investigation of Transport Phenomena by Applied Holography", Proc. of the Sec. Int. Symp. on Flow Visualization, Bochum, West Germany, 535-540

van der Meulen, J.H.J. and Raterink, H.J., 1977, "Flow Visualization of Boundary Layers in water by in-line holography", Proc. of the Int. Symp. on Flow Visualization, Tokyo, Japan, 357-362



AL : Achromatic Lens
B1,B2: Laser Beams
BC : Beam Combiner
BC1,BC2: Beams Combined
C : Beam Colimator
CCD : Digital Camera
DL : Diverging Lens
FOA: Flow-Off-Axis
FIA: Flow-In-Line
HP : Holographic Plate
L1,L2,L3: Converging Lens
M: Mirror
ML: Microscope Lens
PC : Personal Computer
PH : Pinhole
P: Prism
RI: Real Image
S: Shutters
TV : Televideo monitor
VC : Video Camera
VI: Virtual Image

Figure 1. Optical Layout for Holography: (a) Recording, (b) Reconstruction

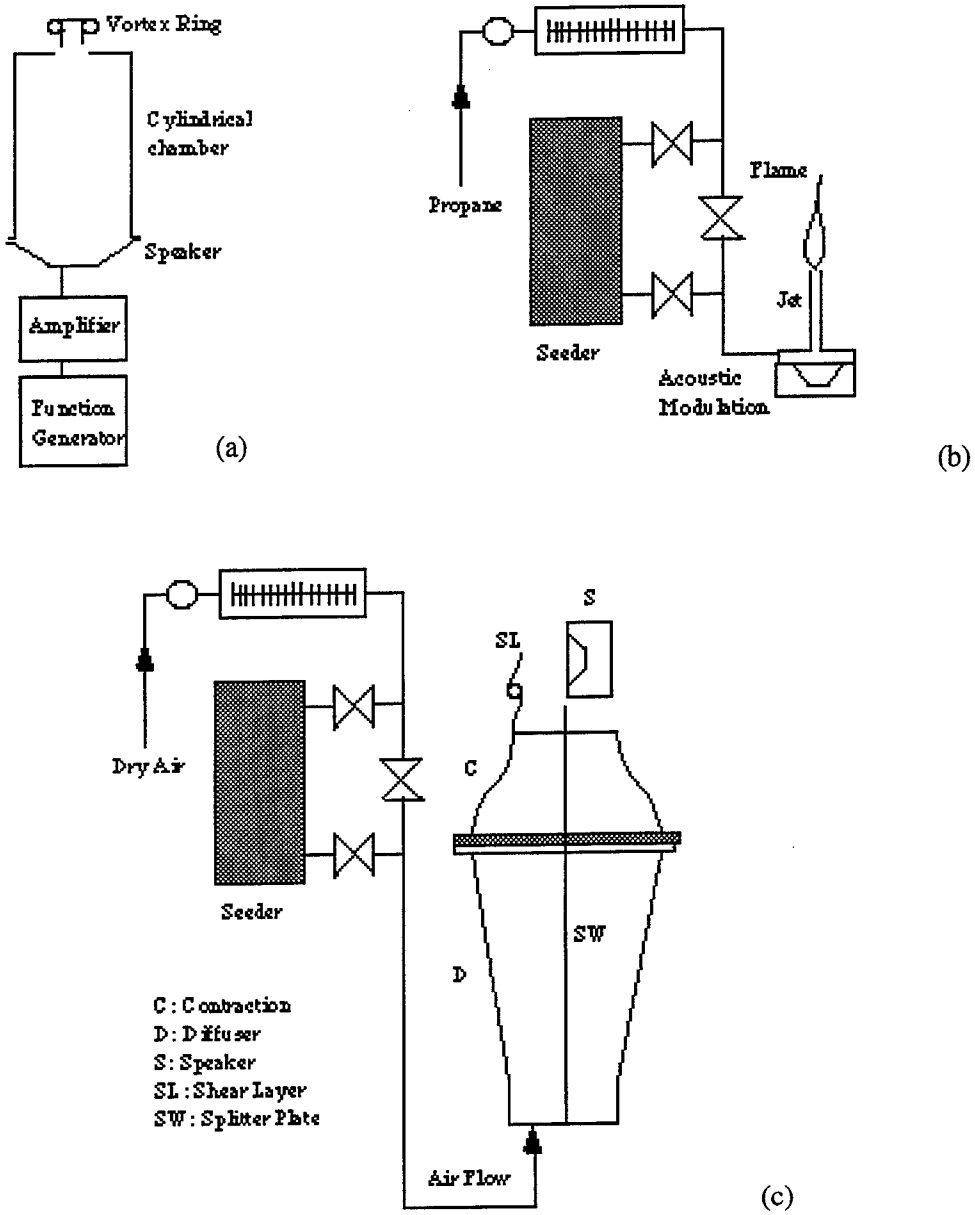


Figure 2. Flow Facility Schematics: (a) Vortex ring generator, (b) Flame generator, (c) Mixing Layer ssytem

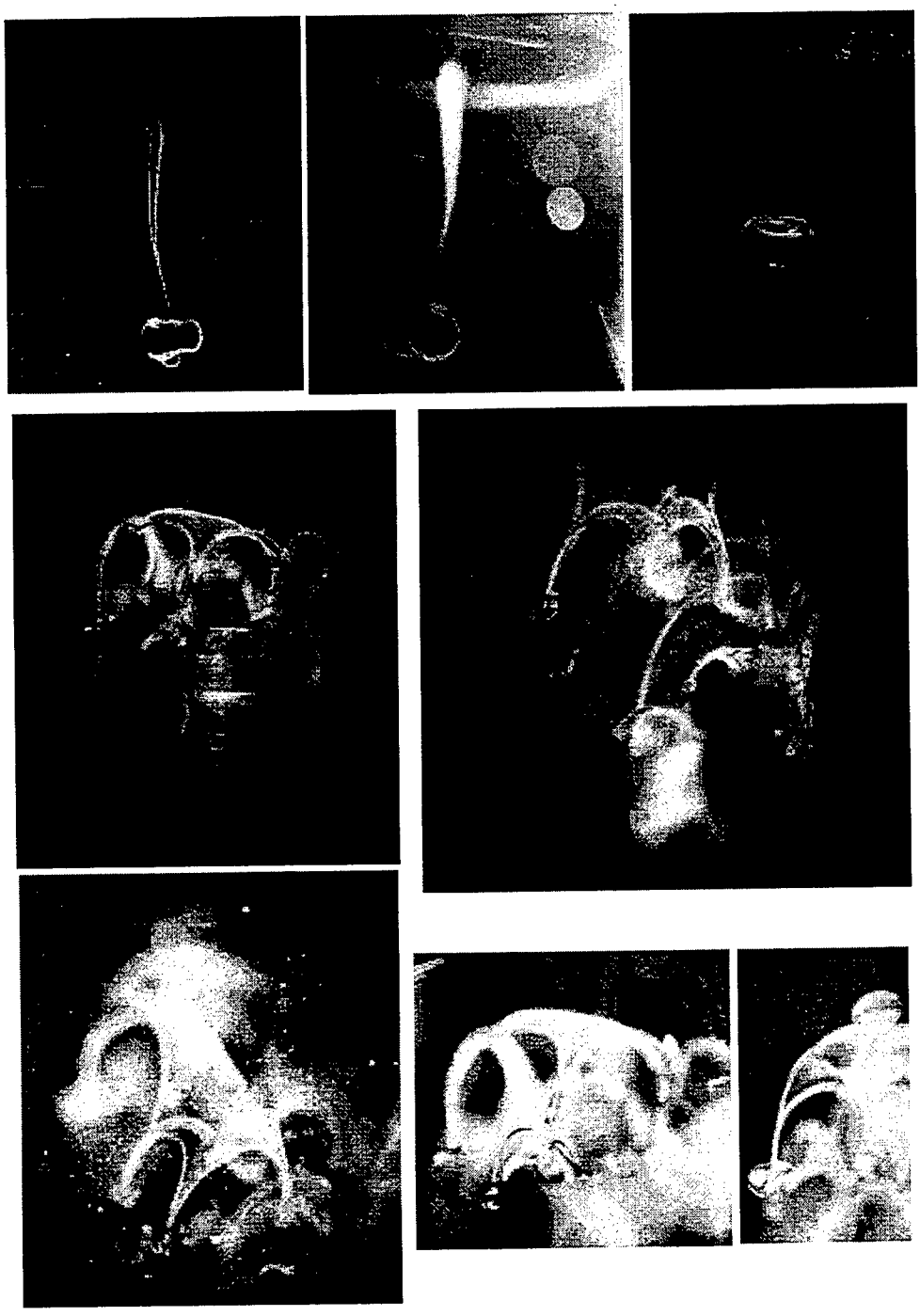
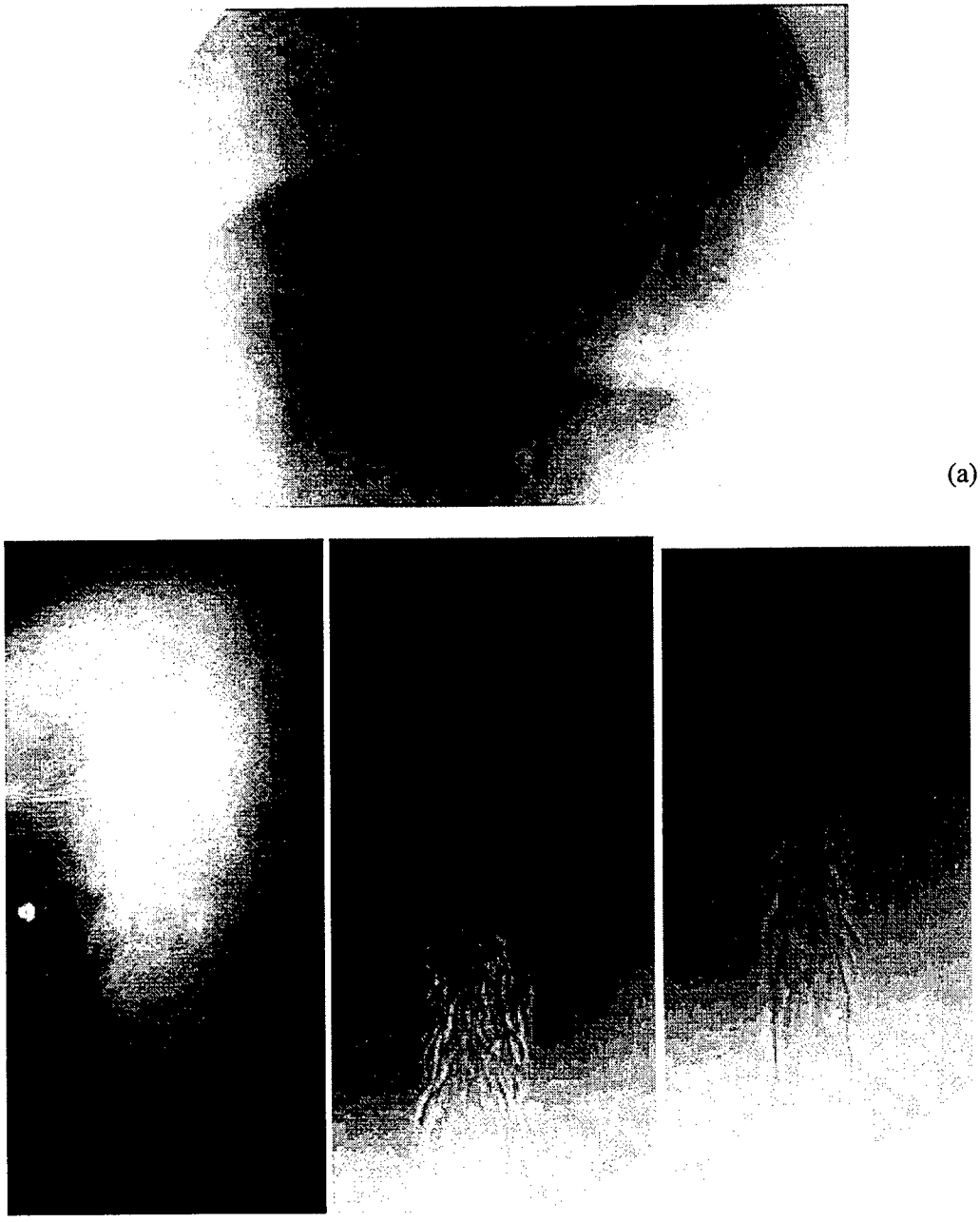


Figure 3. Photographs of reconstructed holographic images of a falling drop viewed from different angles and with different focuses:(a) The donut-shaped drop, (b) The cascade

Figure 4. Two photographs of the reconstructed holographic image of a mushroom-shaped vortex ring in air



(a)

(d) (b) (c)

Figure 5. Photographs of the hologram plates revealing teh shadowgraph of (a) a shear layer, (b) jet with cold propane, and (c) flame. Also shown (d) is reconstructed holographic image of the cold propane jet from hologram (b)

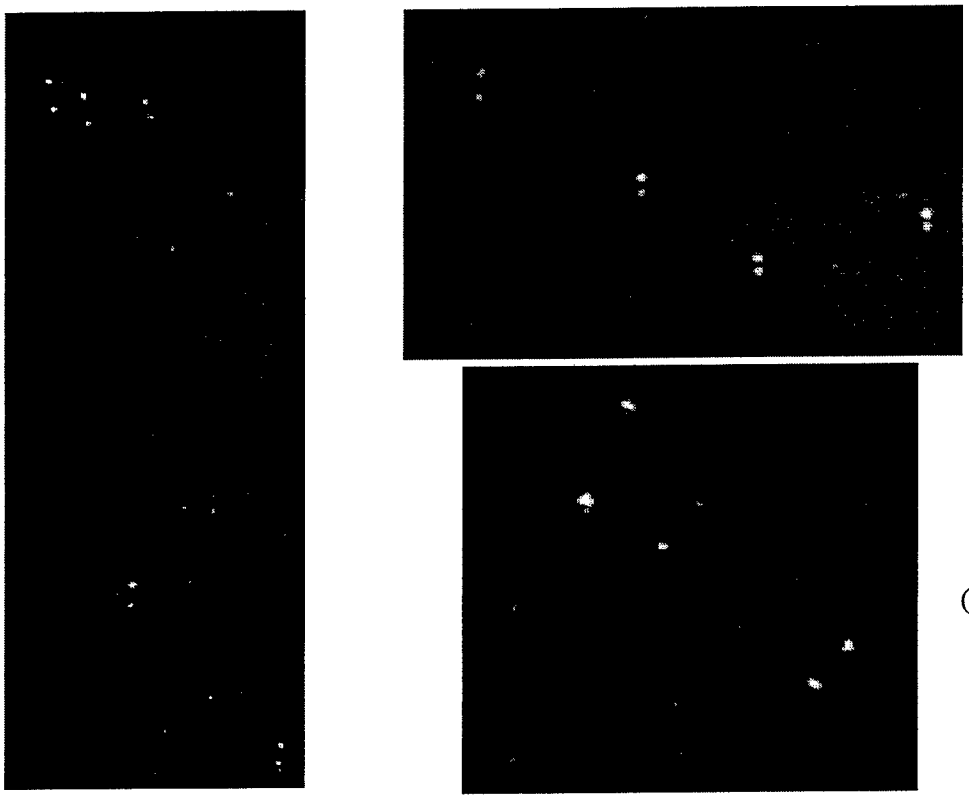


Figure 6. HPIV reconstruction: particle pairs in the cold propane jet reconstructed from an IROV hologram. These photographs were obtained by focusing at different planes and different regions of the 3D field from (a) the uniform flow region, (b) and (c) a velocity-gradient region.

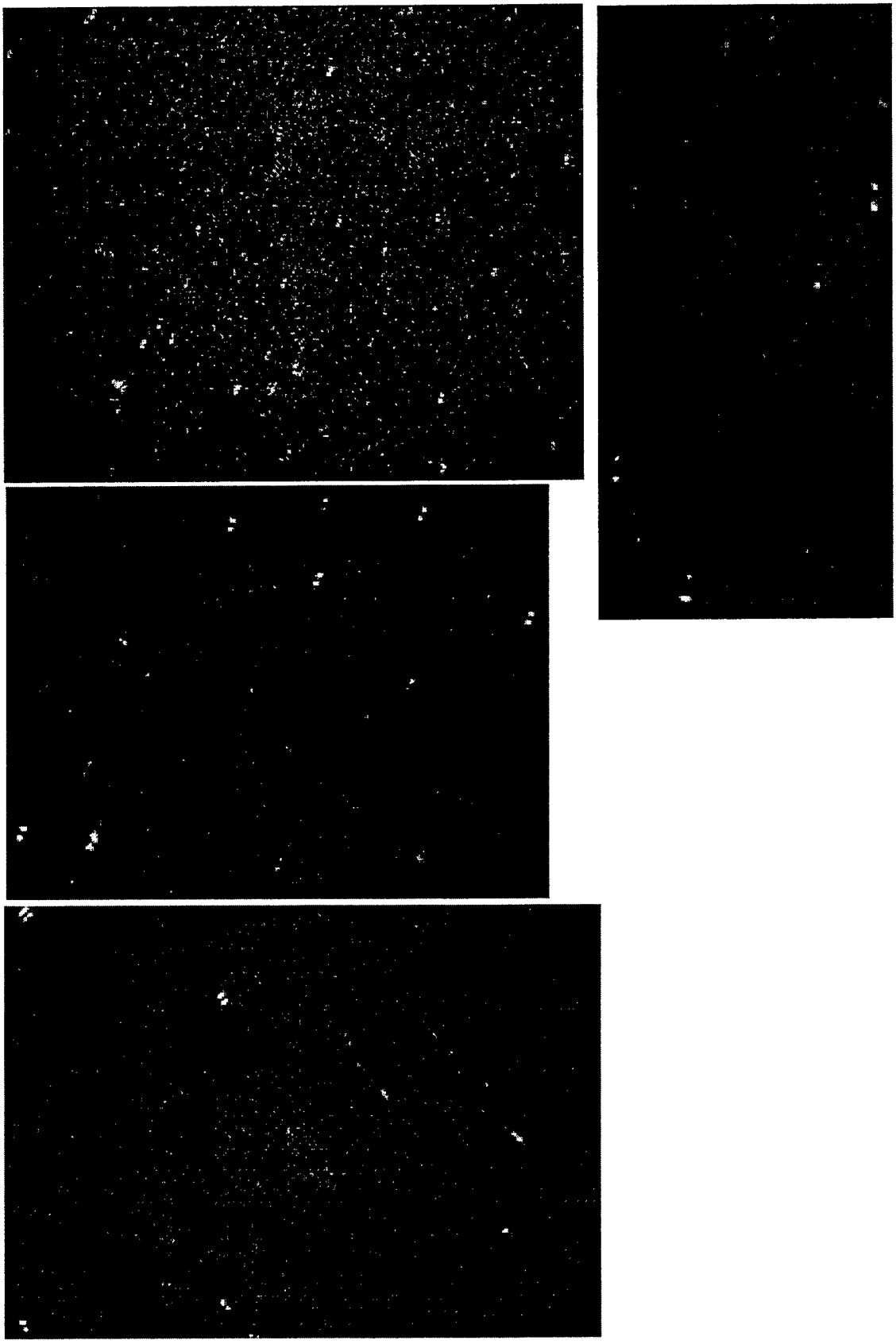


Figure 7. Particle pairs reconstructed from an IROV hologram of a shear layer

BAND GAP CALCULATIONS ON OLIGOMERS WITH AN ALL-CARBON BACKBONE

Douglas J. Miller, PhD
Associate Professor of Chemistry
Science & Mathematics Department

Cedarville College
PO Box 601
Cedarville, OH 45314-0601

Final Report for:
Summer Faculty Research Program
Wright Laboratory/Materials Directorate

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

and

Wright Laboratory/Materials Directorate

September 1996

BAND GAP CALCULATIONS ON OLIGOMERS WITH AN ALL-CARBON BACKBONE

Douglas J. Miller, PhD
Associate Professor of Chemistry
Science & Mathematics Department
Cedarville College

ABSTRACT

Attempts were made to calculate the band gap of various conjugated oligomers with an all carbon backbone. Computational software (Gaussian, Gamess, and HyperChem) was utilized in this work. Extrapolation of the results for the oligomers to the polymeric materials was made to predict candidates possessing desired conductive and nonlinear optical properties.

BAND GAP CALCULATIONS ON OLIGOMERS WITH AN ALL-CARBON BACKBONE

Douglas J. Miller, PhD

INTRODUCTION

Conjugated polymers are important for their applications in electronic technologies.[1] The band gap is a key material property for (semi)conductive and nonlinear optical properties, and much of the current focus is on designing new polymers having small band gaps.[2] Typical undoped conjugated polymers have band gaps in the 1.5 - 3.5 eV range, so polymers possessing values less than or equal to 1.0 eV represent important candidate materials.

We recently proposed[3] a scheme whereby low band gap materials could be designed. The underlying principle is that of symmetric ground states. Briefly, these are polymers in which two equivalent, isoenergetic, isostructural resonance forms contribute equally to the real structure. This present report extends this work to a system having an all carbon backbone, a negative charge, and a very rigid structure.

DISCUSSION AND RESULTS

Computational software (Gaussian[5], Gamess[6], and HyperChem) were utilized in attempts to predict the approximate band gap in potentially conductive polymers. Oligomeric species were studied with the expectation that trends for these species could be extrapolated to the polymeric materials. All of the studied species had a pattern of alternating double bonds. The smallest members of the two series of oligomers studied are shown in Figure 1 (the number after the SQ or PH represent the number of squarate groups in that oligomer). The SQ and PH series differ in their

termination groups. The SQ series was studied first and had structures end-capped by methylene groups. The PH series had a phenyl ring on each end. Each squarate group in either series would "naturally" have a -1 charge shared by the two oxygens of that squarate group. It was for this reason that all species possessing one minus charge per squarate were labeled as the "natural" forms. Both oxidized and reduced species of the oligomers were examined. Oxidized species might show the effect of p-doping. Figure 2 contrasts the bonding of the SQ2 molecule with a -2 charge to that of the neutral molecule. Note that in the neutral compound, the middle ring has lost its aromaticity.

Other differences between the SQ and PH series of oligomers may be noted. The SQ2 molecule in Figure 1 has an aromatic ring in the middle and two quinoidal rings on either side. In contrast, the PH2 molecule has a quinoidal ring in the middle. Each series is expanded to the next member by adding one squarate ring and one six-membered ring to each end, thus forming molecules SQ4, SQ6, SQ8, PH4, PH6, and PH8. All "natural" forms have alternating sequences of aromatic and quinoidal six-membered rings. As mentioned earlier, the SQ molecules are terminated by methylene groups attached to quinoidal six-membered rings, whereas PH molecules are terminated by aromatic phenyl rings. The middle ring in each series will alternate between an aromatic and quinoidal structure as you go from one oligomer to the next larger one (ie, from PH2 to PH4 to PH6 to PH8). The PH4 molecule with a -4 charge is shown in Figure 3 and its structure should be contrasted to that of the PH2 molecule with a -2 charge in Figure 1.

Both series offer a model for a conductive system with an all carbon backbone.

Quantum chemistry software was first used on the SQ series to find optimized geometries using semiemperical (AM1) parameters. Rough z-matrices were used as input for geometry optimizations in the Gaussian 92 program. The molecules were fully optimized within D_{2h} symmetry. Only the SQ2 and PH2 systems could be completely optimized by Gaussian 92. Other oligomers either had their geometries only partially optimized by Gaussian 92 or merely had their z-matrices checked by Gaussian 92. These approximate results were used as input for final geometry optimization in the Gamess program.

The electronic energies of the ground and excited states of the SQ series were then evaluated by semiemperical (AM1) and Hartree-Foch (with the 3-21 basis set) methods in Gamess. Two major difficulties arose at this point. Gamess would not perform the Hartree-Foch energy calculations on the SQ6 molecules (having six squarate groups) due to linear dependencies in the input basis set. In addition, the numbers coming out of the calculations already performed were not following any predictable or expected pattern (see Table 1). This latter result was found to be due to the tendency of the excited electrons to be unbound to the system. An undesirable collection of electron density in the methylene end-groups of the SQ series of molecules was also found, thereby possibly introducing an undesirable end-cap effect. The above factors lead us to alter our approach.

The PH series of oligomers was chosen to avoid the end-cap effect. Geometry optimizations were carried out as before. Each atom of the PH2

molecule with a -2 charge, PH2-2, had an explicit basis set defined for it in the Hartree-Foch calculations. The atoms of the squarate groups, which were found to possess large concentrations of electron density, were provided with a diffuseness (P function) term in the explicit basis set used (which was DH). Only the PH2-2 molecule has been studied so far, and it yielded a HOMO-LUMO gap of 1.46eV. More work on the PH series remains.

HyperChem (Release 3 for Windows) was used both to visualize geometries optimized by other software and to calculate the wavelengths for the HOMO-LUMO electronic transitions. These latter calculations were carried out to compensate for the inability of the earlier methods to produce usable results. The ZINDO/S method with singly excited configurations used in the configuration interaction (CI) calculations was utilized. A symmetric range of occupied and unoccupied orbitals, starting with three of each (3x3), was considered for each initial CI calculation. The range of orbitals was increased (to 4x4, 5x5, etc.) until at least two consecutive results were nearly identical. All of the geometries used were optimized geometries of the appropriate "natural" oligomer, even for oxidized or reduced species. Calculations were done for the "natural" systems and for the systems with more or fewer pairs of electrons (see Table 2 on pages 41-14, 41-15, and 41-16). Oscillator strengths are also shown in Table 2.

The HyperChem results, with nanometers translated into eVs, for the "natural" members of the SQ and PH series were plotted in Figure 4. The x-axis units are 1/n where n is the number of squarate pairs in the oligomer. The graph shows that the transition energy decreases with increasing length (smaller 1/n values) of the oligomer. The point representing the transition

energy for the PH8 molecule (x = 0.25) with a molecular charge of -8 appears to indicate that the curve for the PH series is heeling over at this point. Extrapolation of the PH curve to the polymeric point (large n values, x close to zero) is an encouraging sign that such polymers may be reasonably good conductors. It is expected that the extrapolated curves for both the SQ and PH series should have similar values for their y-intercepts (x = 0).

CONCLUSION

The calculations involving the software packages, Gaussian and Gamess, could not be completed owing to computational difficulties. Linear dependencies in the basis set used for the SQ oligomers precluded carrying out Hartree-Foch calculations on some of the SQ molecules. The values which could be obtained from such calculations proved to be unreliable due to the tendency of excited electrons to become unbound. This latter difficulty was compensated for, but insufficient time remained to complete a new set of calculations.

Efforts were shifted to the use of HyperChem for the computation of the first singlet transition for several molecules in the SQ and PH series. Results plotted for the "natural" members of these two series suggested that polymeric materials derived from such molecules might possess band gaps below 1.5 eV, or even below 1.0 eV. Further computational work might be able to firmly establish this point.

DJM:ljm

REFERENCES

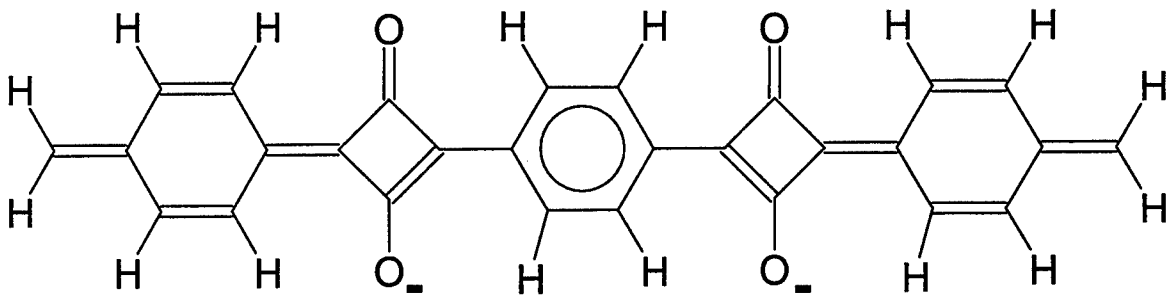
(1) Gorman, C. B.; Grubbs, R. H. In *Conjugated Polymers*; Bredas, J. L.; Silbey, R., Eds.; Kluwer Academic Publishers: Boston, **1991**; pp. 1-48.

(2) Kertesz, M.; Ashertehrani A. *Macromolecules* **1996**, *29*(3), 940-945.

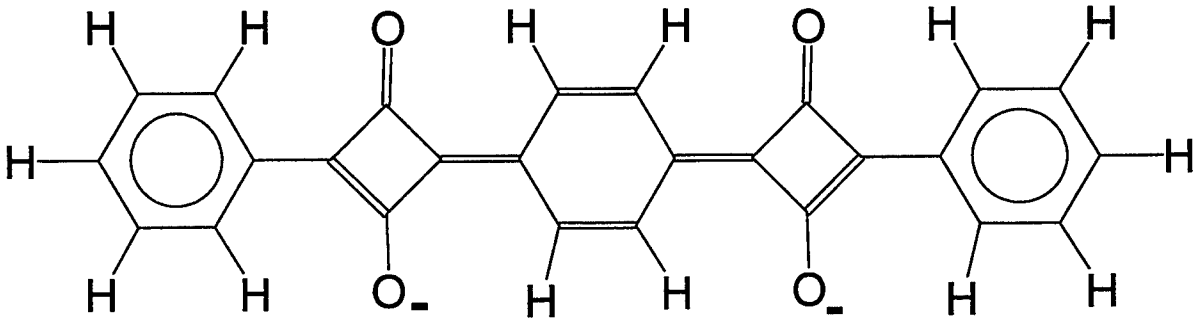
(3) Unpublished Results.

(4) *Gaussian 92*, Revision A, M. J. Frisch, G. W. Trucks, M. Head-Gordon, P. M. W. Gill, M. W. Wong, J. B. Foresman, B. G. Johnson, H. B. Schlegel, M. A. Robb, E. S. Replogle, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzalez, R. L. Martin, D. J. Fox, D. J. Defrees, J. Baker, J. J. P. Stewart, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1992.

(5) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347-1363.

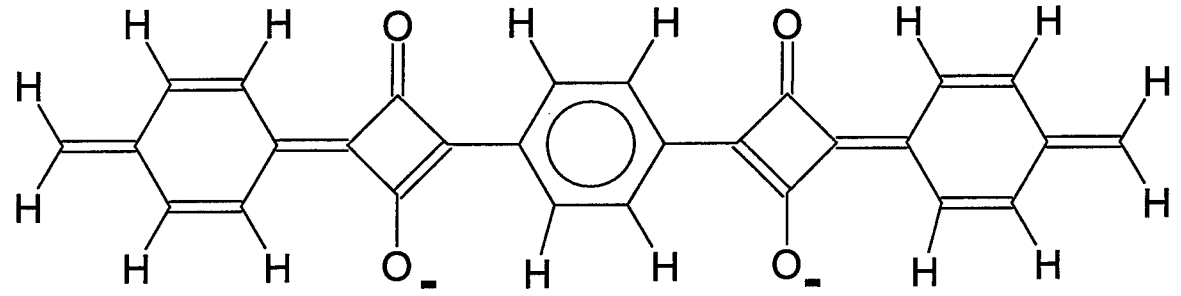


SQ2 Molecule, -2 Charge
("Natural" Form)

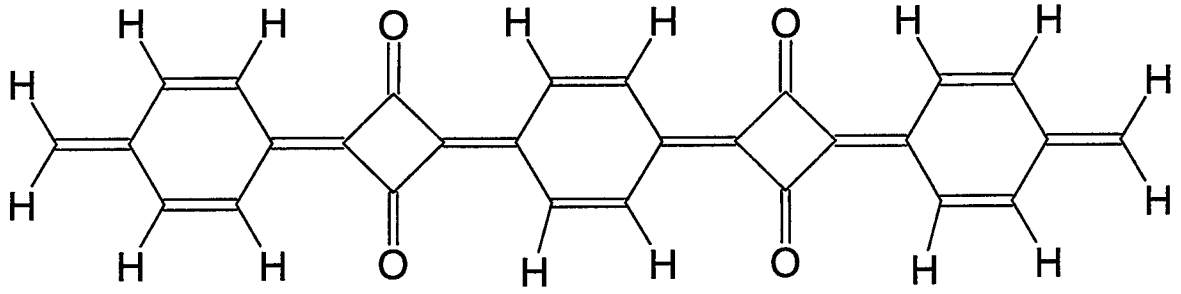


PH2 Molecule, -2 Charge
("Natural" Form)

Figure 1 (SQ and PH Series)

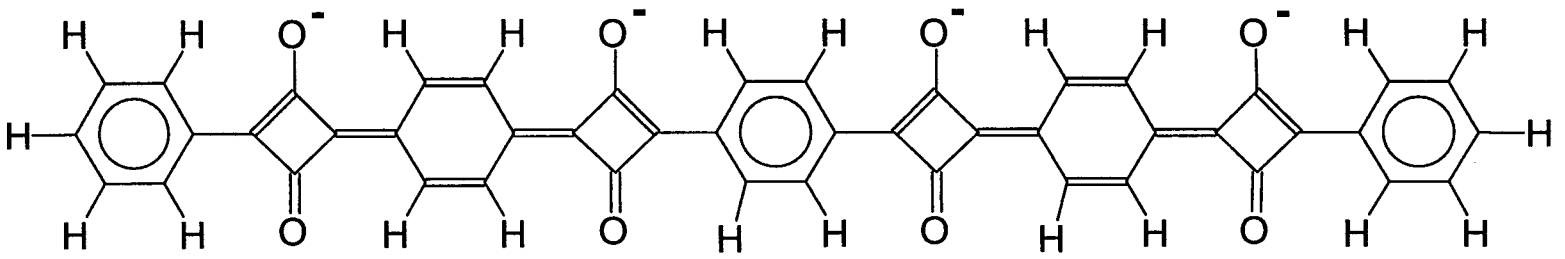


SQ2 Molecule, -2 Charge
("Natural" Form)

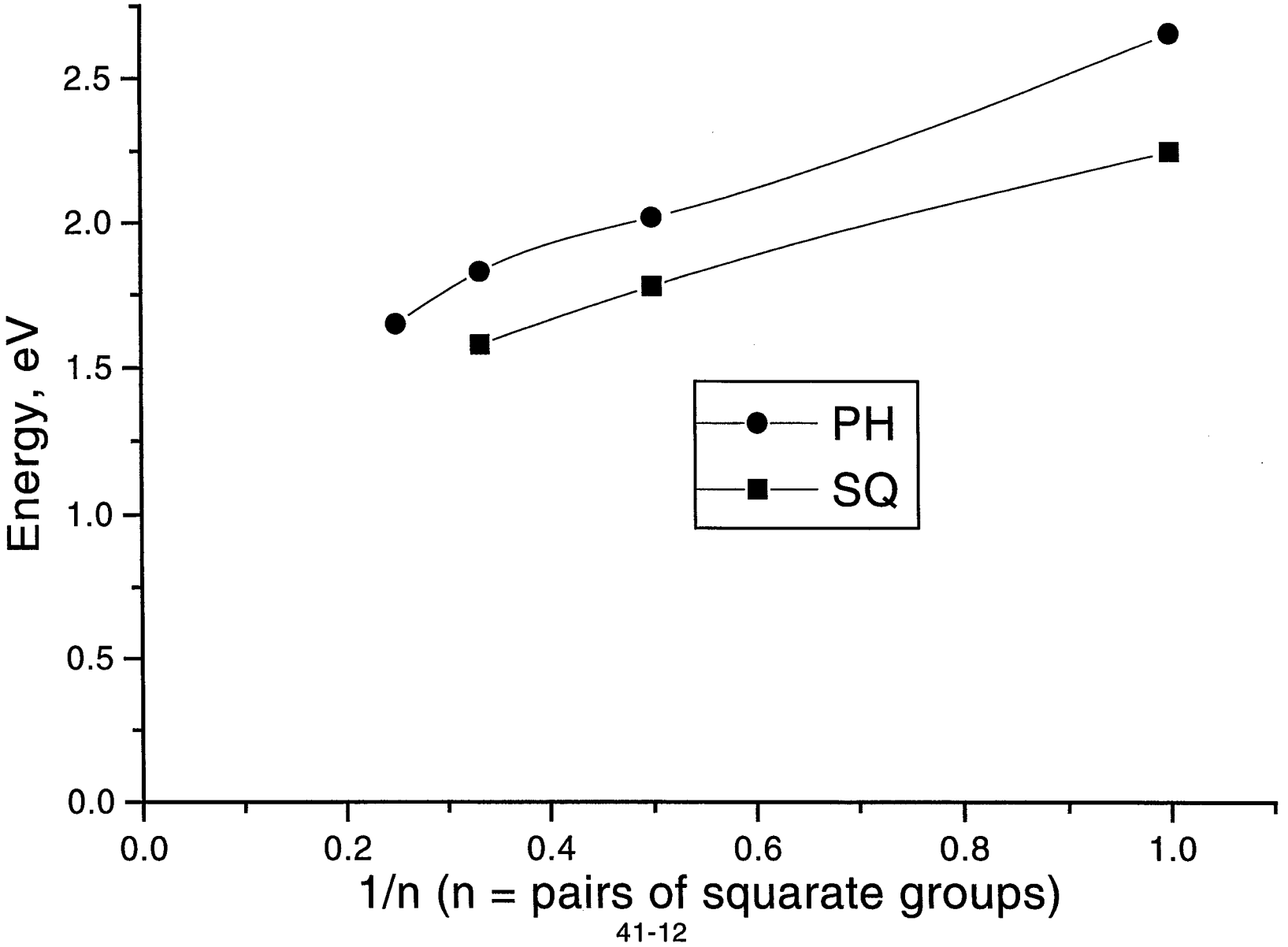


SQ2 Molecule, Neutral
(Oxidized)

Figure 2 (SQ2 System)



PH4 Molecule, -4 Charge
("Natural" Form)


Energy, eV
2.5
2.0
1.5
1.0
0.5
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1/n (n = pairs of squarate groups)
PH
SQ

41-12

Table 1

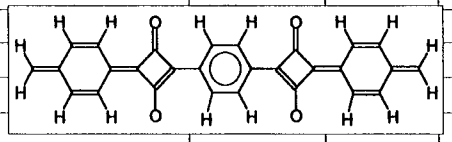
Transition Energies (eVs) for SQ Oligomers

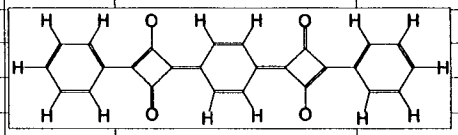
No. Squarates	Charge	Energy (AM1)	Energy (HF[a])
2	-2	1.6496	2.9769
	0	1.1420	1.3564
4	-4	2.2184	3.5112
	-2	1.8420	2.2049
	0	3.3176	3.1437
6	-6	2.0222	b
	-4	2.1675	b
	-2	0.8944	b
	0	3.2893	b

a - Hartree-Foch calculations using 3-21 basis set in Gamess
b - Linear dependencies in the basis set prevented calculation

Table 2

ZINDO/S Calculations of Singlet-Singlet Transitions in Phenyl-Squarate Oligomers





No. Squarates	Charge	CI	Singlet	Methylene-Term. Osc-Str	Singlet	H-terminated Osc-Str
2	0	3x3	616	4.72	690	3.27
		4x4	618	4.67	690	3.27
		5x5	618	4.67	690	3.27
		6x6	618	4.67	692	3.30
	-2	3x3	549	3.32	460	3.10
		4x4	551	3.21	461	3.15
		5x5	551	3.21	466	3.26
		6x6	551	3.21	466	3.26
	-4	3x3	772	6.02	588	2.79
		4x4	773	6.03	620	2.94
		5x5	774	5.99	620	2.94
		6x6	774	5.99	620	2.94
4	0	3x3	768	6.93	979	5.54
		4x4	777	7.04	999	5.38
		5x5	786	7.09	1003	5.39
		6x6	786	7.08	1001	5.37
		7x7			1001	5.37
	-2	3x3	1048	5.97	885	5.66
		4x4	1072	5.84	897	5.63
		5x5	1092	5.70	907	5.56
		6x6	1096	5.65	907	5.55
		7x7	1096	5.65	907	5.55

Table 2

		-4	3x3	677	5.30	591	4.82
			4x4	679	5.41	592	4.73
			5x5	679	5.38	593	4.69
			6x6	696	5.38	603	4.77
			7x7	696	5.38	615	4.80
			8x8			615	4.80
		-6	3x3	1167	11.68	951	7.92
			4x4	1169	11.70	959	7.94
			5x5	1162	12.20	1024	7.71
			6x6	1197	11.81	1070	7.29
			7x7	1197	11.81	1070	7.28
			8x8			1070	7.28
6		0	3x3	833	9.02	1172	7.66
			4x4	863	8.93	1224	7.3
			5x5	878	8.99	1251	7.12
			6x6	886	9.02	1271	6.95
			7x7	895	9.02	1272	6.94
			8x8	895	9.01	1274	6.94
		-2	3x3	1147	7.37	990	8.16
			4x4	1186	7.27	1027	8.14
			5x5	1228	7.10	1056	8.09
			6x6	1249	6.98	1068	8.01
			7x7	1262	6.92	1074	7.97
			8x8	1274	6.84	1078	7.96
			9x9	1274	6.84	1078	7.96
		-4	3x3	1150	7.32	1141	7.28
			4x4	1185	7.25	1203	7.15
			5x5	1214	7.14	1232	6.99
			6x6	1230	7.04	1255	6.91
			7x7	1231	7.02	1262	6.85
			8x8	1243	6.97	1262	6.85
			9x9	1249	6.91		
			10x	1249	6.91		

Table 2

		-6	3x3	724	6.68	647	6.24
			4x4	764	7.30	650	6.47
			5x5	769	7.44	650	6.43
			6x6	771	7.51	665	6.46
			7x7	784	7.50	676	6.54
			8x8	784	7.48	676	6.54
		-8	3x3	1056	23.9	1066	29.7
			4x4	1092	23.4	1067	29.8
			5x5	1137	23.3	1158	28.0
			6x6	1137	23.3	1230	26.4
			7x7	1142	23.7	1285	24.1
			8x8	1162	23.5	1316	22.6
			9x9	1163	23.5	1318	22.6
			10x				
	8	-8	3x3			650	6.08
			4x4			663	6.57
			5x5			672	6.94
			6x6			697	7.44
			7x7			710	7.65
			8x8			728	8.07
			9x9			730	8.09
			10x			733	8.11
			11x			742	8.12
			12x			750	8.08
			13x			750	8.08
			14x			750	8.08

HYDROGEN AND HELIUM ION IMPLANTATIONS FOR OBTAINING HIGH-RESISTANCE LAYERS IN N-TYPE 4H-SILICON CARBIDE

Ravi K. Nadella
Assistant Professor
Division of Engineering and Computer Science
Wilberforce University
Wilberforce, Oh 45384

Final Report for:
Summer Faculty Research Program
Wright Laboratory

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

and

Wright Laboratory

August 1996

HYDROGEN AND HELIUM ION IMPLANTATIONS FOR OBTAINING HIGH-RESISTANCE LAYERS IN N-TYPE 4H-SILICON CARBIDE

Ravi K. Nadella
Assistant Professor
Division of Engineering and Computer Science
Wilberforce University

Abstract

Light ions like hydrogen and helium were used to obtain high-resistance layers in n-type silicon carbide. Ion implantation was used to dope silicon carbide with light ions. Maximum resistivities of the order of 10^7 and 10^8 Ω-cm were observed for hydrogen and helium implantations, respectively. The high-resistance of the implanted layers starts to decrease after annealing at 600°C. Rutherford backscattering measurements indicate the relationship between high-resistance and implantation damage.

HYDROGEN AND HELIUM ION IMPLANTATIONS FOR OBTAINING HIGH-RESISTANCE LAYERS IN N-TYPE 4H-SILICON CARBIDE

Ravi K. Nadella

Introduction

The present emphasis on wide-bandgap semiconductor silicon carbide (SiC) arises primarily from the need for devices operating under extreme conditions [1]. The important polytypes of SiC are 4H and 6H with bandgaps of 3.26 and 3.03 eV, respectively [2]. SiC has high thermal conductivity, high saturation drift velocity, and high breakdown electric field [1,2]. These unique properties together with thermal stability, radiation tolerance, chemical inertness make SiC promising material for high-temperature, high-power, high-frequency, and harsh environments [1,2].

To make different electronic devices in SiC, n-, p-, and i-type layers are required. Both n- and p-type layers were obtained by growth and ion implantation [3,4]. Presently no technology exists for obtaining reliable high-resistance i-type layers in SiC. High-resistance i-type layers are required for applications like semi-insulating substrates for the fabrication of high quality electronic devices such as metal-semiconductor field-effect transistors. Carrier confinement layers needed for interdevice isolation also require these high resistance i-type layers.

Even though, there are few published reports on this topic, detailed investigations were not performed [5-7]. Due to the lack of such a technology, etching is used for device isolation. But, etching loses the planarity of the process and the yield will be low. Inorder to exploit the superior properties of SiC and make high performance devices, it is very important to establish the semi-insulating technology in SiC to obtain high-resistance i-type layers [8,9].

There are mainly two ways of obtaining high-resistance layers in any material, namely damage related and chemical doping related [10]. In the first case, deep levels are created by damage and results in high-resistance layers by annihilating the free carriers. Damage is usually created by ion implantation and the same implant species can be used for obtaining high-resistance layers in both n-type and p-type material. The second method of chemical doping requires appropriate species to introduce deep donor or deep acceptor levels to compensate p-type or n-type material, respectively.

In this work, we used the first method and performed hydrogen and helium implantations to obtain high-resistance layers in n-type 4H-SiC by implantation damage.

Experiment

Hydrogen and helium implantations were performed into n-type 4H-SiC with electron concentrations of approximately $2x10^{19}$ cm^{-3} and $2x10^{18}$ cm^{-3}, respectively. The implant energies are 350 keV and 65 keV, respectively, for an implant dose of $1x10^{14}$ cm^{-2}. The implantation was performed at room temperature. Samples were annealed in the temperature range 200-1400^{o} C for 10 min in Ar ambient to study the stability of the high resistance behavior. For measuring the resistivity of the implanted samples, nickel dots of 1 mm diameter were deposited on the implanted surface and two large area nickel contacts were deposited on the un-implanted surface of the samples. The nickel contacts were alloyed at 1000°C for 1 min in Ar ambient for the samples annealed at 1000°C or more. If the samples were annealed at lower than 1000°C, the alloying temperature was limited to the annealing temperature. Simple Current-voltage (I-V) measurements were performed using HP4145B parameter analyzer to obtain the resistance between the nickel dot on the implanted surface and one nickel contact on the bottom. The thickness of the implant profile was estimated using the He atom concentration depth profiles obtained by SRIM software [11] whih uses the transport equation. Secondary ion mass spectrometry (SIMS) measurements on the sample were unsuccessful due to the high background concentration of H and low sensitivity for He in our system. From the resistance and the thickness of the layer, the resistivity of the implanted layer (ρ) was calculated using the formula, $\rho=R*A/d$, where R is the resistance, A is the area of cross section of the nickel dot, and d is the thickness. Rutherford backscattering (RBS) measurements were performed to study the link between implantation damage and the resistivity in the samples.

Results and Discussion

Resistivity variation with annealing temperature is shown in Figure 1 for both H and He implants. For alloying temperatures of less than 1000°C, we were skeptical about the Ohmic nature of the contacts. But, most of the contacts were nearly Ohmic. The contact resistance was measured by performing I-V measurements between either the top or bottom contacts. Even though the contact reistance was high, it is still negligible compared to the resistance of the implanted layer. The resistance of the conducting substrate is only few Ohms and was neglected. As seen from Figure 1, maximum resistivities of the order of 10^{7} and 10^{8} were observed for H and He, respectively. Higher resistivity for He compared to H is due to the lower electron concentration in the substrate used for He implants ($2x10^{18}$ cm^{-3} versus $2x10^{19}$ cm^{-3}) and the higher lattice damage caused by He due to its heavier mass (4 versus 1). As seen from Figure 1, resistivity is almost constant up to 600°C and decreases for higher temperatures. But, even after 1000°C/10 min annealing, the resistivity is 2x106 Ω-cm for H implant and after 1400°C/10 min annealing, the resistivity is $1x10^{6}$ Ω-cm for He implant

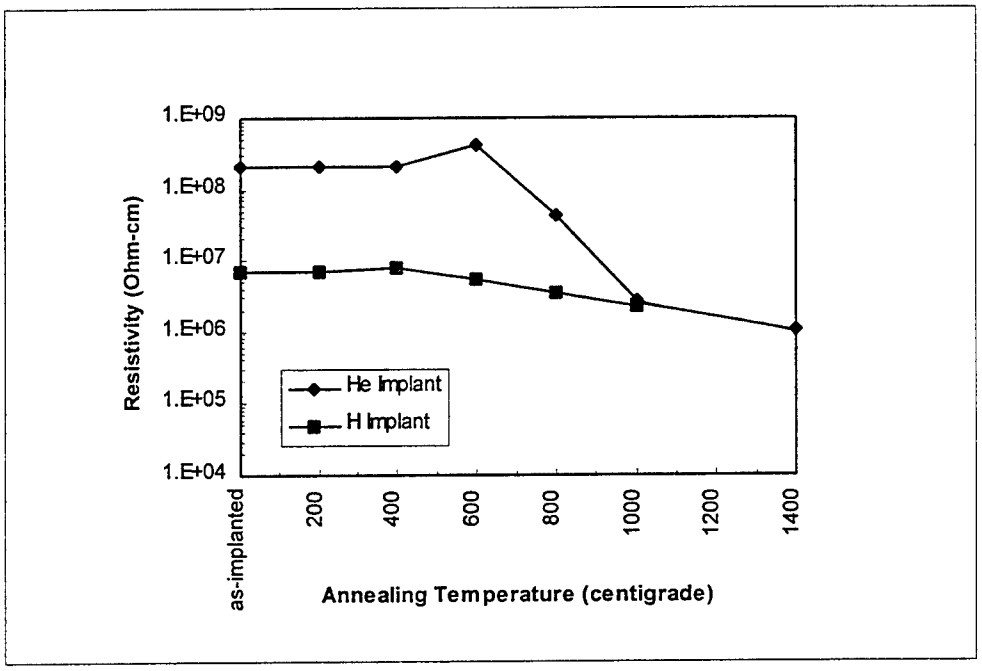


Figure 1. Resistivity variation with annealing temperature for 65 keV/10^{14} cm^{-2} He implants and 350 keV/10^{14} cm^{-2} H implants for 10 min. FA.

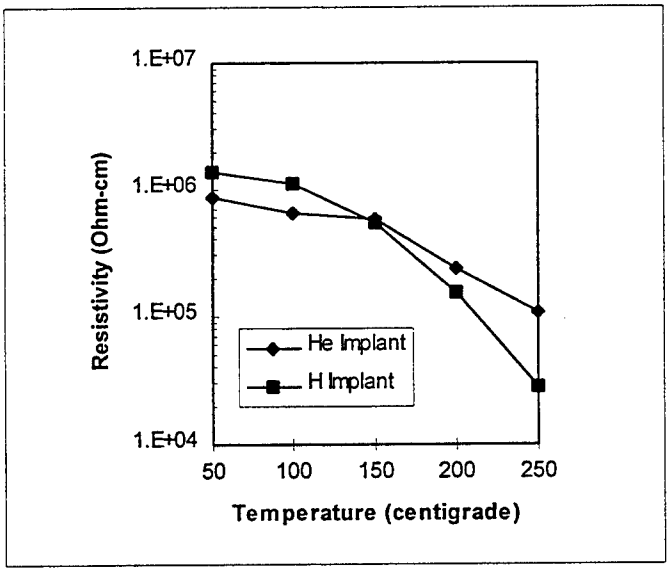


Figure 2. Resistivity variation with measurement temperature for 65 keV/10^{14} cm^{-2} He and 350 keV/10^{14} cm^{-2} H implants after 1400°C/10 min FA and 1000°C/10 min FA, respectively.

Since SiC is pursued for its high temperature applications, the high-resistance layers have to withstand high temperatures. To check this, I-V measurements were performed at temperatures up to 250°C using a heating stage. Figure 2 shows the results for both H and He implanted samples. Resistivities of the order of 10^4 and 10^5 Ω-cm were observed even at 250°C for H and He, respectively. These are high enough for many device applications. In fact these are higher than the best resistivities reported in SiC [5-7].

RBS results for H implants are shown in Figure 3. After annealing at 1000°C for 10 min., the damage is very close to the virgin limit. At some energies, the yield for annealed sample is lower than that for virgin. This may be due to the non-uniformities in the substrate quality. Even though the spectra for virgin and annealed samples are very close, still there may be considerable amount of damage in the annealed sample. RBS results can be directly correlated to the resistivity variation with temperature (Figure 2) where resistivity decreases with increasing annealing temperature. Hence, it can be interpreted that the resistivity obtained in this study is caused by implantation damage. The RBS data for He implants is also similar.

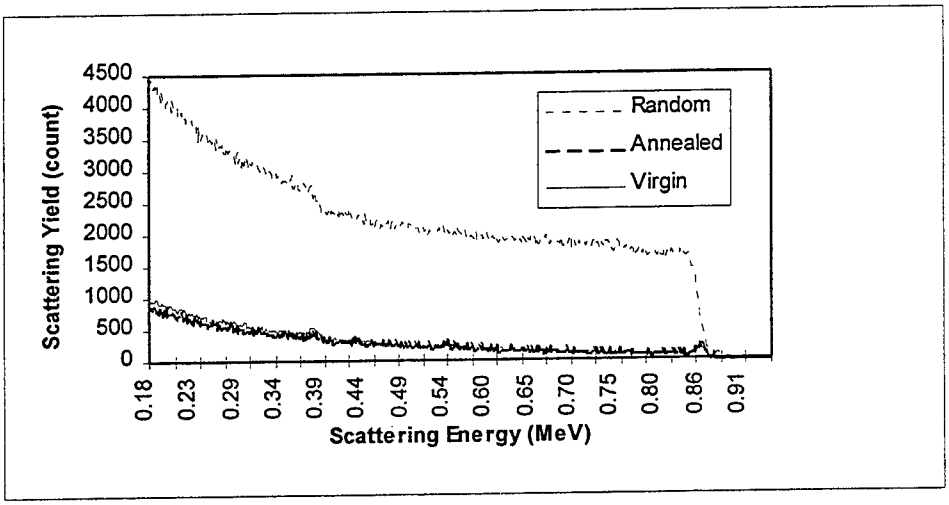


Figure 3. RBS data on 350 keV/10^{14} cm^{-2} implanted SiC. Scattering specra for random, virgin, and 1000°C/10 min FA are shown.

The results of this study can be used to obtain high-resistance layers where lattice damage is not of concern. Where lattice damage needs to be low, transition metals have to be used for high-resistance layers. Once these technologies are established, silicon carbide on insulator (SCOI) like silicon on insulator (SOI) is possible. This will open up for the fabrication of a new class of devices to be made in SiC.

Conclusions

High-resistance layers were obtained in n-type SiC with resistivities of 10^8 Ω-cm. Resistivity is almost constant for annealing temperatures up to 600°C. Even at 250°C measurement temperature resistivities of the order of 105 Ω-cm were observed. RBS data links the high-resistance behavior to the implantation damage. When high-resistance layers with low damage and high operating temperatures are required, transition metals have to be used.

Acknowledgments

The author would like to acknowledge the help of M.A. Capano, J. Landis, R. Bertke, J. Holinger, C. Cerny, and R. Niedhart of Wright Patterson Air Force Laboratory, J. Gardner of George Mason University, D.S. Simons and P.H. Chi of National Institute of Standards and Technology, O.W. Holland of Oak Ridge National Laboratory, and M. Stan of Case Western Reserve University during the course of this work. This work was supported by the Air Force Office of Scientific Research.

References

1. H. Morkoc, S. Strite, G.S. Gao, M.E. Lin, B. Sverdlov, and M. Burn, "Large-band-gap SiC, III-V nitride, and II-VI ZnSe-based semiconductor device technologies", J. Appl. Phys. 76, 1363(1994).
2. Technical notes of Cree Research Inc., Durham, NC 27713.
3. J. Gardner, M.V. Rao, O.W. Holland, G. Kelner, D.S. Simons, P.H. Chi, J. Andrews, J. Kretchmer, and M. Ghezzo, "Elevated temperature nitrogen implants in 6H-SiC", J. Electron. Mater. 25, 1(1996)
4. S. Ahmed, C.J. Barbero, and T.W. Sigmon, "Activation of ion implanted dopants in alpha-SiC", Appl. Phys. Lett. 66, 712(1995).
5. H.M. Hogwood, R.C. Glass, G. Augustine, R.H. Hopkins, J. Jenny, M. Skowronski, W.C. Mitchel, and M. Roth, Appl. Phys. Lett. 66, 1364(1995).
6. J.R. Jenny, M. Skowronski, W.C. Mitchel, H.M. Hogwood, R.C. Glass, G. Augustine, and R.H. Hopkins, J. Appl. Phys. 78, 3839(1995).
7. T. Kimoto, T. Nakajima, H. Matsunami, T. Nakata, and M. Inoue, "Formation of semi-insulating 6H-SiC layers by vanadium ion implantation", Appl. Phys. Lett. 69, 1113(1996).

8. D. Caputo, G. Cesare, F. Irrera, and F. Palma, "Solar-blind UV photodetectors for large area applications", IEEE Trans. On Electron. Dev. 43, 1351(1996).
9. S. Sriram, G. Augustine, A.A. Burk, R.C. Glass, H.M. Hobgood, P.A. Orphanos, L.B. Rowland, T.J. Smith, C.D. Brandt, M.C. Driver, and R.H. Hopkins, ", IEEE Electron Dev. Lett. 17, 369(1996).
10. R.K. Nadella, J. vellanki, M.V. rao, and O.W. Holland, "MeV B compensation implants into GaAs and InP", J. of Appl. Phys. 72, 2179(1992).
11. J.F. Ziegler and J.P. Biersack, "Stopping and range of ions in matter", software to calculate the dopant profiles in matter (1996).

Least Squares Marquardt Calibration of Dielectric Resonator Measurements

Krishna Naishadham

Associate Professor

Department of Electrical Engineering

Wright State University

Dayton, OH 45435

Final Report for:

Summer Faculty Research Program

Wright Laboratory

Sponsored by:

Air Force Office of Scientific Research

Bolling Air Force Base, DC

and

Materials Directorate

Wright Laboratory, OH

September 1996

Least Squares Marquardt Calibration of Dielectric Resonator Measurements

Krishna Naishadham

Associate Professor

Department of Electrical Engineering

Wright State University

Dayton, OH 45435

Abstract

Dielectric resonators (DRs), consisting of high temperature superconducting (HTS) thin-films for the end caps, are useful as microwave characterization tools, and as components in microwave systems (*e.g.*, sharp-skirt miniature filters in cellular base stations). In HTS DR measurements, because of the extremely high Q's possible (of the order of 10^6), the measured parameters are very sensitive to the background "noise" contributed by the coupling mechanism, case modes, radiation, etc. Therefore, it becomes important to properly calibrate all the background information in order to accurately measure the unloaded Q factor of the resonator. The research reported here pertains to the application of a non-linear curve-fitting procedure, the least squares Marquardt (LSM) algorithm, to accurately filter out the background noise and extract the unloaded Q of a DR. It has been observed that the background noise manifests out of the resonant band as a sinusoidal envelope. Therefore, an appropriate model of the measured data that we employ is a linear fractional transformation corresponding to mapping of a pure Lorentzian in the complex plane, multiplied by a sum of complex exponentials pertaining to the standing wave modes of the loop-coupling mechanism which corrupt the resonance curve. We have used this composite non-linear transformation to model raw data from a DR with copper end-plates. Preliminary analysis shows that the LSM fit is quite satisfactory even in the case of highly corrupted data. In future, we anticipate to vaildate this algorithm with measured data on HTS DRs.

Least Squares Marquardt Calibration of Dielectric Resonator Measurements

Krishna Naishadham

1 Introduction

The discovery of high temperature superconductivity in LaBaCuO at 30K by Bednorz and Müller (1986) [1], and in YBaCuO (YBCO) at temperatures above 90K by Chu and several others (1987) [2], continues to impact on the design of microwave systems. Because of extremely small losses (or high Q-factor), low noise, low power consumption, potential for circuit miniaturization, and uniform electrical behavior over a wide temperature range, high temperature superconductor (HTS) materials are becoming increasingly useful in aerospace industry, where size, weight and performance need to be optimized. Several designs of passive HTS microwave circuits, such as ultra low-loss transmission lines, sharp skirt microwave filters, high-gain antenna arrays, etc., have been reported [3]. In addition, HTS materials exhibit non-linear field effects at the macroscopic level (*e.g.*, Josephson tunneling effect), which make possible a number of active devices, such as field effect transistor (FET) and heterojunction bipolar transistor (HBT), operating with improved performance over their room-temperature normal conductor counterparts [4].

Most of the microwave applications of HTS materials employ thin-film technology, in contrast to bulk materials. Microwave characterization of HTS thin films is essential to the development of these applications. The most important thin-film HTS parameter for such characterization is the surface resistance, R_s, which determines the dissipation in the film. This collaborative research with WL/MLPO addresses how the surface resistance can be measured accurately with a microwave network analyzer using miniaturized dielectric resonators (DRs) comprising of 1 cm^2 HTS films [5]. The author's objective of the research is to develop a robust and accurate model for extracting the unloaded Q of the DR from narrow-band measurements of scattering (S) parameters of loop-coupled resonators. Therefore, the details of measurement procedure *per se* are not dealt with in this report. The interested reader is referred to [6] - [8] for these details. The reflection and transmission measurements on the films were carried out by Dr. Eric Moser at WL/MLPO at several frequency bands in the range 19 GHz to 40 GHz, and at several temperatures.

In HTS DR measurements, because of the extremely high Q's possible (of the order of 10^6), the measured parameters are very sensitive to the background "noise" contributed by the

coupling mechanism, case modes, radiation, etc. Therefore, it becomes important to properly calibrate all the background information in order to accurately measure the unloaded Q factor of the resonator. The research reported here pertains to the application of a non-linear curve-fitting procedure, the least squares Marquardt (LSM) algorithm, to accurately filter out the background noise and extract the unloaded Q of a DR. It has been observed that the background noise manifests out of the resonant band as a sinusoidal envelope. Therefore, an appropriate model of the measured data is a linear fractional transformation corresponding to mapping of a pure Lorentzian in the complex plane, multiplied by a sum of complex exponentials pertaining to the standing wave modes of the loop-coupling mechanism which corrupt the resonance curve. We have used this composite non-linear transformation to model raw data from a DR with copper end-plates. Preliminary analysis shows that the LSM fit is quite satisfactory even in the case of highly corrupted data.

The measurement procedure is still evolving, and technical difficulties involved in the accurate microwave characterization of DRs comprising of small-area HTS thin films precluded the validation of the LSM algorithm on HTS samples. These difficulties are anticipated to be mitigated by the recent acquisition of smaller sapphire pucks and coaxial cables for better confinement of energy within the DR. In future, we anticipate to validate this algorithm with measured data on the improved HTS DRs.

The report is organized as follows. In the next section, a brief survey of two existing methods to extract surface resistance from dielectric resonator measurements of S-parameters is presented, and the limitations of the two methods are discussed. In the third section, we describe an equivalent circuit representation of the dielectric resonator cavity and coupling mechanism, and discuss how the reflection and transmission coefficients of the equivalent circuit circumscribe circles in the complex plane as a function of frequency. These circles, known as Q-circles [5], [9], characterize dielectric resonators in a narrow frequency range around the resonant frequency. Subsequently, we introduce the LSM curve-fitting method to model the measured data and extract the unloaded Q of the DR. The algorithm's utility and robustness in the face of measurement-corrupting noise are demonstrated by its application to a DR with copper end-plates. The report concludes with a summary and discussion of the results. Some details on the follow-up effort to meet the HTS microwave characterization needs of WL/MLPO are also discussed.

2 Dielectric Resonator Measurements

Cavities containing dielectric resonators are very useful in measuring the surface resistance of HTS thin films. Fig. 1 shows a sapphire dielectric resonator with HTS end caps. Energy is coupled into and out of the cavity through two coupling loops. By proper design and placement of these loops, one can ensure that only the dominant TE_{011} mode is excited within the resonator. Because the fields are well-trapped in the dielectric and within a small cylindrical region outside the sapphire, the losses in the resonator system can be minimized, with the result that extremely large Q's (of the order of a million) can be measured.

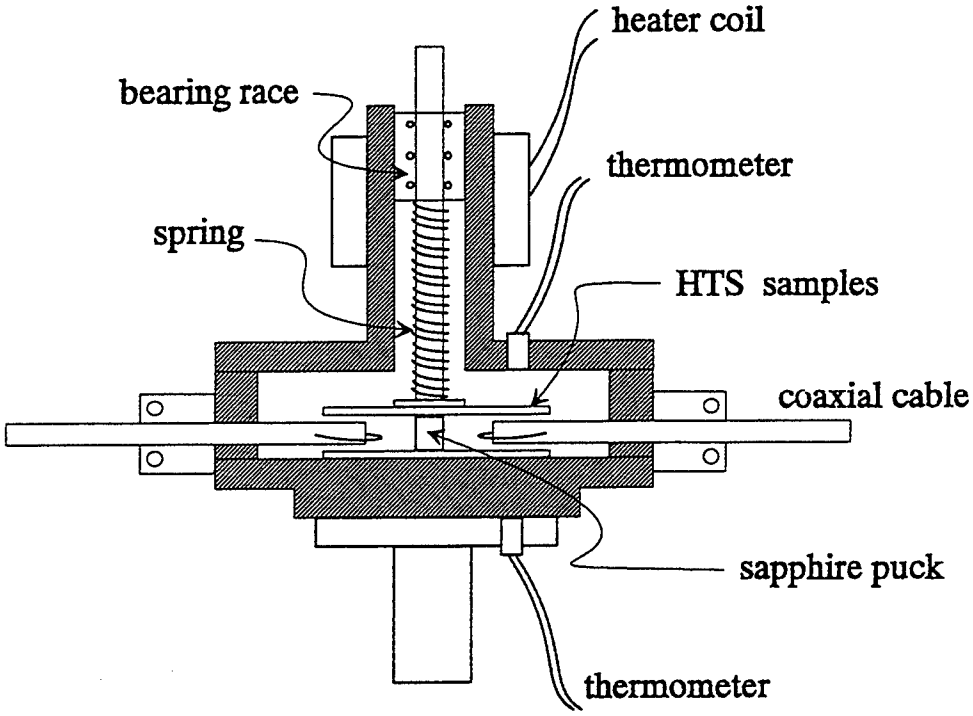


Figure 1: Cross-sectional view of the HTS-sapphire-HTS dielectric resonator fixture.

In order to measure the surface resistance of YBCO HTS films over a wide frequency range, six cylindrical saphire pucks of different diameters were employed. The HTS films were placed non-destructively on either end of the saphire puck, and the resonant frequencies as well as the insertion loss around the resonant frequency were measured by monitoring two-port S-parameters of the DR. The loops were made using coaxial cables, and connected to the 50 Ω test ports of HP8722 network analyzer for automated measurement.

The fields trapped within the dielectric (Fig. 1) are oscillatory and described by Bessel functions of the first kind, $J_n(x)$. The evanescent fields along the radial direction are specified by (decaying) Bessel functions of the second kind, $K_n(x)$. These fields and their behavior are analyzed in [5]. By imposing boundary conditions on the tangential fields at the dielectric

interface $\rho = a$, one obtains the transcendental equation

$$\frac{J_1(\xi_1 a)K_0(\xi_2 a)}{\xi_1 a} + \frac{K_1(\xi_2 a)J_0(\xi_1 a)}{\xi_2 a} = 0 \tag{1}$$

where ξ_1 and ξ_2 are radial wavenumbers in the dielectric ($\rho \leq a$) and air ($\rho \geq a$), respectively. In order to solve eq. (1) numerically for the resonant frequency, we provide initial guesses of the frequency and $\xi_1 a$ and calculate

$$\xi_2 a = \pi\sqrt{(a/L)^2 - (2a/\lambda)^2} \tag{2}$$

where λ is the operating wavelength. The resonant frequencies for six different puck diameters were computed from (1) using Mathcad, and are listed in Table 1. The dielectric constant of sapphire is assumed as $\epsilon_r = 9.3$. These resonant frequencies are in good agreement with measured values.

Table 1. Computed Resonant Frequencies of Sapphire Dielectric Resonators.

Radius a	Length L	Res. Freq. (GHz)
0.09"	0.137"	37.276
0.095"	0.137"	35.888
0.1"	0.137"	34.628
0.12"	0.137"	30.607
0.15"	0.137"	26.587
0.25"	0.137"	20.364

We have also investigated the evanescent field decay along the radial direction external to the dielectric puck, to provide insight into determining a puck size that would ensure a small perturbation of the resonant mode of the dielectric resonator [11]. These studies resulted in changing the puck geometry and the cable size to minimize radiation leakage and reduce the measurement noise. At the time of writing this report, the measurements on these new DRs were not available.

Two methods have been in use to extract the unloaded Q-factor of the dielectric resonator, namely, Ginzton method [12] and Kobayashi's method [13]. Both of these methods are applicable to the processing of S-parameters measured by the microwave network analyzer, and will be briefly discussed next. The limitations of these two methods will also be presented.

2.1 Ginzton Method

For a resonant cavity, the magnitude of the insertion loss (S_{21} expressed in dB) follows the peaked behavior shown in Fig. 2. Ginzton's method [12] entails the observation of the resonant frequency, f_L, and Δf, the spread between half-power (3 dB) points, to determine the loaded Q-factor, Q_L, as

$$Q_L = \frac{f_L}{\Delta f}. \tag{3}$$

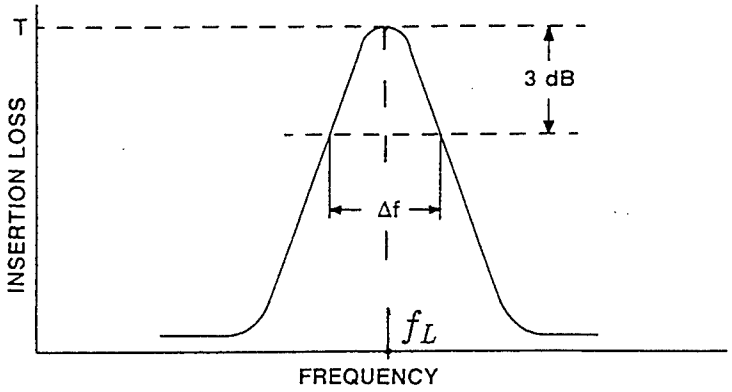


Figure 2: Resonant curve measurement of the loaded Q-factor of a dielectric resonator.

Ginzton's method employs measured data from only two frequencies, and thus, suffers from the following limitations. First, the unloaded Q-factor cannot be calculated because the magnitude response lacks information on the coupling coefficients, which determine the proportion of source power actually coupled into the resonator at each port. Second, if the data is either asymmetric around the peak or corrupted by measurement noise, an extraction procedure based only on magnitudes may yield very unreliable results. The phase of the measured S-parameters also becomes important in these situations.

2.2 Kobayashi Method

Kobayashi method [13] also employs an HTS dielectric resonator operating in the TE_{011} mode to determine the surface resistance of HTS films. The extraction procedure in Kobayashi's method improves upon the Ginzton method by providing the coupling coefficients, from which the unloaded Q-factor may be determined. Essentially, the loaded Q-factor is still computed from the resonant peak and the two 3 dB points, as in Ginzton's method (see (3)). Kobayashi, however, assumes that the input and output coupling coefficients are equal, and determines the

unloaded Q-factor from the insertion loss, T, at the resonant frequency f_L (see Fig. 2):

$$Q_0 = \frac{Q_L}{1-T}, \quad T = |S_{21}| = |S_{12}| = \frac{2\kappa_c}{2\kappa_c + 1} \tag{4}$$

where κ_c is the coupling coefficient at either port. We have found that Kobayashi's method requires moderate coupling for accurate prediction of the unloaded Q. It is difficult to ensure that the loops are always correctly positioned for equal coupling, especially with the small resonator fixtures that we employ at higher frequencies. Since Kobayashi's method is also based on magnitude measurements, it suffers from the same limitations as Ginzton's method, namely that the Q can be adversely affected by a few errant or inconsistent points in the sweep.

3 Equivalent Circuit Modeling

A resonator, in principle, has many modes with different resonant frequencies. However, if attention is focused on the dominant mode, which is the only one typically excited, the dielectric resonator can be conveniently represented by a parallel tuned circuit [9]. Thus, microwave circuit theory can be employed to formulate a robust extraction alogorithm for the determination of the unloaded Q-factor. Unlike Ginzton's and Kobayashi's methods, such an algorithm would utilize both magnitude and phase of the measured S-parameters.

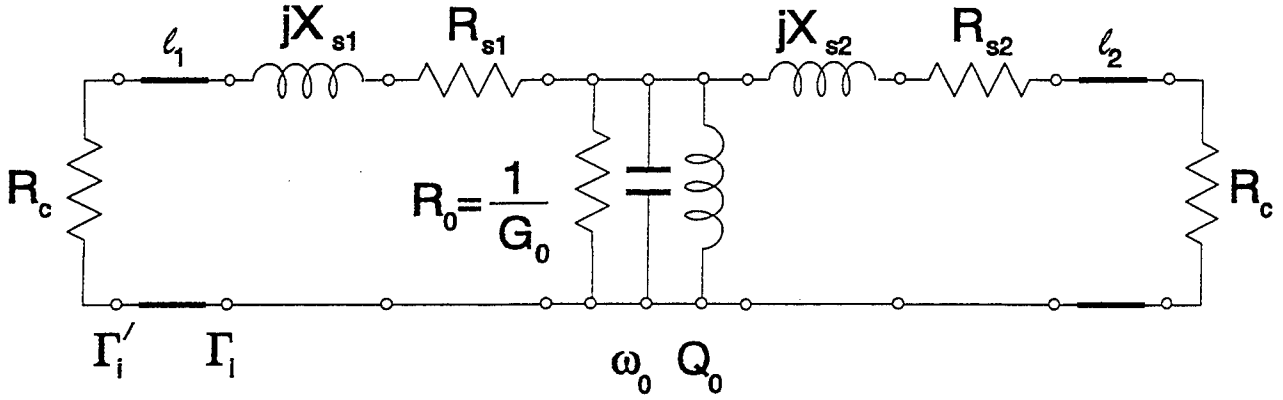


Figure 3: Equivalent circuit of the dielectric resonator configuration, including the input and output coupling loops.

An equivalent circuit of the dielectric resonator configuration of Fig. 1, including the coupling loops, is shown in Fig. 3. The resonator is completely specified by the resonant frequency ω_0, the unloaded Q factor Q_0, and the conductance G_0 (or the resistance R_0). The input and output coupling loops are each modeled by a series resistance R_s and reactance X_s. The series resistance accounts for the power dissipated in the coupling loop. The series reactance includes

the reactance of the loop, and also encompasses the influence of all higher-order resonant modes with distant resonant frequencies. This influence is usually negligible. Therefore, the equivalent circuit is valid only near the first (fundamental) resonance. The analyzer is connected to the loops by means of two transmission lines with characteristic impedance R_c. For modeling purposes, these lines are assumed to have lengths ℓ_1 and ℓ_2, respectively, at the input and output ends. The attenuation and phase shift introduced by these cables will be considered in Sec. 4. The unloaded admittance of the resonator is calculated as

$$Y_0 = \frac{1}{R_0}\left[1 + jQ_0 2\frac{\omega - \omega_0}{\omega_0}\right] \tag{5}$$

where ω is the operating frequency and ω_0 is the unloaded resonant frequency. This *unloaded* admittance does not consider the external loading of the coupling loops and the connecting transmission lines. The loaded admittance of the resonator is given by

$$Y_L = \frac{1}{R_0}(1 + \kappa_1 + \kappa_2)\left[1 + jQ_0 2\frac{\omega - \omega_L}{\omega_0}\right] \tag{6}$$

with the coupling coefficients due to external loading calculated as

$$\kappa_1 = \frac{(R_c + R_{s1})R_0}{(R_c + R_{s1})^2 + X_{s1}^2}, \qquad \kappa_2 = \frac{(R_c + R_{s2})R_0}{(R_c + R_{s2})^2 + X_{s2}^2}. \tag{7}$$

Here, ω_L is the loaded resonant frequency. It is observed from (7) that each coupling coefficient may be written as

$$\kappa_k = \kappa_k^l + \kappa_k^c, \quad k = 1, 2, \tag{8}$$

$$\kappa_k^l = \frac{R_c R_0}{(R_c + R_{sk})^2 + X_{sk}^2}, \qquad \kappa_k^c = \frac{R_{sk} R_0}{(R_c + R_{sk})^2 + X_{sk}^2}. \tag{9}$$

Notice that superscript l denotes coupling associated with the transmission line, whereas superscript c denotes that caused by the loop. Physically, each coupling coefficient equals the ratio of power dissipated in the external component to power dissipated in the resonator. Using standard circuit theory [9], the input impedance at each port can be calculated as

$$Z_1 = Z_{e1} + \frac{1}{Y_o + Y_{e2}}, \qquad Z_2 = Z_{e2} + \frac{1}{Y_o + Y_{e1}} \tag{10}$$

$$Z_{ek} = \frac{1}{Y_{ek}} = R_c + R_{sk} + X_{sk}, \quad k = 1, 2. \tag{11}$$

The port reflection coefficients are then given by

$$S_{kk} \equiv \Gamma_k = \Gamma_{dk} + \frac{2\kappa_k^l}{1 + \kappa_1 + \kappa_2}\frac{e^{j\gamma_k}}{1 + jQ_L 2\frac{\omega - \omega_L}{\omega_0}} \tag{12}$$

with the unloaded and loaded Q factors related by

$$Q_0 = Q_L(1 + \kappa_1 + \kappa_2). \tag{13}$$

In the limit as the resonator is detuned to an extremum on either side of ω_L, it is evident from (12) that the reflection coefficient approaches a value Γ_{dk} given by

$$\Gamma_{dk} = \frac{R_{sk} + jX_{sk} - R_c}{R_{sk} + jX_{sk} + R_c}. \tag{14}$$

The transmission coefficient can be derived in a similar manner, and is given by

$$S_{21} = S_{12} = \frac{2\sqrt{\kappa_1^l \kappa_2^l}}{1 + \kappa_1 + \kappa_2} \frac{e^{-j\phi}}{1 + jQ_L 2\frac{\omega - \omega_L}{\omega_0}}. \tag{15}$$

The phase angles γ_k and ϕ are functions of loop parameters R_s and X_s:

$$\gamma_k = 2 \arctan \frac{X_{ks}}{R_c + R_{ks}}, \quad \phi = \arctan \frac{X_{1s}}{R_c + R_{1s}} + \arctan \frac{X_{2s}}{R_c + R_{2s}}. \tag{16}$$

3.1 Q Circles

As the frequency deviates from ω_0, the reflection and transmission coefficients describe circles in the complex plane, known as *Q circles* [9]. The unloaded Q-factor can be accurately computed from the center and diameter of the Q-circle. As an example, Fig. 4 shows the Q circle for S_{11}, plotted from (12), for the case:

$$\frac{R_o}{R_c} = 2, \quad Q_0 = 1000,$$

$$\frac{R_{s1}}{R_c} = 0.2, \quad \frac{R_{s2}}{R_c} = 0.4, \quad \frac{X_{s1}}{R_c} = 0.5, \quad \frac{X_{s2}}{R_c} = 1.5,$$

$$\beta_0 \ell_1 = 36 \text{ deg.}, \quad \beta_0 \ell_2 = 40 \text{ deg.}, \quad f_0 = 1 \text{ GHz},$$

where β_0 is the free space wavenumber at f_0. The circle is obtained by plotting S_{11} over a frequency band of $\pm f_0(3/Q_0)$ around the unloaded resonant frequency. For lossy coupling loops, the energy coupled into the resonator is reduced by the dissipation in the loops, with the result that the Q-circle is tangential to a circle, known as the *coupling loss circle*, at the detuned point. The loss circle is shown by the dashed curve in Fig. 4. Similar Q-circles and loss circles can be drawn for S_{22} and S_{12}. The reader is referred to Appendix for further details on extraction of the unloaded Q from the geometrical attributes of these circles.

The Q-circles are usually not smooth for measured data because of extraneous noise and other limitations of the measurement system. These imperfections cannot be calibrated with

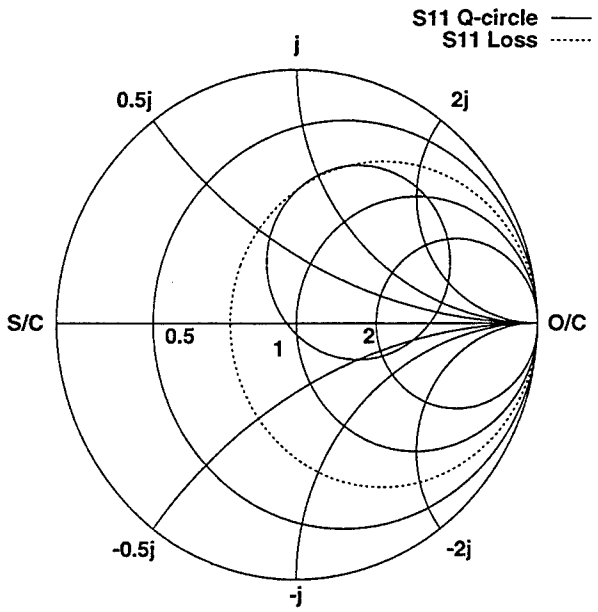


Figure 4: Q-circle for the simulated reflection coefficient data.

hardware. However, because of the physical reasoning that the measured S-parameters of loop-coupled resonators must follow the Q-circles as a function of frequency [9], a curve-fitting procedure may be used to model the measured data and filter out the imperfections.

4 Least Squares Marquardt Curve Fitting Procedure

The LSM method for parameter extraction of the two-port loop-coupled DR is based on a similar method for the analysis of one-port measurements [9] and multi-mode resonators [14]. Our improved technique involves enhancing the expressions for the reflection and transmission coefficients, given by eqs. (12) and (15), respectively, by considering the complete equivalent circuit including the coupling mechanism, and using least squares minimization to fit the full sweep through the resonance. The fit functions to these coefficients are of the form

$$w_i = \frac{a_1 t_i + a_2}{1 + a_3 t_i} \sum_{j=1}^{p} A_j e^{\gamma_j t_i} \tag{17}$$

with the normalized frequency

$$t_i = 2\frac{f_i - f_L}{f_0}. \tag{18}$$

Eq. (17) may be viewed as a non-linear fractional transformation mapping the normalized frequency variable, t_i, to the space spanned by w_i. The complex transformation constants a_1, a_2, a_3, the amplitudes A_j and the propagation constants γ_j of the p transmission line modes existing on the connecting cables because of impedance discontinuity at the loop interface, are to be determined from the set of i measurements, f_i, w_i, $i = 1, 2, \cdots N$, where w_i denotes

theoretical approximation to the measured parameter at the frequency f_i. The functional dependence of these constants on physical parameters of the resonator may be determined by comparing the right hand side of (12) or (15) with that of (17). For example, from (15) we obtain

$$a_1 = 0, \quad a_2 = \frac{2\sqrt{\kappa_1^l \kappa_2^l} e^{-j\phi}}{1 + \kappa_1 + \kappa_2}, \quad a_3 = jQ_L. \tag{19}$$

The terms in the summation in (17) account for cable losses, multiple reflections, and relative phase shifts (dispersion) introduced by spurious discontinuities in the cables leading to the resonator. Normally, these factors are calibrated out using multiple sets of independent measurements. One such measurement is the transmission between the two ports without the resonator. However, this measurement is not reliable for a loop-coupled resonator, because the loops are designed to weakly couple to each other. Uncalibrated or poorly calibrated data is better analyzed by using (17) as a fit function. The inclusion of transmission line modes renders the fit non-linear.

Since the measured data is overdetermined, the problem of calculating the transformation constants in the fit function may be cast as minimization of the square error between measurements w_{m_i} and the model w_i, defined by

$$\chi^2(\mathbf{a}) = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \mid w_{m_i} - w_i \mid^2 . \tag{20}$$

Here, we assume that $N \gg m$, where m is the number of independent parameters of the fit function. For notational convenience, these parameters of (17) are stored in an array $\mathbf{a} = [a_1 \quad a_2 \quad \cdots a_m]^T$, with the superscript T denoting transpose. The factor $1/\sigma_i^2$ denotes the weighting constant for the ith sample, and may be assumed as unity without any loss of generality. When the parameters are allowed to vary from their initial estimates by differential increments, δa_j, the model w_i can be approximated by first-order Taylor's series expansion as

$$w_i \approx w_{i0} + \sum_{j=1}^{m} \frac{\partial w_{i0}}{\partial a_j} \delta a_j \tag{21}$$

where w_{i0} and the derivative are evaluated at the initial guess $\mathbf{a} = \mathbf{a}_0$. Although the model is non-linear with respect to the parameter vector $\mathbf{a}$, the Taylor's series approximation in (21) effectively linearizes the function, so that linear least squares theory can be applied.

Expressing the data, w_{m_i}, and the function, w_i, in complex form as $w_{m_i} = r_{m_i} + jx_{m_i}$ and $w_i = r_i + jx_i$, respectively, the chi-squared error in (20) may be written as

$$\chi^2 = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \left[(r_{m_i} - r_i)^2 + (x_{m_i} - x_i)^2 \right] . \tag{22}$$

We minimize χ^2 with respect to each of the parameter increments, δa_j, by setting the parametric derivatives equal to zero:

$$\frac{\partial \chi^2}{\partial(\delta a_k)} = -2\sum_{i=1}^{N} \frac{1}{\sigma_i^2}(r_{m_i} - r_i)\frac{\partial r_i}{\partial(\delta a_k)} - 2\sum_{i=1}^{N} \frac{1}{\sigma_i^2}(x_{m_i} - x_i)\frac{\partial x_i}{\partial(\delta a_k)} = 0, \mathrm{k} = 1, 2, \cdots m. \quad (23)$$

Noting from (21) that

$$\frac{\partial w_i}{\partial(\delta a_k)} = \frac{\partial w_{i0}}{\partial a_k} \quad (24)$$

and replacing w_i with its Taylor's series expansion, we obtain from (23)

$$\sum_{i=1}^{N} \frac{1}{\sigma_i^2}(r_{m_i} - r_{i0})\frac{\partial r_{i0}}{\partial a_k} + \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(x_{m_i} - x_{i0})\frac{\partial x_{i0}}{\partial a_k} = \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\sum_{j=1}^{m} \frac{\partial r_{i0}}{\partial a_j}\frac{\partial r_{i0}}{\partial a_k}(\delta a_j) + \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\sum_{j=1}^{m} \frac{\partial x_{i0}}{\partial a_j}\frac{\partial x_{i0}}{\partial a_k}(\delta a_j), \quad \mathrm{k} = 1, 2, \cdots m. \quad (25)$$

Eq. (25) can be concisely expressed in matrix form as $[\mathrm{C}][\alpha] = [\beta]$ with the elements given by

$$C_{kj} = \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\left[\frac{\partial r_{i0}}{\partial a_j}\frac{\partial r_{i0}}{\partial a_k} + \frac{\partial x_{i0}}{\partial a_j}\frac{\partial x_{i0}}{\partial a_k}\right], \quad \alpha_j = \delta a_j, \quad \mathrm{j}, k = 1, 2, \cdots m, \quad (26)$$

$$\beta_k = \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(r_{m_i} - r_{i0})\frac{\partial r_{i0}}{\partial a_k} + \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(x_{m_i} - x_{i0})\frac{\partial x_{i0}}{\partial a_k}. \quad (27)$$

The matrix [C] and the vector $[\beta]$ can be completely calculated from the initial guess $\mathbf{a}_0$. The derivatives can be evaluated analytically from the model (17). This procedure can be iterated, with the corrective offset of the parameter vector computed at each iteration from $[\alpha] = [\mathrm{C}]^{-1}[\beta]$. If the initial guess is close to the solution vector, then, this linearized least squares implementation suffices. However, for the noise-corrupted DR measurements, we found that this procedure does not converge well because of the sensitivity to initial guess, and deviations from an ideal Q circle. Therefore, the correction to vector [a] at each iteration is implemented using the Marquardt algorithm [15], which is formulated to seek a global minimum in the parameter space from a relatively crude initial guess. Marquardt's algorithm employs an interpolating parameter, λ, to influence the direction of search at each iteration. The LSM algorithm is implemented as follows:

1. Compute $\chi^2(\mathbf{a}_0)$ given the initial guess.

2. Modify the diagonal elements of [C] as $C'_{kj} = C_{kj}(1 + \lambda)$, with an initial value of $\lambda = 0.001$.

3. Compute the parametric correction, $\delta\mathbf{a}$, from $[\alpha] = [\mathbf{C}]^{-1}[\beta]$.

4. If $\chi^2(\mathbf{a} + \delta\mathbf{a}) > \chi^2(\mathbf{a})$, then, set $\lambda_{new} = 10\lambda_{old}$; else, set $\lambda_{new} = 0.1\lambda_{old}$. Recompute [C] and the new correction, $\delta\mathbf{a}$.

5. Repeat the previous step until the iterations converge, as indicated by the weighted variance changing by less than 0.01 from one iteration to the next.

4.1 Initial Guess Estimate

Two critical parameters which provide the initial guess to the LSM algorithm are the loaded resonant frequency, f_L, and the Q-factor, Q_L. The estimation scheme for f_L is based on the fact that the magnitude of either S_{12} or S_{21}, when plotted against frequency, exhibits the maximum slope at the resonant frequency. For the measured data, it has been found that these two parameters have resonant frequencies which are slightly shifted. Therefore, the arithmetic average of these two parameters is examined for maximum slope of magnitude against frequency. Specifically, the derivative $\mid dw_i/df \mid$ is calculated numerically using the central difference approximation on neighboring frequency points (except at the end points where either forward or backward difference is employed), and the values are arranged in descending order to detect the resonant frequency, f_L. The detected value is confirmed by plotting the derivative against frequency. The unloaded resonant frequency f_0 is approximated as f_L in calculating t_i as per (18). The loaded Q-factor may be estimated from the raw data using [16]

$$Q_L = \frac{\omega_L \mid dw_i/d\omega \mid}{2Re(w_i)} \tag{28}$$

where w_i represents the measured parameter. The form in (28) is convenient because the loaded resonant frequency and the derivative estimated in the previous step can be used to evaluate (28) at several frequencies in the resonant band. An average of all these closely spaced values is assumed as the best estimate of Q_L.

As an example of how these estimates of f_L and Q_L can be used in the LSM algorithm, consider the fit function in (17) with only one exponential, whose amplitude is normalized to unity. Then, one needs initial guesses for a_1, a_2, a_3 and γ_1. Clearly, a_2 can be set to the value of the function, w_i, at $f_i = f_L$. From either (12) or (15), it follows that $a_3 = jQ_L$. We have

found that the convergence of the algorithm is not sensitive to the estimate of γ_1. Therefore, we start with an estimate of $\gamma_1 = 0$. For the transmission parameters, the model implies $a_1 = 0$ for all iterations (see (19)), while it is set equal to a_3 initially for the reflection parameters.

5 Examples of Curve Fitting

We have determined the unloaded Q of the same resonator as described in Section 3.1, with the raw data for curve-fitting obtained by shifting the reference planes in eqs. (12) and (15) in accordance with the length of the input and output lines. Curve-fitting of the simulated data provides an intuitive validation of the computer program which has been written in C to implement the LSM algorithm described in the previous section. The computer program has been developed in LabWindows with a convenient graphical user interface. In order to make the validation over a wideband, simulated data lying within a band of $\pm 10(f_0/Q_0)$ is input to the LSM program. One exponential is used for the standing wave mode. Three iterations were used to correct the parameters. From the parameters of the fitted Q-circle, we have calculated $\kappa_1 = 1.39012$, $\kappa_2 = 0.65508$, $Q_L = 326.579$, which yield an unloaded Q of 994.5 (within 0.5% of the specified $Q_0 = 1000$). The phase shifts to compensate for the line lengths are estimated to be $\beta_0\ell_1 = 36.012$ deg., $\beta_0\ell_2 = 40.013$ deg., yielding a fitting error of $\chi^2 = 0.001$.

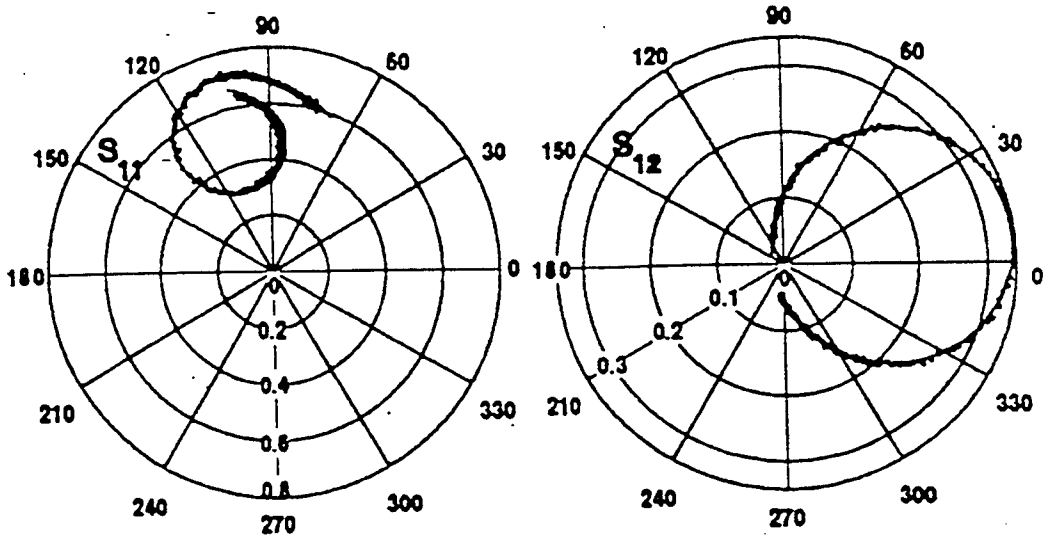


Figure 5: Comparison of the Least Squares Marquardt (LSM) curve-fit result with measured data. The fit functions for S11 and S12 are (16) and (17), respectively.

An example of the improved results obtained by using the LSM method on measured data is shown in Fig. 5, for data obtained from a dielectric resonator formed using copper endplates with a 12mm radius. The resonant frequency is 26.45 GHz. Using the LSM algorithm, the calculated Q_0 is 3710, with coupling coefficients $\kappa_1 = 0.532$ and $\kappa_2 = 0.219$, and $\chi^2 = 0.02$, as evidenced by the good match in both magnitude and phase (Fig. 5). Kobayashi's method gives $Q_0 = 4336$ for the same dataset, while Shen's method gives $Q_0 = 4206$. The discrepancies are not surprising, given that the raw measurements do not trace out perfect circles in a polar representation.

6 Conclusions

An efficient algorithm, based on least squares minimization of the square error between an assumed fractional non-linear transformation and the measured data, has been developed for the extraction of unloaded Q from dielectric resonator measurements. The convergence of the alogorithm has been enhanced using the Marquardt method. The circuit equivalent of the primary DR mode has been employed to develop idealized expressions for the resonator response, which trace out circles in the complex plane. Starting with the approach of matching Q-circles to the resonator data, we have augmented these expressions to compensate for the undesirable influence of the coupling feed structure. The resulting improved method for analyzing two-port resonators is useful when calibration is difficult. This procedure is more reliable and accurate than previous methods based on three-point resonant curve measurements (*e.g.*, Ginzton's and Kobayashi's methods), because both magnitude and phase of a wide data sweep around the resonant frequency are used to fit the measured data. Poorly calibrated data with a few errant points in the sweep can be analyzed using the LSM method. The extraction program has been applied to compute the unloaded Q of dielectric resonators consisting of copper end plates.

We conclude the report with a brief discussion on extension of this work to meet the needs of WL/MLPO. Temperature-dependent measurements of HTS films need to be implemented in the program. Although we anticipate no insurmountable difficulties in extracting the data from HTS film measurements, the non-linear transformation may need modification because of the extremely high Q's of HTS resonators. Second, detailed analysis of the electromagnetic fields in the resonator and its feed structure needs to be carried out to better understand the coupling mechanism and to minimize the influence of package modes, radiation loss. The field analysis also facilitates improved design of the DR without the lateral reflective wall (i.e., *open* HTS dielectric resonators). Third, the LSM program needs to be augmented with the capability

to adaptively change the model in accordance with the raw data. For example, the program should be able to handle arbitrary number of data points and reasonable number of standing wave modes. Error analysis also needs to be done. It would be desirable to combine the four S-parameters into one data set such as the impedance. It is anticipated that funds for this work will be provided by AFOSR SREP program and WL/MLPO.

Appendix
Determination of Coupling Coefficients

With reference to Fig. 4, let d_{11} and d_{22} denote diameter of the Q-circle for input and output reflection coefficients, respectively, while that of the corresponding coupling loss circle is denoted by d_{1c} and d_{2c}, respectively. The diameter of the transmission Q-circle is d_{12}. These diameters are obtained from the corresponding transformation vector **a** generated by the fitted curves following the procedure discussed in [9]. The diameter of the loss circle is computed as

$$d_{kc} = \frac{d_{kk}[1 - |\Gamma_{dk}|^2]}{d_{kk} - (d_{kk}/2)^2 - |\Gamma_{dk}|^2 + |\Gamma_{ck}|^2}, \quad k = 1, 2 \tag{29}$$

where Γ_{ck} denotes the center of the corresponding reflection Q-circle. The various coupling coefficients are calculated as (see eqs. (8) and (9))

$$\kappa_1^l = \frac{d_{11}/2}{1 - d_{11}\left[d_{1c}^{-1} - (d_{12}/d_{11})^2 d_{2c}^{-1}\right]} \tag{30}$$

$$\kappa_2^l = \kappa_1^l \left(\frac{d_{12}}{d_{11}}\right)^2 \tag{31}$$

$$\kappa_k^c = \kappa_k^l \left(\frac{2}{d_{kc}} - 1\right), \quad k = 1, 2. \tag{32}$$

The unloaded Q factor follows from these coupling coefficients and the loaded Q, as per (13).

References

[1] J. G. Bednorz and K. A. Müller, "Possible high T_c superconductivity in the Ba-La-Cu-O system," *Z. fur Phys.*, vol. 64, p. 189, 1986.

[2] M. K. Wu, J. R. Ashburn, C. W. Chu et al., "Superconductivity at 93K in a new mixed-phase Y-Ba-Cu-O compound system at ambient pressure," *Phys. Rev. Lett.*, vol. 58, p. 908, 1987.

[3] *IEEE Trans. Microwave Theory Tech.*, Special Issue on Microwave Applications of Superconductivity, vol. 39, no. 9, pp. 1445-1594, Sep. 1991.

[4] T. M. Klapwijk, D. R. Heslinga, and W. M. van Huffelen, "Superconducting field effect devices," in *Superconducting Electronics*, H. Weinstock and M. Nisenoff (eds.), Berlin Heidelberg: Springer-Verlag, pp. 385-408, 1989.

[5] D. Kajfez and P. Guillon (eds.) *Dielectric Resonators*, Norwood, MA: Artech House, 1986.

[6] Hewlett Packard Product Note No. 8510-3, "The measurement of both permittivity and permeability of solid materials," no. 5954-1535, Aug. 1985.

[7] W. B. Weir, "Automatic measurement of complex dielectric constant and permeability at microwave frequencies," *Proc. IEEE*, vol. 62, no. 1, pp. 33-36, Jan. 1974.

[8] L. P. Ligthart, "A fast computational technique for accurate permittivity determination using transmission line methods," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-31, no. 3, pp. 249-254, March 1983.

[9] D. Kajfez, *Q Factor*, Oxford, MS: Vector Fields, 1994.

[10] Z.-Y. Shen, *High-Temperature Superconducting Microwave Circuits*, Norwood, MA: Artech House, 1994.

[11] E. K. Moser and K. Naishadham, "HTS dielectric resonator as a microwave characterization tool," Applied Superconductivity Conference, Pittsburgh, PA, Paper No. MSD-7, August 1996.

[12] E. L. Ginzton, *Microwave Measurements*, New York, NY: McGraw-Hill, 1957.

[13] Y. Kobayashi, T. Imai, and H. Kayano, "Microwave measurement of temperature and current dependences of surface impedances for high-T_c superconductors," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-39, no. 9, pp. 1530-1538, Sep. 1991.

[14] W. P. Wheless and D. Kajfez, "Experimental characterization of multi-moded microwave resonators using automated network analyzer," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-35, no. 12, pp. 1263-1269, Dec. 1987.

[15] D. W. Marquardt, "An algorithm for least-squares estimation of non-linear parameters," *J. Soc. Ind. Appl. Math.*, vol. 2, pp. 431-441, 1963.

[16] D. Kajfez and W. P. Wheless, "Invariant definitions of the unloaded Q factor," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-34, no. 7, pp. 840-841, July 1986.

Final Report of
A SUMMER FACULTY PROJECT FOR ANATOMICAL FEATURE EXTRACTION FOR REGISTRATION OF MULTIPLE MODALITIES OF BRAIN MR

Timothy S. Newman

Assistant Professor
Department of Computer Science
University of Alabama in Huntsville
Huntsville, AL 35899

Final Report for:
Summer Faculty Research Program
Wright Laboratory (WL/AARA)
Wright-Patterson AFB, OH

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

September 1996

A Summer Faculty Project for Anatomical Feature Extraction for Registration of Multiple Modalities of Brain MR

Timothy S. Newman
Department of Computer Science
University of Alabama in Huntsville
Huntsville, AL 35899

Abstract

In this report, we introduce and discuss techniques developed for the extraction of anatomical structures in head MR images. The techniques include methods for the extraction of the eyes, brain lateral ventricles, and brain longitudinal fissure. The ultimate goal of the feature extraction is to allow the registration of different modalities of MR datasets of the same individual (for example, registration of a T1-weighted dataset with MRA, T2-weighted, and proton density datasets). After describing the methods, preliminary results of their application to a real dataset are exhibited.

A Summer Faculty Project for Anatomical Feature Extraction for Registration of Multiple Modalities of Brain MR

Timothy S. Newman

1 Introduction

In conventional medical practices, tomographic datasets, such as CT (computerized tomography) or MR (magnetic resonance imaging), are not used directly during surgery, but rather are used in an indirect mode for diagnosis and treatment planning. The traditional approach involves examining the tomographic data as a series of cross-sectional (two-dimensional) slices rather than as a volume of data. Medical staff examining the series of slices must attempt to mentally combine the slice information into a volume of data. A capability to view the data as a volume rather than as a series of slices may allow improved chances of accurate diagnosis and treatment. A number of medical centers have increasing capabilities for volumetric visualization of tomographic data and for image guidance during surgery, but, in general, capabilities are limited. One goal of our work is to more fully exploit the data through volumetric visualization.

Visualization involves the use of computer-synthesized images or pictures to discover or highlight data geometry and topology (that is, the relationships between the data) [20]. A common component of visualization is the hi-resolution animation of time-varying data [29]. By volumetric visualization, we mean extracting, modeling and rendering volumetric data, for example, by classifying anatomical structures or by viewing the interior of volumes. In this report, we present techniques for anatomical structures extraction for the purpose of registering datasets together. Ultimately, the registered datasets will be used in a volumetric visualization.

In this project, our primary application area is image-guidance for neurosurgery. Conventional neurosurgery usually involves the use of stereotactic frames that act as fixtures that keep the head rigidly in place during surgical procedures[1]. Surgical instruments are often attached directly or indirectly to the frame to allow operative procedures to be planned and executed using the stereotactic coordinate system. The CT and/or MR images acquired of a patient wearing a frame can be registered together, thus allowing potential comparison of images for diagnosis, treatment planning, and surgical planning. (Some frames provide registration by rigidly fixing the patient in place while other frames are made of materials that produce fiducial markers in the acquired CT and/or MR images. The datasets can then be registered by alignment of fiducial markers.) Some clinical centers attach fiducial markers (often small beads) directly to the scalp of a patient rather than use frames. For example, Malison et al. [24] have registered SPECT and MR datasets

[1]The frames can be somewhat unwieldy and are usually attached to screws drilled into the skulls of patients.

using external fiducial markers attached to a patient's scalp. Barnett et al. [3] of Cleveland Clinic have used tomographic images for surgical guidance using external fiducial markers and an ultrasonic probe for registration of probe measurements to pre-operative images using the markers as registration landmarks. Philips Medical Systems has developed a system with Aachen University that also uses external fiducial markers [1]. Their system tracks infrared emitters attached to a surgical probe and, like the Cleveland Clinic approach, uses the markers as registration landmarks. One of the goals of the larger project of which our work is a part is to provide frameless image-guided neurosurgery (although our concern is more on image-guidance than on frame replacement). One of the barriers to more effective use of tomographic data is the difficulty of registering images, particularly in studies conducted over time [25].

Background information about stereotaxy is available in a number of books and papers. Readers with a computer science background may wish to consult [30] for a discussion of stereotaxy, CT, and the role of computer graphics or to consult [31] for an introduction to recent advances and issues in computer-assisted surgery. General introductions to neurosurgery and stereotaxy are available in [21] and [32].

In this report, we discuss methods for extracting three anatomical structures in different MR modalities. The extracted features support the registration of the different image modalities with one another. By registering different modality images, it may be possible to supply a richer array of information to medical staff for planning and carrying out neurosurgery. For example, in MRA (magnetic resonance angiography) data, certain blood vessels appear as high-contrast features whereas in T1-weighted images, white matter has a distinct contrast from gray matter. T2-weighted images are appropriate for detection of cerebral fluid's boundary with white or gray matter [18] and for the discovery of regions of edema caused by tumor cell invasions, particularly for glial neoplasms [21]. If MRA, T1-, and T2-weighted images, and proton density data can be registered together, the different classes of information present (for example, the locations of blood vessels, white matter, and gray matter) can be integrated. In neurosurgery, trajectory planning involves the avoidance of major cortical blood vessels as well as areas of functional cortex [21, 32]. Trajectory planning may seek to exploit the sulci, also; the sulci may provide an access path to a region that has less invasion of white or gray matter [21, 32]. For image-guidance in neurosurgery, these areas must be known. Information about the tissues in the vicinity of the operative site is also helpful to the surgeon [28]. By registering and then exploiting data from multiple inputs (i.e., different modes of MR), we hope to make automated or semi-automated extraction of needed information for trajectory planning more reliable. We begin the remainder of the report with a background on volume registration and then discuss the methods that we have developed for feature extraction. Experimental results for one patient are included.

2 Volume Registration Background

The problem of registering datasets is an important one and has attracted attention from a number of researchers. The methods in the literature include experimental methods of registering MR data with PET or CT data from the same patient, registering MR patient data with a generalized atlas, and registering MR datasets (of the same MR modality) collected over time from the same patient. Van den Elsen, Pol, and Viergever have reviewed many of the registration methods in [33]. In this report, we do not discuss the MR-PET registration literature but instead focus on the MR-MR registration methods.

A number of efforts have been reported recently to register MR data from a patient with atlas data (or vice versa). The atlas data consists of labeled tissue of a "model" MR image (usually of a normal patient) or from textbook atlases built from histological cross-section information. Many, perhaps most, of the methods build an individualized patient atlas by deforming the patient data to the atlas data. Christensen at al. [6, 7] have recently developed a technique that deforms individual normal patient MR data to an atlas. Their atlas is a reference MR image of another normal patient. Like many of the deformation-based approaches, their method is based on models of elastic solids and viscous fluids. In [7], they presented a massively parallel implementation of the deformation method on the MasPar MP2. Gee, Reivich, and Bajcsy [12] have developed a technique that elastically deforms an atlas constructed from a normal patient to a candidate MR dataset from another normal patient. Their atlas consisted of manually-created outlines of a number of brain structures. The surface of the brain and edges of the ventricles were extracted from the candidate data after manual removal of the skulls from the datasets. Davatzikos et al. [8] have registered a 2D slice of a patient MR dataset to a digitized representation of a slice of the Talairach brain atlas. Their method achieved the registration by deforming the atlas image. They did not address the issue of determining the full 3D transformation of one dataset to another, however.

Methods have also been presented to register MR and CT data. Hemler et al. [16], for example, registered MR and CT data of the same patient by matching polygonalized boundaries that approximate the shape of the skin surface. Their system was not always able to properly extract the skin boundary automatically, thus an interactive editing of the dataset was necessary to fully and accurately extract the skin boundary.

Grimson et al. [13] have also addressed the registration of data sets, although their interest is registering a surgeon's view of a live patient with MR or CT data of the patient. Their work's aim is providing image guidance for neurosurgery. Their method consists of the acquisition of a laser range image of the patient's head and the registration of the range data with the MR or CT dataset. In their technique, a coarse estimate of the registration transformation is first generated by comparing pairs of points in the range data with pairs of points extracted from the outline of the skull in the MR or CT data. The transformation is then refined by minimizing a measure of the mismatch between the remainder of the range and MR (or CT) dataset's

extracted skull boundary.

At the Montreal Neurological Institute (MNI), MR and PET data are commonly registered before neurosurgery to aid in planning and then registered with vascular images (DSA) acquired during surgery [28]. MR and PET data are registered through least-squares matching of manually extracted corresponding feature points. Intra-surgical monitoring is accomplished either through the known transformation of a stereotactic frame, or (for frameless surgeries) by manually picking points on the surface of the head and their corresponding points on pre-operative MR data and then computing the transformation via least-squares.

Höhne et al. [17] have developed a multi-media tool that contains a labeled hierarchical solid model of the brain that is useful for surgical rehearsal. The tool could potentially serve as a brain atlas as well.

2.1 Tissue Classification

Many papers have appeared in the literature discussing classification of tissue in tomographic images. For brain MR, many of the methods locate major tissue classes (e.g., white matter, gray matter, and cerebro-spinal fluid (CSF)) but not necessarily anatomical structures. Hornak and Fletcher [18], for example, have had some success in segmenting major classes of tissue in brain MR using multi-dimensional histograms that compare inverse T1, inverse T2, and proton-density values for corresponding pixels. Kapouleas [19] has extracted the outline of the brain and suspected Multiple Sclerosis lesions using local edge detection in proton density and T2-weighted data. Arata et al. [2] have segmented ventricles, sulci, and white matter lesions from proton density and T2-weighted data. They first detected white matter lesions using a histogram-based approach. They next used region growing to find ventricular regions from hand-selected seed points. After removing the extracted ventricular regions from the data, they segmented the sulci using K-means clustering. Li et al. [23] have segmented major tissue classes (white matter, gray matter, cerebro-spinal fluid, and potential tumors) using symmetry, morphological operations, and measures of segmented region size.

3 Feature Extraction

In our summer work, we have developed techniques for extraction of the eyes, lateral ventricles, and the longitudinal fissure. We extract these structures to support the registration of T1-weighted, T2-weighted, proton density, and MRA datasets of one patient. We assume that the patient's position between data collections is not controlled; the patient's position may change between scans. The eyes, ventricles, and fissure were chosen because they appear to be distinct in all of the dataset types, increasing the likelihood of their successful extraction. In this section, we first discuss our extraction of the eyes, then the lateral ventricles, and finally the longitudinal fissure.

3.1 Eye Extraction — A New Technique for Spherical Fitting

Our method extracts the eye by exploiting the spherical shape of the eye. We also exploit model information including eyeball size and the homogeneity of the vitreous humor signal response in an MR dataset. The goal in eye extraction is to determine the position of the center point of each eye. The eye center points serve as registrable landmark features.

The spherical shape of the eye is exploited in a number of ways in our eye detection technique. Since the eyeball is approximately spherical, the intersection of the parallel planar tomographic slices with the eyeball forms circular boundaries. We detect circular boundaries in each slice using a Hough Transform-based approach, then use model information to prune the list of boundaries, and finally cluster hypothesized circles together to gather evidence of spherical structures in the data.

Hough-based approaches have been popular for parameter estimation problems in the literature. For example, Hebert and Ponce [15] have used Hough techniques for classifying surfaces as conical, cylindrical, or planar. Previously, the author has developed Hough-based approaches for spherical and cylindrical parameter estimation in range images [26].

Our eye extraction method consists of the following steps. First, we use a 2D edge detector to locate edges in all the slices of the datasets that are to be registered. Next, for each slice, we perform Hough circle detection. From this set of hypotheses, circles unlikely to be part of the eye are then pruned. The datasets are then resliced and the three above steps applied to each of the resliced datasets. The three initial steps are then repeated for each reslicing. Finally, the surviving circle hypotheses are clustered using a hierarchical clustering technique. For each cluster, least squares estimation is used to obtain the parameters of the best-fitting sphere. In the remainder of this section, the details of the method are presented and then some results are exhibited.

3.1.1 Edge Detection

The first stage of the eye extraction technique consists of edge detection in each slice of the MRA, T1-weighted, T2-weighted, and proton density datasets. To minimize the effects of noise, each slice is median filtered before edge detection. The edge detection is accomplished using a 2D Sobel edge detector. The lowest-intensity edge pixels in the edge image are discarded to improve robustness to noise and also to help limit the processing time for downstream stages of the extraction algorithm. We have experimented with several thresholds on edge intensity. Preliminary results indicate that discarding the weakest 20% of the edge pixels allows downstream processing to avoid some processing while also eliminating some image artifacts. For some of the datasets, it may be possible to discard the weakest 40% of the edges without affecting final results.

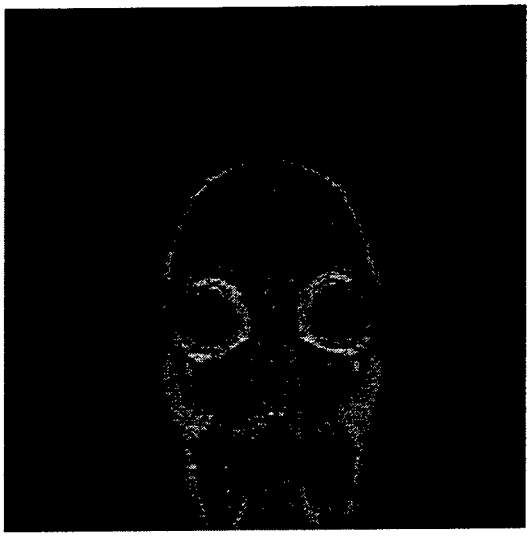


Figure 1: One Slice (containing eyes) from a Proton Density Dataset.

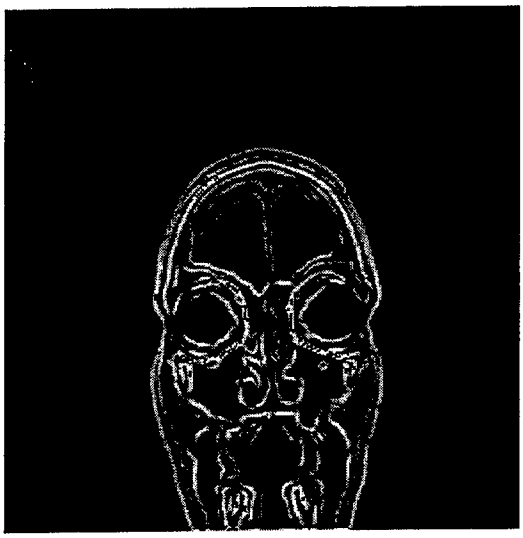


Figure 2: Edges Detected in Slice of Median-Filtered Proton Density Dataset.

The image in Figure 1 is one slice from a proton density dataset and Figure 2 shows the edges detected in the slice after median filtering.

3.1.2 Circular Cross-section Detection

The next stage of our procedure involves detection of circular features in the slice images. Since the eyeball is spherical in shape, the intersection of the parallel planar slices with the eyeball should be circular. For each slice, we detect circular features using the well-known Hough circle detection.

The Hough transform was originally proposed as a method to detect linear structures in images by Hough [9]. The Hough method accumulates evidence for linear structures through the use of a finite number of parameter bins. Evidence in the image is used to increment bins and at the end of evidence accumulation, the highest-count bin(s) is (are) used as the estimates of the structure's parameters. For linear structures, the Hough method consists of incrementing bins associated with the parameters of all lines that pass through each pixel in the edge image. Accumulation methods (i.e., Hough methods) can be used to recover parameters of other shapes and surfaces in the image as well. For example, Kimme, Ballard, and Sklansky [22] were the first to extend the Hough technique to the detection of circles. We have previously used other generalizations of the Hough technique to detect and fit both spherical and cylindrical surfaces in images [26][2].

In the Hough circle detection technique, a circle is parametrized by its centroid, (x_0, y_0, z_0), and radius, r. For each pixel P_i in the filtered edge image, the parameter bins associated with all circles that could pass

[2]The spherical fitting method developed in our summer work is a new technique distinct from our prior spherical fitting method. The current method includes an extraction step that the prior method did not have and also uses a different approach to estimate the sphere parameters.

through P_i are incremented. In practice, this is accomplished by conducting Hough circle centroid detection for each desired radius r_j. Parameter z_0 can be taken as a known value using the offset of the slice from the base slice (coupled with the known inter-slice spacing). Thus, for each (r_j, z_0), the Hough circle method estimates (x_0, y_0).

The heart of Hough circle detection is that the counts of parameter bins associated with each circle whose boundary contains edge pixel P_i are incremented. Kimme et al. determined circle center location by computing $(rcos\theta, rsin\theta)$ for fixed r while θ was increased in increments of $\Delta\theta$ from 0 to 2π. This is equivalent to generating the circular collection of pixels centered at each P_i. However, there are more efficient ways to generate the sequence of pixels comprising this circular collection. We have improved the performance of the method by using the Bresenham Circle Generation Algorithm [4] as formulated by [11] to generate the sequence. The Bresenham method uses only integer calculations and exploits 8-way symmetry to greatly reduce the number of calculations for circle generation. Bresenham's method employs second-order partial differences to calculate the change in y for each increment of x in one octant of a circle's boundary. Unlike an increment by $\Delta\theta$, the Bresenham method results in each boundary pixel being visited exactly once.

In each slice, we retain the top five bins associated with each radius. The radii tested were the integer radii ranging from 2 to 18 pixels, inclusive. In the datasets we had access to, intra-slice pixel size was 1mm. Thus, our method is designed to detect eyes between 4 and 36mm in diameter. We believe that this range of values should in principle allow our method to be used for nearly all human datasets.

3.1.3 Hypothesis Pruning

Following generation of a set of circle centroid hypotheses, model information is used to prune the set. The region interior to the circle boundaries of with eyes contains the vitreous humor. Vitreous humor is homogeneous in the MR data, exhibiting signal amplitudes within a narrow range of values. We exploited this information by eliminating hypothesized circles whose interior contained less homogeneous tissue. It was our goal for the choice of thresholds in signal amplitude to be non-critical to the step, so we deliberately selected thresholds outside the typically observed range of values for vitreous humor tissue.

The pruning stage eliminates over 80% of the hypotheses. In particular, it eliminates hypotheses generated from edge pixels near the eyeball. There appear to be tissue inhomogeneities outside but near to the eye which cause the generation of many edge pixels. Since some of the detected circle boundaries contain some of these edge pixels as well as some true eye edge pixels, the pruning is effective in eliminating these detected circles as well as other circles in the datasets whose interior contain tissue that is not vitreous humor.

3.1.4 Reslicing and Non-isotropic slices

The next stage of processing involves re-slicing the datasets and then re-applying the edge detection, circle detection, and pruning steps. By re-slicing the data, we mean that the data is viewed as slices with an inter-slice axis perpendicular to the original inter-slice axis. In principle, we could perform two re-slicings of the data, one using the original x axis as the new inter-slice axis and the other using the original y axis. In practice, we use a single re-slicing where the original x axis serves as the new inter-slice axis. Essentially, this means that following reslicing, the processing is applied to the $y - z$ slice planes of data (earlier processing operated on the $x - y$ planes of data). In most MR datasets, the original inter-slice spacing is greater than the intra-slice spacing, thus the re-sliced images usually are non-isotropic.

For the non-isotropic slices, a larger number of hypotheses tend to be generated and to survive pruning. It is not unusual for over 1000 hypotheses to survive pruning for the non-isotropic slices. This is because more circles can be fit through the larger pixels of the more coarse-resolution data. We further prune this resliced data to reduce the amount of down-stream processing using two steps. First, we cluster each slice's circle hypotheses using single-link clustering. We discuss the single-link clustering method under the next heading.

The result of the single-link clustering on each resliced slice is a grouping of hypotheses into sets that contain circles of similar size and similar positions. In each of these groups, we remove the high-variance hypotheses. Specifically, if there are multiple circle hypotheses, we retain only the lowest-variance hypothesis. We also retain any other hypothesis whose variance is less than 100. This allows the removal of redundant data while speeding up subsequent processing.

Single-Link Clustering

The Single-Link Clustering method (also called agglomerative hierarchical clustering) [10] is a well-known classification technique. The method is a greedy algorithm that merges pattern vectors of data into classes. The technique's steps are as follows. Initially, each pattern vector P_i is the sole member of a cluster C_i. Clusters are then iteratively merged until only one cluster remains or until no clusters can be merged. At each step of the iterative merging, exactly two clusters are merged. The two clusters merged are the closest two clusters, as defined by a distance measure. However, if the closest clusters are greater than a distance T apart, no clusters are merged and the algorithm terminates.

In Single-Link Clustering, the distance measure of cluster distance between a cluster C_k and C_l is the distance between whichever pattern vector P_m in C_k is closest to a vector P_n in C_l. That is, $d(C_k, C_l) = min\ d(P_a, P_b), \forall\ P_a \in C_k$ and $P_b \in C_l$. We define distance as Euclidean distance.

In our usage of the Single-Linkage Clustering, our pattern vectors are of the form (x_0, y_0, z_0, R), where (x_0, y_0, z_0) is a hypothesized circle centroid and R is that circle's hypothesized radius. We have experimented

with a number of distance thresholds, T. We have found that a squared distance between 19 and 41mm^2 allows circles from each eye to be clustered separately while not splitting hypotheses from the same eye into multiple clusters. Squared distances larger than 41mm tend to allow circles that are not part of the eyeball's cross-section to be merged with hypotheses that are part of the eyeball's circular cross-sections. Squared distances less than 19 mm can cause some circles from the same eye to be in different clusters. Typically, this results if a circle is detected in a slice i and slice $i+2$ but not slice $i+1$.

3.1.5 Clustering Circles to Find Spheres

The single-link clustering is also used to cluster the combined collection of hypotheses from both the original and the resliced data. In this subsection, we discuss this usage of the clustering in our extraction and fitting method.

The clustering of hypotheses from the combined collection groups nearby circles to build evidence for spherical structures in the data. Grouping is achieved through the use of the Single-Link Clustering applied to all surviving hypotheses. We reject any clusters that contain circles of only small radii as well as any clusters that do not contain circles detected in both the original data and in the resliced data. These steps help eliminate non-spherical structures from consideration. For example, small cylindrical structures in the data will be pruned from consideration.

3.1.6 Fitting Spheres to Clustered Data

In the final stage of our eye extraction method, we determine the sphere that best fits each surviving cluster of circular hypotheses. Our sphere-fitting method fits spheres to spherical boundaries, thus we apply the sphere-fitting to the boundaries of the hypothesized circles in each cluster. For each cluster, the boundary points we use are the collection of edge pixels that originally "voted" for any of the cluster's hypothesized circles. Duplicate pixels are removed to avoid biasing the spherical fitting. Spheres are fit to each collection of boundary points using the least squares fitting. Although the sphere-fitting problem might appear to be non-linear, Hall et al. [14] have observed that quadric fits are linear in the coefficients of the surface equation for spheres (and other quadric surfaces).

The spherical equation that is fit is

$$x^2 + y^2 + z^2 + g'x + h'y + i'z + j' = 0,$$

where $g' = -2x_0$, $h' = -2y_0$, $i' = -2z_0$, and $j' = -r^2 + {x_0}^2 + {y_0}^2 + {z_0}^2$, with (x_0, y_0, z_0) representing the center of the sphere and r representing the sphere's radius. Ultimately, the goal of our fitting is to recover (x_0, y_0, z_0) as a feature for registration of datasets.

Dataset	x0	y0	z0	rad0	x1	y1	z1	rad1	Distance
MRA	99.6	46.5	2.5	17.25	160.2	49.2	3.1	17.3	60.64
Proton Density	163.9	151.2	171.0	15.51	98.5	152.3	174.1	15.79	65.47
T1-weighted	155.9	144.9	193.1	12.53	93.1	146.5	193.5	12.53	62.81
T2-weighted	162.0	149.4	165.7	12.97	99.6	151.2	173.5	12.87	62.86

Table 1: Sphere Centers, Radii, and Distances in 4 datasets of one individual.

If there are n boundary points (x_i, y_i, z_i), the fitting problem can be formulated as

$$\begin{bmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ & \vdots & & \\ x_n & y_n & z_n & 1 \end{bmatrix} \begin{bmatrix} g' \\ h' \\ i' \\ j' \end{bmatrix} = \begin{bmatrix} -{x_1}^2 - {y_1}^2 - {z_1}^2 \\ -{x_2}^2 - {y_2}^2 - {z_2}^2 \\ \vdots \\ -{x_n}^2 - {y_n}^2 - {z_n}^2 \end{bmatrix}$$

We solve for g', h', i', and j' using least squares, and then determine x_0, y_0, z_0, and r from these. We accept the two hypothesized spheres of largest radius as the eyeballs.

The results of applying our methods to non-registered datasets collected from one individual are shown in Table 1. In the table, all measurements are shown in millimeters. Although the estimated eye radii vary from dataset to dataset, the radii for each eye within a dataset are very similar. Because different data types may exhibit different response to tissue types, the edges that are detected and grouped into spheres may not be identical across datasets. Since the ultimate goal of the fitting is the registration of eye centroids, perhaps a better measure of the algorithm's capability is the distance between the eye centers. We would expect this value to be more constant among the datasets. The difference in these distances in the test data may be partially caused by the eye lenses being more visible in certain datasets. When the lens is visible in a dataset, it tends to cause the generation an edge in the edge image, which biases the sphere center calculation to some extent. Ultimately, our method may need some refinement to address or exploit the effects of the lens. Furthermore, the method needs to be tested on more data.

Our eyeball extraction method is designed to find a common eye center point for use as a landmark fiducial feature among the datasets. We believe that the method may also be useful to extract other spherical features from range or tomographic data.

3.2 Ventricle Extraction

For extraction of the lateral ventricles, we have applied a technique that we earlier developed for extraction of the kidney, kidney collecting system, and kidney tumors [27]. The technique relies on a human-supplied

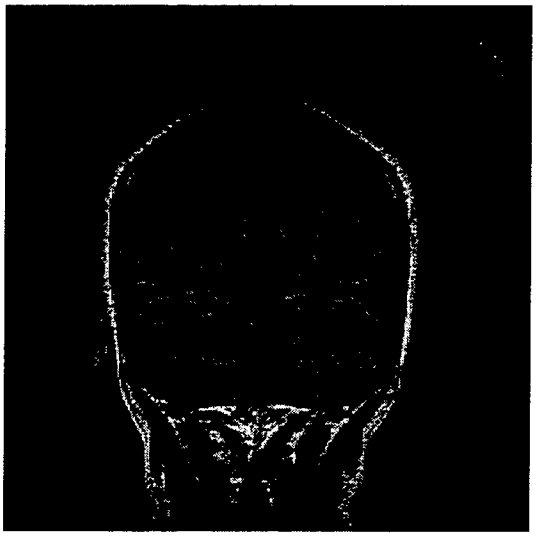


Figure 3: T1-weighted slice containing ventricles.

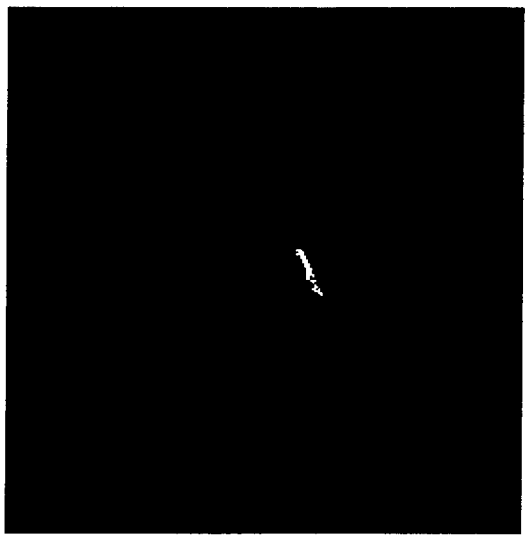


Figure 4: Extracted Left Ventricle using Volume Growing

initial seed. A single seed is selected in each of the lateral ventricles. The method for extraction of the ventricles has three stages, consisting of volume-growing, followed by morphological refinement, and finally locally adaptive histogramming refinement. For registration purposes, we propose that the center of mass of each ventricle be used.

Our volume-growing method consists of finding a set of connected voxels that contains a seed voxel V. The location of the seed is taken as the point chosen by human input. However, as the intensity value for the seed, we use the median voxel value M in a three-dimensional window about the seed voxel V. The use of the window makes the method more robust. We have found that a $3 \times 3 \times 3$ window produces good segmentations. The set of connected voxels is formed from all voxels whose intensity values are within T units of the median window value M. The volume is grown using 8-connected component growing within each slice plane and 2-connected component growing out of the slice plane (i.e., only into voxels directly above or directly below another voxel already in the grown volume).

A slice from a T1-weighted data set that contains the lateral ventricles is shown in Figure 3. Our segmentation of the left ventricle in the slice using volume growing is shown in Figure 4.

We refine this original segmentation by performing a three-dimensional close operation followed by a single dilation. The close removes isolated interior voxels and the dilation expands the region boundary to help capture voxels partially within the ventricle region.

Next, we refine the segmentation further using locally adaptive histogramming. Specifically, we dilate the segmented volume six times. We then compute the histogram for the slice containing the original seed voxel V. The histogram mode that contains the initial seed intensity is used to define the ventricular region.

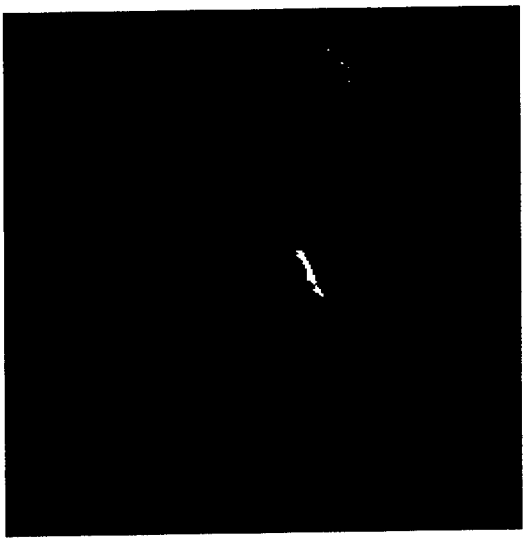


Figure 5: Final segmentation of left ventricle in slice.

The peak of this histogram is then used as the seed for application of the histogramming method to the slice adjacent (above and below) the starting slice. This is repeated until histograms have been computed and a mode detected in each slice. The mode of each histogram is defined using the local minima nearest the value of the seed. If a local minima is a small "dip" in the histogram, the minima is ignored and the next minima is searched for. The final segmentation of the left ventricle in the slice from Figure 3 is shown in Figure 5. Shaded 3D renderings of the extracted ventricle volume are shown from three viewpoints in Figure 3.2.

Our method for extracting the lateral ventricles uses volume growing, morphological operations, and locally adaptive histogramming. For registration purposes, we propose that the center of mass of the extracted volume be used.

3.3 Fissure Extraction

For extraction of the longitudinal fissure, we are currently adapting the method of Brummer [5]. Brummer's method uses a Hough-based approach to locate a planar collection of linear structures. The fissure is the largest planar structure in the brain region. Brummer locates the fissure using two 2D Hough transforms, producing a method with a complexity that is one order of magnitude less than a full 3D Hough method. In this section, we discuss our work in progress to locate the fissure.

Brummer's method focuses on an area of the head likely to contain the fissure by eliminating from consideration parts of the dataset close to the sides of the head. Brummer eliminated all voxels whose distance from the skull is less than 15% of the total width of the skull. We have found that a threshold of 10% eliminates less voxels that are part of the longitudinal fissure and use that threshold instead. We detect

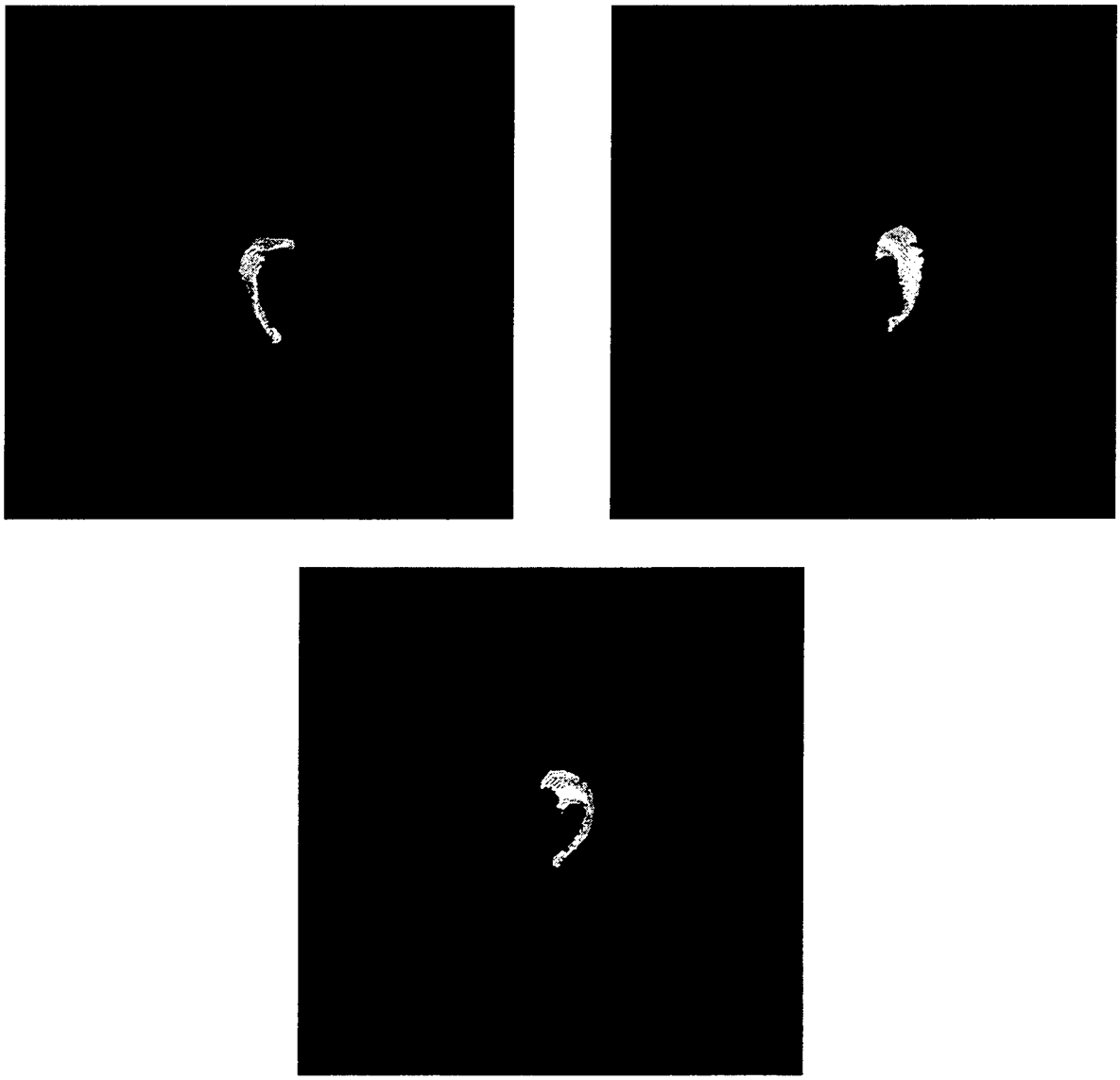


Figure 6: Shaded 3D Renderings of Extracted Ventricle.

the outline of the head (i.e., the skull) as follows. We construct a mask image for each slice that contains the skull and nearby tissue. The mask initially is a copy of the original dataset. We then remove from consideration all voxels in the dataset that have very small intensities. All pixels whose intensity is above this threshold (typically a threshold of a few percent (e.g., 4%) of the maximum image intensity is used) are used to form a binary mask dataset and are set to a value of 1 in the mask. The mask image is then eroded once to remove outliers. We further remove background by eliminating any remaining outliers, that is, isolated points adjacent only to zero-valued pixels in the mask image. The mask image is then dilated twice to fill holes and close gaps in the remainder of the image. Any isolated solitary zero-valued pixels in the mask image are then filled. The outline of the largest connected region in the mask image is a reasonable approximation of the skull. We compute the width w of the bounding box containing this region outline in each slice and finally set to zero any pixel in the mask that is less than a distance of $(0.10)w$ from the region outline. Most of the background noise outside the head is thus removed from consideration through use of the mask. The remaining non-zero pixels in the mask form the region where fissure detection will take place in edge images formed from the original datasets.

We use 2D Sobel edge detection to form an edge image from each slice. As Brummer has required in his algorithm, we require that the slicing of the data have a slice axis that is not perpendicular to the fissure; the fissure is not easily detectable in slices that are parallel to the fissure. We detect line features in each slice's edge image using a Hough-based method where the parametrization for each line is the distance of the line from the origin and the angle of the line with the horizontal.

Once dominate line features have been detected in each slice plane, the fissure plane can be detected. The fissure plane orientation can be parametrized by an angle α and an angle θ, where α is the angle of intersection of the plane with the x axis in each of the slices and where θ is the angle of the intersection of the slice normal $\vec{n}$ with the fissure plane. To clarify, if $\vec{n}$ is taken as the axis of a cone with circular cross-section, where the "tip" of the cone is the intersection of the fissure plane with the $\vec{n}$ axis, then θ is the maximum half-angle of the cone at the tip.

Following the method of Brummer, we estimate α as follows. For each slice, the largest-count Hough accumulator bin associated with each distance d is determined, where d is the distance of the line of fissure intersection from the slice origin. The angle α_i associated with the bin from each slice S_i is then used in building a global histogram. The histogram contains the sum (over all the slices) of bins associated with angles α_i for each distance d. The peak of this histogram is taken as the estimate $\hat{\alpha}$ of the angle. θ can then be estimated by using a second Hough transform. This Hough transform uses a subset of the bin counts of the first Hough transform. Specifically, all bins in all slices that were associated with angle $\hat{\alpha}$ are used. Since the fissure plane could pass through points in the slices associated with these bins, each bin is used to increment the collection of parameter bins (θ, n_i) (where the θ parameter is the angle of intersection between

the fissure and the $\vec{n}$ axis (as explained above) and n_i is the location where a (fissure) plane passing through slice S_i would intersect the $\vec{n}$ axis) for all possible planes. The distance d's in the original Hough bins are used to compute the appropriate n_i associated with each possible intersection angle θ, where $0 \leq \theta \leq 180$ with increments of θ by $\Delta\theta$ (we use 5 degree increments). The new Hough space is then searched for a maximal bin. This bin determines the plane's θ parameter and the point n_i where the plane intersects the $\vec{n}$ axis, uniquely completing the fissure plane fitting.

We are currently completing implementation of the second Hough stage and will evaluate the utility of this method following completion.

4 Future Work

The fissure plane, eye centroids, and lateral ventricle centroids together define registrable landmarks useful for registration. Our future work will include computing the rigid transformation that aligns datasets by registration of corresponding structures. We anticipate that it is likely that this global transformation may require local adjustment of transformation parameters, perhaps using a gradient descent method. Furthermore, it may be necessary to apply some small local deforming transformations to bring local regions of the datasets into better correspondence. Future testing should include applying the methods to more datasets; to fully analyze the utility of our methods, a rich set of test data is needed.

5 Conclusions

We have presented techniques to extract three anatomical features—the eyeballs, the lateral ventricles, and the longitudinal fissure—from MR datasets. For extraction of the eyeballs, we have developed a new technique that uses circular Hough transforms, model-based hypothesis pruning, hierarchical clustering, and least squares fitting. This technique is also generally applicable to the problem of sphere fitting in volumetric data. For extraction of the lateral ventricles, we use a method based on volume growing with morphological operations and locally adaptive histogramming. The longitudinal fissure extraction technique is being completed through the use of Brummer's method [5]. Preliminary indications are that the eyes and ventricles can be reliably extracted using the methods, but more data is needed to test the methods. The methods are aimed at extraction and identification of features for dataset registration. By registering multiple datasets, enhanced visualizations for neurosurgical planning and guidance may be possible.

References

[1] Ludwig Adams, Achim Knepper, Dietrich Meyer-Ebrecht, Rainer, Rüger, and Willem van der Brug, "An Optical Navigator for Brain Surgery," *IEEE Computer*, Vol. 29, No. 1, 1996, pp. 48-54.

[2] Louis K. Arata, Atam P. Dhawan, Joseph P. Broderick, Mary F. Gaskil-Shipley, Alejandro V. Levy, and Nora D. Volkow, "Three-Dimensional Anatomical Model-Based Segmentation of MR Brain Images Through Principal Axes Registration," *IEEE Transactions on Biomedical Engineering*, Vol. 42, No. 11, Nov. 1995, pp. 1069-1078.

[3] G. H. Barnett, D. W. Kormos, C. P. Steiner, D. Piraino, J. Weisenberger, F. Hajjar, C. Wood, and J. McNally, "Frameless Stereotaxy Using a Sonic Digitizing Wand: Development and Adaptation to the Picker ViStar Medical Imaging System," in *Interactive Image-Guided Neurosurgery*, edited by R. J. Maciunas, American Association of Neurological Surgeons: Park Ridge, IL, 1993, pp. 113-119.

[4] J. E. Bresenham, "A Linear Algorithm for Incremental Digital Display of Circular Arcs," *Communications of the ACM*, Vol. 20, No. 2, 1977, pp. 100-106.

[5] Marijn E. Brummer, "Hough Transform Detection of the Longitudinal Fissure in Tomographic Head Images," *IEEE Transactions on Medical Imaging*, Vol. 10, No. 1, March 1991, pp. 74-81.

[6] Gary E. Christensen, Richard D. Rabbit, and Michael I. Miller, "3D Brain Mapping using a Deformable Neuroanatomy," *Physics in Medicine and Biology*, Vol. 39, 1994, pp. 609-618.

[7] Gary E. Christensen, Michael I. Miller, Michael W. Vannier, and Ulf Grenander, "Individualizing Neuroanatomical Atlases using a Massively Parallel Computer," *IEEE Computer*, Vol. 29, No. 1, 1996, pp. 32-38.

[8] Christos Davatzikos, Jerry L. Prince, and R. Nick Bryan, "Image Registration Based on Boundary Mapping," *IEEE Transactions on Medical Imaging*, Vol. 15, No. 1, 1996, pp. 112-115.

[9] Richard O. Duda and Peter E. Hart, "Use of the Hough Transformation to Detect Lines and Curves in Pictures," *Communications of the ACM*, Vol. 15, No. 1, 1972, pp. 11-15.

[10] Richard O. Duda and Peter E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons: New York, 1973.

[11] J. Foley, A. van Dam, S. Feiner, and J. Hughes, *Computer Graphics: Principles and Practice, Second Edition in C*, Addison-Wesley: New York, 1996.

[12] Jim C. Gee, Martin Reivich, and Ruzena Bajcsy, "Elastically Deforming 3D Atlas to Match Anatomical Brain Images," *Journal of Computer-Assisted Tomography*, Vol. 17, No. 2, 1993, pp. 225-236.

[13] W. Eric L. Grimson, G. J. Ettinger, S. J. White, T. Lozano-Pérez, W. M. Wells, and R. Kikinis, "An Automatic Registration Method for Frameless Stereotaxy, Image Guided Surgery, and Enhanced Reality Visualization," *IEEE Transactions on Medical Imaging*, Vol. 15, No. 2, 1996, pp. 129-140.

[14] Ernest L. Hall, James B. K. Tio, Charles A. McPherson, and Firooz A. Sadjadi, "Measuring Curved Surfaces for Robot Vision," *IEEE Computer*, Vol. 15, 1982, pp. 42-54.

[15] Martial Hebert and Jean Ponce, "A New Method for Segmenting 3-D Scenes Into Primitives," *Proceedings of the Sixth International Conference on Pattern Recognition*, 1982, pp. 836-838.

[16] Paul Hemler, Thilaka Sumanaweera, Petra van den Elsen, Sandy Napel, and John Adler, "Quantified Registration Error Versus the Accuracy of Registered Surfaces for a Multimodality Surface-Based Registration System," *Proc., SPIE International Symposium on Medical Imaging 96*, Newport Beach, CA, Feb. 1996.

[17] K. H. Höhne, B. Pflesser, A. Pommert, M. Riemer, T. Schiemann, R. Schubert, and U. Tiede, "A 'Virtual Body' Model for Surgical Education and Rehearsal," *IEEE Computer*, Vol. 29, No. 1, 1996, pp. 25-31.

[18] Joseph P. Hornak and Lynn M. Fletcher, "Multispectral Image Segmentation in Magnetic Resonance Imaging," in *Digital Image Processing Methods*, ed. by Edward R. Dougherty, Marcel Dekker: New York, 1994.

[19] Ioannis Kapouleas, "Segmentation and Feature Extraction for Magnetic Resonance Brain Image Analysis," *Proc., Tenth International Conference on Pattern Recognition*, Atlantic City, NJ, June 1990, pp I-583 to I-590.

[20] A. Kaufman, "Trends in Volume Visualization and Volume Graphics," in *Scientific Visualization: Advances and Challenges*, ed. by L. Rosenblum et al., Academic Press: San Diego, 1994.

[21] P. J. Kelly, *Tumor Stereotaxis*, W. B. Saunders: Philadelphia, 1991.

[22] Carolyn Kimme, Dana Ballard, and Jack Sklansky, "Finding Circles by an Array of Accumulators," *Communications of the ACM*, Vol. 18, No. 2, Feb. 1975, pp. 120-122.

[23] Chunlin Li, Lawrence O. Hall, and Dmitry B. Goldgof, "Knowledge-Based Classification and Tissue Labeling of Magnetic Resonance Images of Human Brain," *Proc., SPIE Conf. on Biomedical Image Processing and Biomedical Visualization, SPIE Vol. 1905*, San Jose, 1993, pp. 554-565.

[24] Robert T. Malison, E. G. Miller, R. Greene, G. McCarthy, D. S. Charney, and R. B. Innis, "Computer-Assisted Coregistration of Multislice SPECT and MR Brain Images by Fixed External Fiducials," *Journal of Computer-Assisted Tomography*, Vol. 17, No. 6, 1993, pp. 952-960.

[25] Timothy S. Newman, "Issues and Directions for Medical Applications of Computer Vision and Data Visualization," *Proceedings of the IEEE Workshop on Visualization and Machine Vision*, Seattle, June 1994, p. 106.

[26] Timothy S. Newman, P. J. Flynn, and A. K. Jain, "Model-Based Classification of Quadric Surfaces," *Computer Vision, Graphics, and Image Processing: Image Understanding*, Vol. 58, No. 2, 1993, pp. 235-249.

[27] Timothy S. Newman, Ning Tang, Stephen Bacharach, and Peter Choyke, "A Volumetric Segmentation Technique for Diagnosis and Surgical Planning in Lower Torso CT Images," *Proc., Thirteenth International Conference on Pattern Recognition*, Vienna, August 1996, pp. C-553 to C-557.

[28] Terry Peters, Bruce Davey, Patric Munger, Roch Comeau, Alan Evans, and André Olivier, "Three-Dimensional Multimodal Image-Guidance for Neurosurgery," *IEEE Transactions on Medical Imaging*, Vol. 15, No. 2, 1996, pp. 121-128.

[29] Report of the NSF Advisory Panel on Graphics, Image Processing, and Workstations, 1987.

[30] Michael L. Rhodes, William V. Glenn, Yu-Ming Azzawi, Robert S. Howland, and Don L. Hibbard, "Computer Graphics and an Interactive Stereotactic System for CT-Aided Neurosurgery," *IEEE Computer Graphics and Applications*, Vol. 3, No. 5, 1983, pp. 31-38.

[31] Michael L. Rhodes and Douglas D. Robertson, "Computers in Surgery and Therapeutic Procedures," *IEEE Computer*, Vol. 29, No. 1, 1996, pp. 20-23.

[32] Joseph R. Smith and Herman F. Flanigin, "Stereotactic Neurosurgery," in *Essentials of Neurosurgery: A Guide to Clinical Practice*, edited by Marshall B. Allen and Ross Miller, McGraw-Hill: New York, 1995, pp. 463-478.

[33] Petra A. van den Elsen, Evert-Jan D. Pol, and Max A. Viergever, "Medical Image Matching—A Review with Classification," *IEEE Engineering in Medicine and Biology*, Vol. 12, No. 1, 1993, pp. 26-39.

FPGA IMPLEMENTATION OF THE XPATCH RAY TRACER

Mohammed Y. Niamat
Assistant Professor
Department of Engineering Technology

The University of Toledo
2801 W. Bancroft Street
Toledo, OH 43606

Final Report for:
Summer Faculty Research Project
Wright Laboratory Avionics Directorate (WL/AASH)
Wright Patterson Air Force Base
Dayton, Ohio

August 1996

FPGA IMPLEMENTATION OF THE XPATCH RAY TRACER

Mohammed Y. Niamat
Associate Professor
Department of Engineering Technology
The University of Toledo

Abstract

The ray tracing algorithm, often used in computer graphics for creating a 2-D picture of a 3-D world, is computationally intensive. Many attempts have been made in the past to reduce the processing time taken by this algorithm. In this research project, the ground work for implementing the algorithm onto Field Programmable Gate Array(s) is carried out. A systolic array architecture suitable for implementing the ray tracing algorithm onto FPGA(s) is proposed. VHDL codes for the different modules in the architecture are also developed. Since FPGA(s) are programmable in real-time, this method has a definite advantage over previous ones. It is hoped that the work will lead to the successful implementation of the algorithm on both the CHAMP and DRASTIC field programmable boards. Since the ray tracing algorithm is also used in the Xpatch software package used for the accurate electromagnetic scattering predictions, this work will also contribute toward reducing the processing time of Xpatch.

FPGA Implementation of the XPATCH Ray Tracer

The Ray Tracing Problem

For a given starting point R(o), and a starting direction R(d), the ray tracing problem is to determine if the ray defined by

$$R(t) = R(o) + R(d) * t, \qquad t > 0$$

hits the target. If it does, the next question that needs to be answered is "where is the hit-point ? " To accomplish these before ray tracing a target, a "tree" of the target is built. The tree contains information on how the facets are distributed in the space and is used to identify those facets that are likely to be hit by a ray. In *xpatchf*, the tree is based on an algorithm called 'binary space partitioning' (BSP).

In order to speed the computations for ray-object intersection, the object is often enclosed in a bounding volume. The bounding volume can be of different shapes. Common bounding volumes being spheres, and rectangular boxes. Kay and Kajiya [1], used the concept of slabs to form rectangular boxes as bounding volumes. A slab is simply the space between two parallel planes. The intersection of a set of slabs defines the bounding volume. The method relies on intersection of each pair of slabs by the ray, keeping track of the *near* and the *far* intersection distances. If the largest near value is greater than the smallest far value, then the ray misses the bounding volume; otherwise, it hits.

The Algorithm

One of the simplest finite bounding volumes is the intersection of two parallel planes each aligned so that their normals are in the same direction as the X, Y, and Z axes. The following algorithm determines whether the bounding box is hit by the ray , or not (miss).

Define the orthogonal box by two coordinates [2]:

box's minimum extent = B(0) = [BBP(0,0) BBP(0,1) BBP(0,2)]
box's maximum extent = B(1) = [BBP(1,0) BBP(1,1) BBP(1,2)]

The origin and direction of the ray is defined by:

R(origin) = R(0) = [rs(0) rs(1) rs(2)]
R(direction) = R(d) = [rd(0) rd(1) rd(2)]

For this application the direction vector R(direction) must be a unit vector. The ray can then be defined as the set of points on a line described by the following equation:

$$R(t) = R(0) + R(d) * t, \text{ where } t > 0$$

The above equation can be broken up into its three axes components as follows:

Along X- Axis:	$X(t) = rs(0) + rd(0) * t$
Along Y- Axis:	$Y(t) = rs(1) + rd(1) * t$
Along Z_ Axis:	$Z(t) = rs(2) + rd(2) * t$

In the above equations, 0, 1, and 2 represent the X, Y, and Z axes, respectively.

Since X(t), Y(t), and Z(t) satisfy all points on the X, Y, and Z components of the ray, they must also satisfy the points on the minimum and maximum extents of the box. Therefore,

Along X- Axis:	$BBP(0,0) = rs(0) + rd(0) * t$
	$BBP(1,0) = rs(0) + rd(0) * t$
Along Y- Axis:	$BBP(0,1) = rs(1) + rd(1) * t$
	$BBP(1,1) = rs(1) + rd(1) * t$
Along Z-axis:	$BBP(0,2) = rs(2) + rd(2) * t$
	$BBP(1,2) = rs(2) + rd(2) * t$

From the above equations, 't' can be computed for each of the above six equations. For example, for the X-axis:

$t1 = [BBP(0,0) - rs(0)] / rd(0)$, for the minimum extent of the box, and
$t2 = [BBP(1,0) - rs(0)] / rd(0)$, for the maximum extent of the box.

The ray algorithm for the intersection of the box , can be described as follows:

Set *tnear* = minus infinity, and *tfar* = plus infinity (i.e., arbitrarily large values).
For each pair of planes associated with X, Y, and Z (described here for the set of X planes) :

If the direction rd(0) is equal to zero, then the ray is parallel to the planes, so:

If the origin is not between the slabs, i.e., $rd(0) < BBP(0,0)$, or $rd(0) > BBP(1,0)$, then the ray misses the box.

Else, if the ray is not parallel to the planes, then

begin:
Calculate intersection distances of planes:

$t1 = [BBP(0,0) - rs(0)] / rd(0)$

t2 = [BBP(1,0) - rs(0)] / rd(0)

If t1 > t2, swap t1 and t2
If t1 > *tnear*, set t1 = *tnear*
If t2 < *tfar*, set t2 = *tfar*.

If *tnear* > *tfar*, box is missed
If *tfar* < 0, box is behind ray, so box is missed
end.
end of for loop.

If the box is hit, the intersection distance is equal to *tnear* and the ray's exit point is *tfar*. Figure 1(a) illustrates the case where the ray hits the box. Notice that *tnear* is less than *tfar* for this case. In Figure 1(b), *tnear* is greater than *tfar*, and therefore the ray misses the box.

The various symbols used in calculating the ray/box intersection are tabulated below.

Symbols	Meaning
BBP(0, 0)	Minimum extent of the Bounding Box Parent in the x-direction
BBP(1, 0)	Maximum extent of the Bounding Box Parent in the x-direction
BBP(0, 1)	Minimum extent of the Bounding Box Parent in the y-direction
BBP(1, 1)	Maximum extent of the Bounding Box Parent in the y-direction
BBP(0,2)	Minimum extent of the Bounding Box Parent in the z-direction
BBP(1,2)	Maximum extent of the Bounding Box Parent in the z-direction
rs(0)	X-component of the origin (starting point) of the ray
rs(1)	Y-component of the origin (starting point) of the ray
rs(2)	Z-component of the origin (starting point) of the ray
rd(0)	X-component of the directional vector of the ray
rd(1)	Y-component of the directional vector of the ray
rd(2)	Z-component of the directional vector of the ray
tnear	Distance of the nearest ray/box intersection point from the origin of ray
tfar	Distance of the farthest ray/box intersection point from the origin of ray

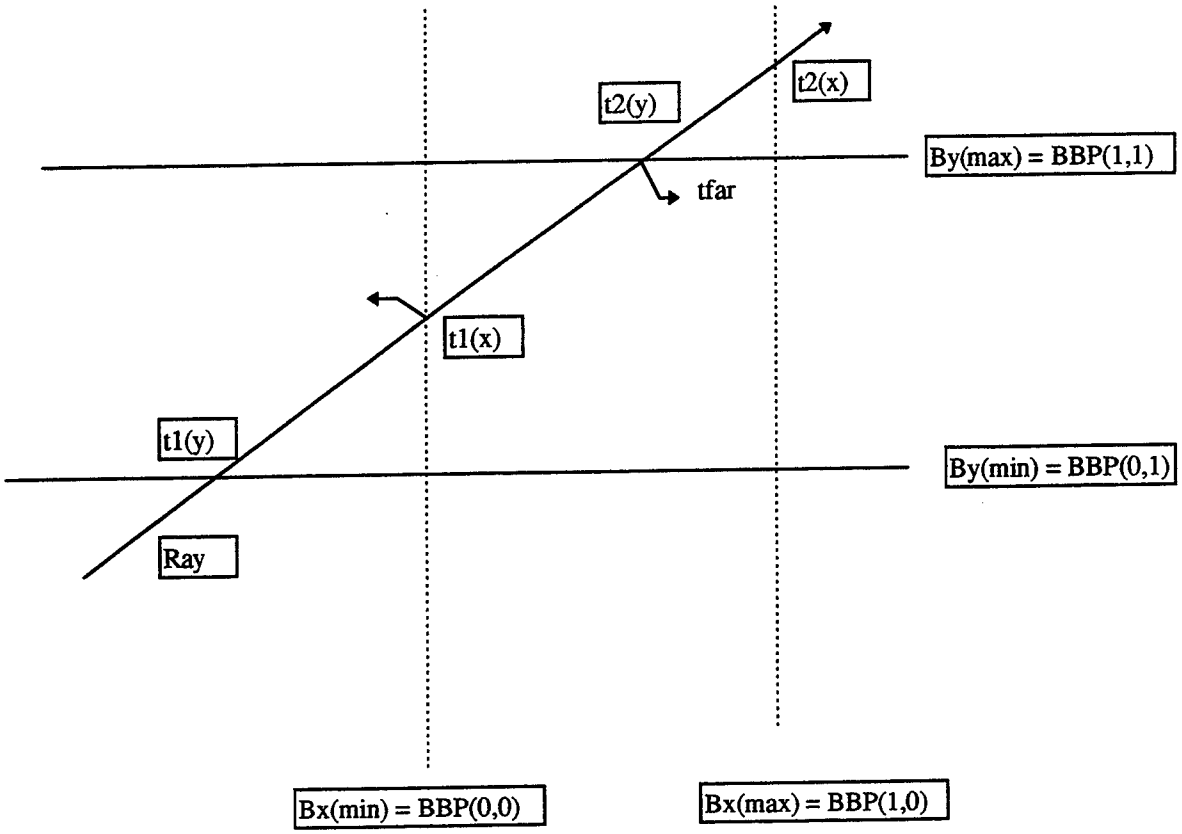


Figure 1 (a) : Ray Intersects Box

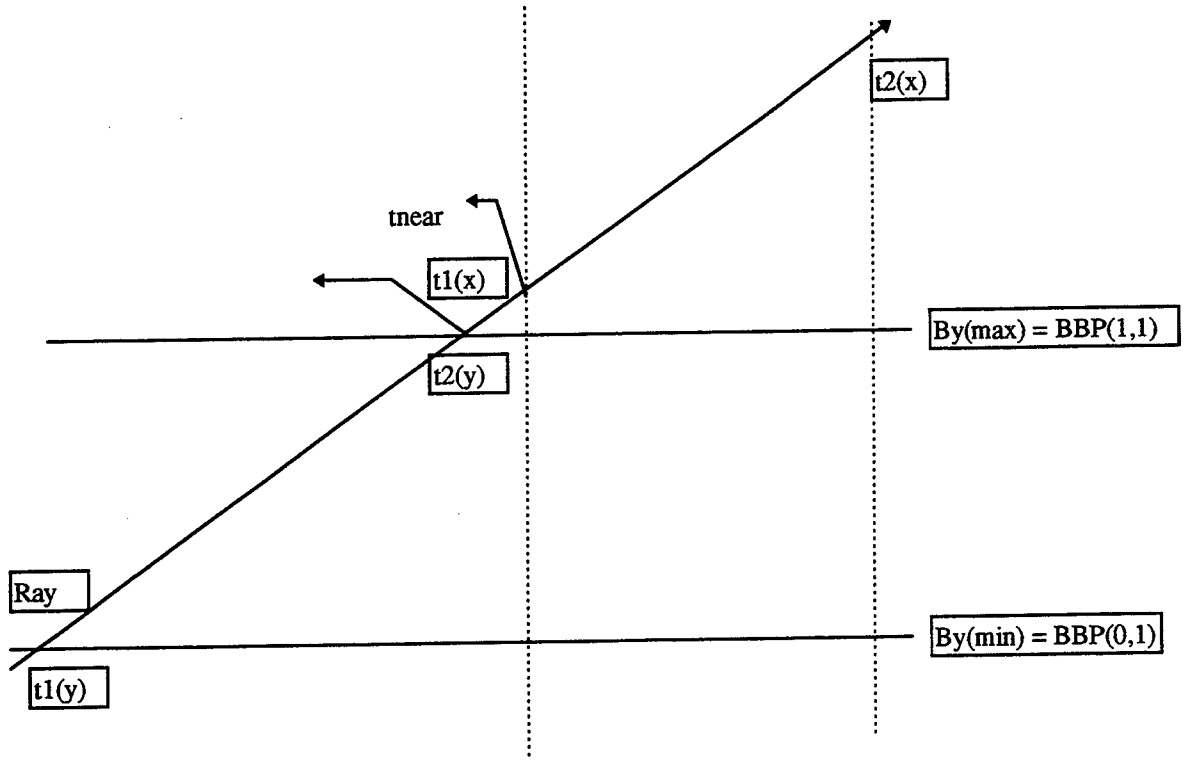


Figure 1 (b) : Ray Misses Box

The following numerical example illustrates the ray/box intersection problem.

Example

Given a ray with origin [0 4 2] and direction [0.218 - 0.436 0.873] and a box with corners:

B(minimum) = [-1 2 1]
B(maximum) = [3 3 3]

find if the ray hits the box.

Here, rs(0) = 0, rs(1) = 4, and rs(2) = 2;
rd(0) = 0.218, rd(1) = - 0.436, and rd(2) = 0.873;
BBP(0,0) = -1, BBP(0,1) = 2, and BBP(0,2) = 1,
BBP(1,0) = 3, BBP(1,1) = 3, and BBP(1,2) = 3.

The algorithm begins by looking at the X slab, defined by BBP(0,0) and BBP(1,0). The distances t1 and t2 are computed as:

t1 = [BBP(0,0) - rs(0)] / rd(0)
= [-1 - 0] / 0.218 = - 4.59

t2 = [BBP(1,0) - rs(0)] / rd(0)
= [3 - 0] / 0.218 = 13.8

After this stage, t1 is set to *tnear*, and *tfar* to t2. Therefore *tnear* becomes - 4.59, and *tfar* 13.8. Neither *tnear* > *tfar* nor *tfar* < 0, so the Y slab is examined. For the Y slab:

t1 = [BBP(0,1) - rs(1)] / rd(1)
= [2 - 4] / -0.436 = 4.59

t2 = [BBP(1,1) - rs(1)] / rd(1)
= [3 - 4] / -0.436 = 2.29

Since t1 > t2, swap these values. The new *tnear* = 2.29, and the new *tfar* = 4.59. Next, we check for the Z slab.

t1 = [BBP(0,2) - rs(2)] / rd(2)
= [1 -2] / 0.873 = -1.15

t2 = [BBP(1,2) - rs(2)] / rd(2)
= [3 - 2] / 0.873 = 1.15.

At this stage, *tnear* is not updated, and remains 2.29. However, *tfar* is updated to 1.15. Since *tnear* > *tfar*, the ray misses the box.

Design Methodology

Earlier attempts to reduce the processing time of the ray tracing algorithm can be found in references [3] to [7]. In this research project, the goal is to map the ray tracing algorithm described above onto Field Programmable Gate Arrays (FPGAs). Specifically, the intention is to use the CHAMP (Configurable Hardware Algorithm Mappable Processor) board which contains 16 Xilinx 4013 FPGAs, on-board 512 K dual-ported RAM, VME external bus, and a crossbar interconnection.

The following design steps are outlined to implement the ray tracing algorithm:

1. Parallelize the algorithm.

2. Develop systolic arrays for the algorithm.

3. Synchronize the data items maintaining nearest neighbor connection.

4. Check theoretical output and obtain performance parameters (pipelining period, etc.).

5. Develop VHDL codes for individual modules used in the overall design.

6. Interface the various modules to obtain the complete design.

7. Simulate and verify design.

8. Map design on Xilinx FPGA(s) using XACT and other CAD tools.

Proposed Architecture

The proposed architecture for finding the ray/box intersection is shown in Figure 2. It consists of nine P.E s (processing elements) connected in a systolic-type configuration. Each P.E. contains a subtractor and a multiplier. The first row computes the parameters t1 and t2 for the X- component, as defined earlier. The second and third rows compute parameters t1 and t2 for the Y and Z axes. The inputs rs(0), rs(1), and rs(2), which define the starting point or origin of the ray, are fed from the left-hand sides of row 1, row 2, and row 3, respectively. The inverses for each of the directional data items, namely, rd(0), rd(1), and rd(2) are first computed (not shown in figure). The results are then fed from the right hand side of the array as shown in the figure. The data items which define the minimum and maximum extent of the bounding box -- BBP(0,0), BBP(1,0), BBP(0,1), BBP(1,1), BBP(0,2), and BBP(1,2) are fed from the top of the array. The outputs t1 and

t2, for each of the axis, are obtained from the bottom of the cells. The data flow for the first three time snapshots are shown in figures 2(a), 2(b), and 2(c). The outputs t1 for each of the axis are fed into a comparator circuit. The maximum of the three t1 parameters is determined. This value is designated as tnear. In a similar fashion, the minimum of the three t2 parameters is determined and designated as tfar. If tnear is greater than tfar, the ray misses the bounding volume box, otherwise, the ray intersects with the box and a hit is registered.

In a real target recognition application, it is expected that a considerably large number of rays will be fired in parallel at the target as shown in Figure 3. This number could be as high as 25 million. In order to increase the speed of computations for a large number of rays, the architecture shown in Figure 2 can be duplicated. This will reduce the time for processing the rays to half. However, the number of processing elements will double to 18. If the architecture is replicated one more time, the number of cells would increase to 27. If computation of the 'plandist' parameter is neglected in the architecture, the number of cells used would be 6, 12, and 18 for the three cases under discussion.

In order to implement the proposed architecture, VHDL codes for the following components were developed.

1. Subtractor/Adder
2. Pipelined Multiplier
3. Divider, and
4. Comparator

The VHDL codes have been analyzed and tested using the Synopsys CAD package. These codes are included in Appendix A. Out of all the above modules, it was found that the divider takes the maximum time. Also, when mapping on the 4013 Xilinx FPGA, the divider occupied 500 CLB's (configuration logic blocks) out of a total of 576 CLB's available on the 4013. The clock cycles needed to perform the various operations are tabulated below:

Subtractor/Adder	1 clock cycle
Pipelined Modified Booth Multiplier	8 clock cycles
Comparator	1 cycle
Divider	28 cycles

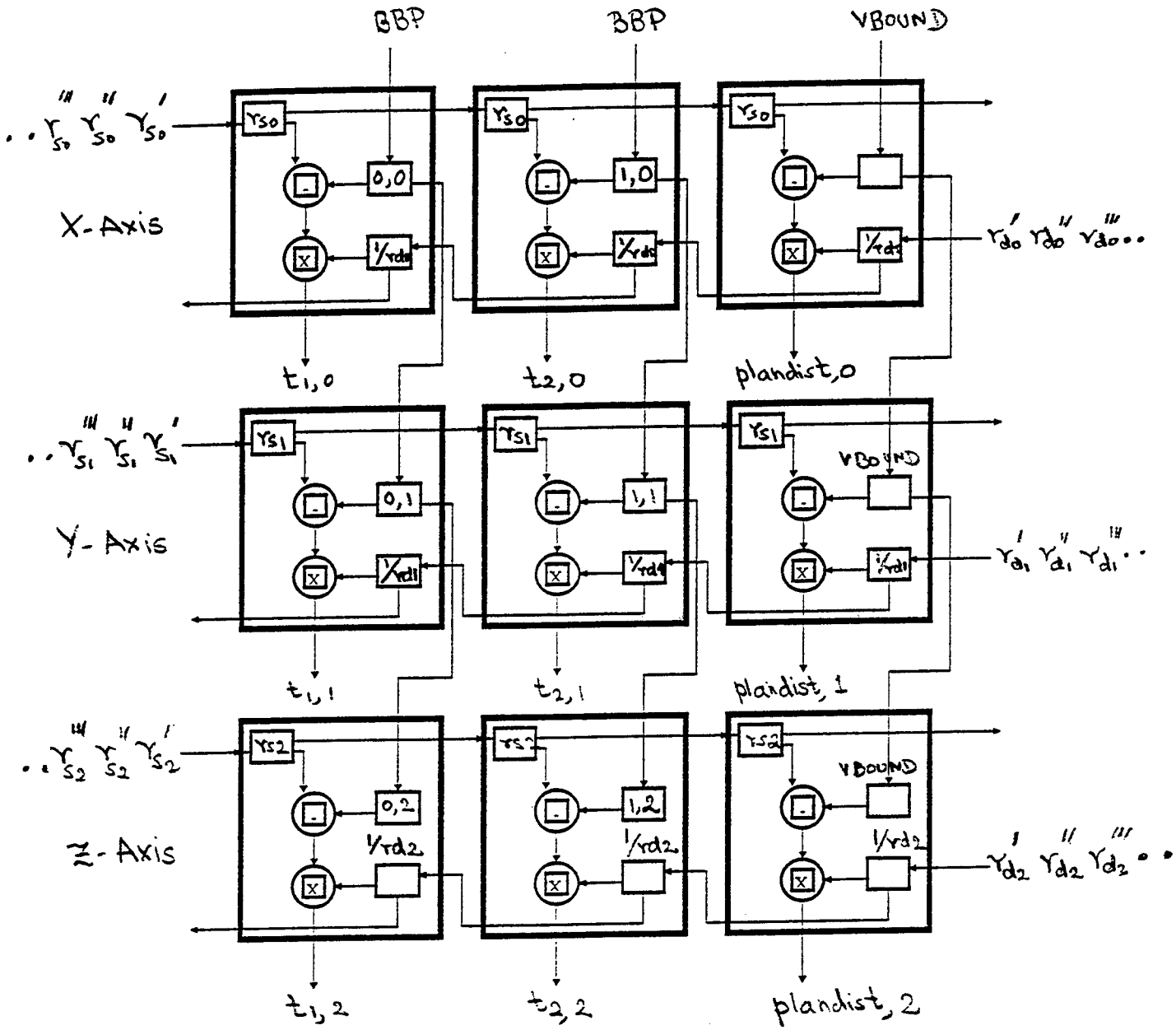


Figure 2(a): Systolic Array, Time Snapshot 1.

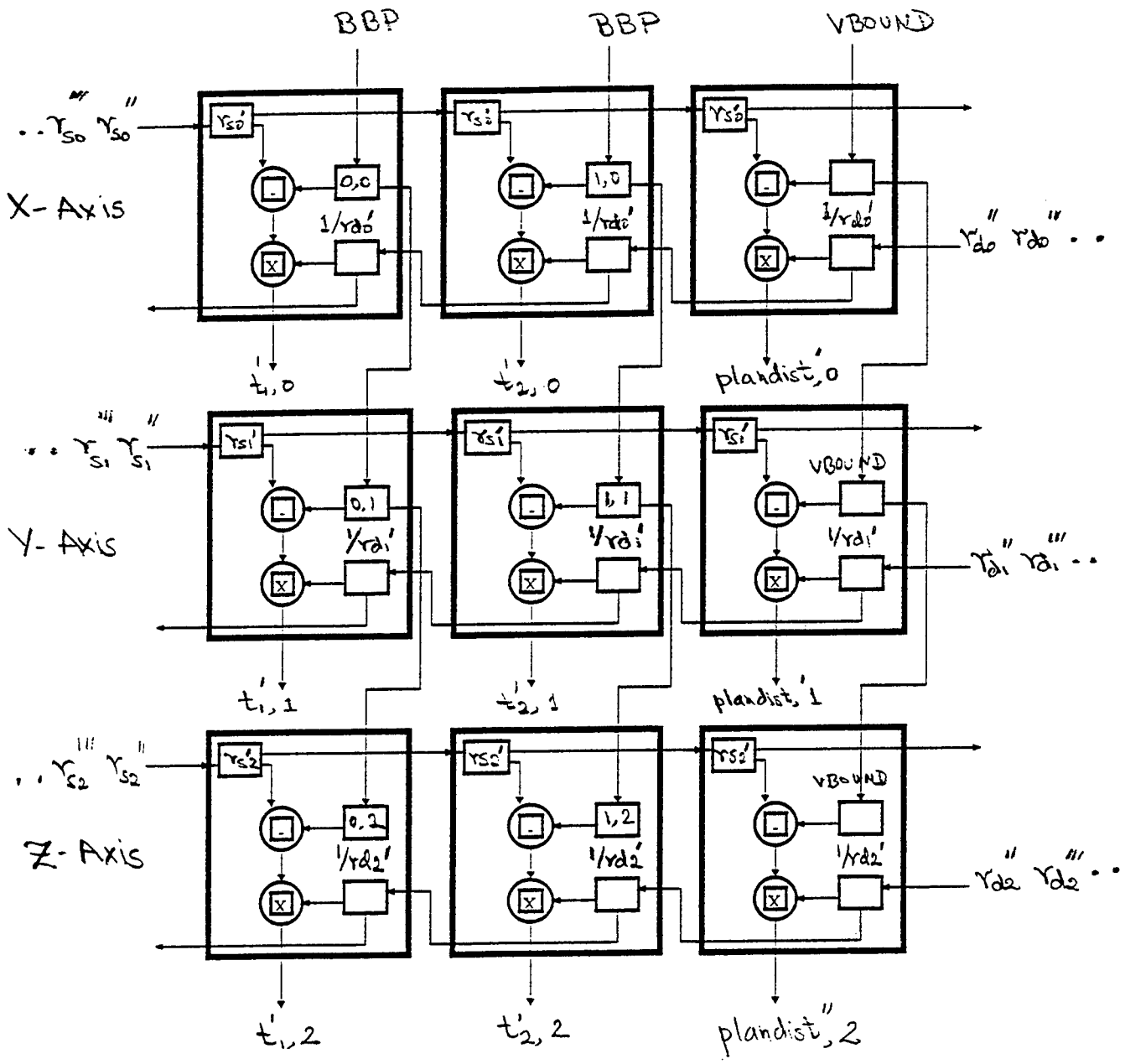


Figure 2(b): Systolic Array, Time Snapshot 2.

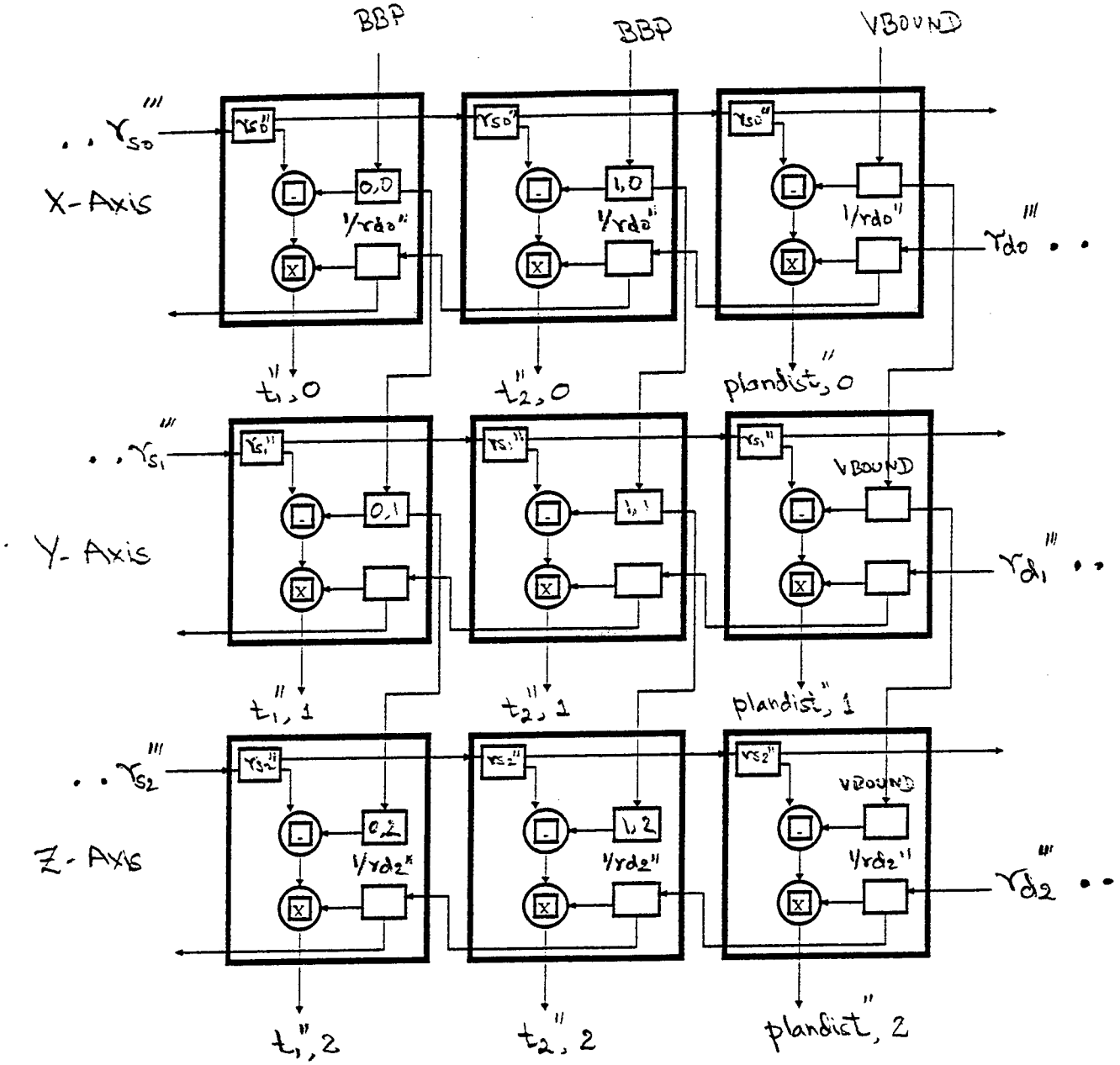


Figure 2(c): Systolic Array, Time Snapshot 3.

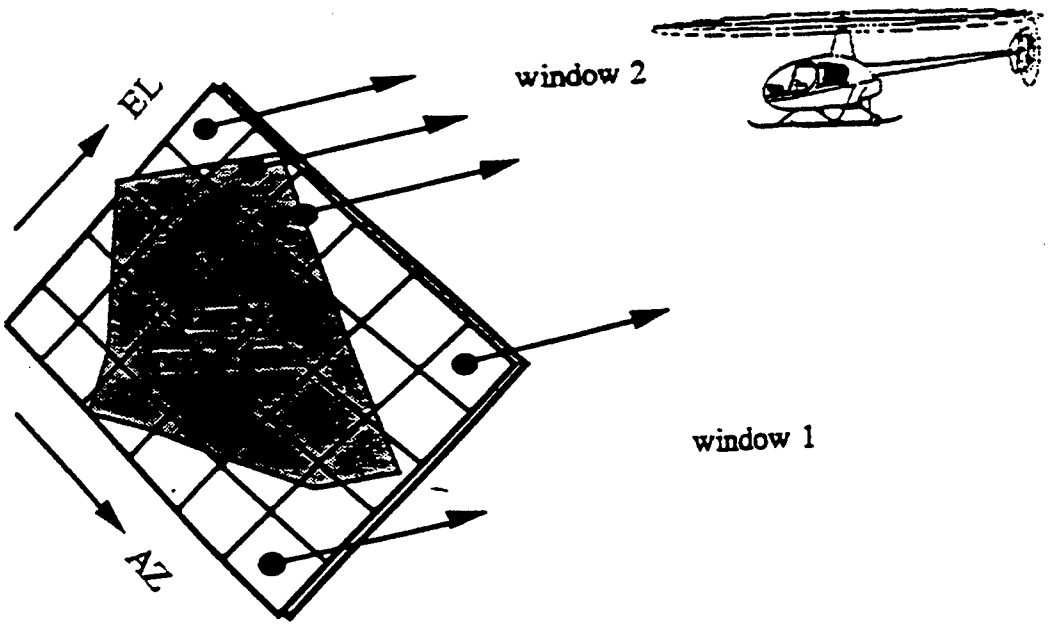


Figure 4: Rays being fired in parallel at the object.

Current Status of the Project and Future Work

A systolic array architecture, suitable for FPGA implementation of the ray tracing algorithm, has been proposed. The VHDL codes for each of the modules used in the design have been developed and tested using the Synopsys CAD tools. The next step is to map the architecture onto FPGA's, preferably on the CHAMP board. The soon to be completed DRASTIC board, with its advanced features, will also form an attractive target for implementing this design. Also, at this stage, the proposed architecture is designed for finding whether the ray intersects the bounding box, or not. However, it can easily be modified to find the co-ordinates of the hit point on the bounding box. This information can then be used to find the co-ordinates of all the successive bounces (multibounce case) of the ray until it exits the environment. Far-field scattering (of interest to people in electromagnetics) at the exit point can then be determined in real-time.

Acknowledgments

The author wishes to acknowledge the financial support given by AFOSR for making this work possible. At the Wright Laboratory Avionics Directorate, several persons assisted the author in developing this project. Foremost, the author wishes to thank Kerry Hill for defining and supervising the entire project. Without her help, this project would not have been possible. Thanks are also due to Tony Kadrovach for devoting countless hours in explaining the ray tracing algorithm, Jeff Summers for developing the VHDL codes, and Brittan Read for verifying some of the VHDL constructs. Thanks are also due to Darrell Barker for serving as the team leader in the various group meetings held to monitor the progress of the project.

References

1. T.L. Kay and J.T. Kajiya, "Ray Tracing Complex Scenes," Siggraph 86, vol. 20, No. 4, pp. 269 - 276, Aug. 1986

2. Andrew Glassner, An Introduction to Ray Tracing, Academic Press, 1989.

3. H. Ling, R. Chou, and S.W. Lee, "Shooting and Bouncing Rays: Calculating the RCS of an Arbitrarily Shaped Cavity," IEE Trans. On Antennas and Propagation, vol. 37, 194 - 205, 1988.

4. Users Manual for XPATCH, Demaco, 100 Trade Center Drive, Champaign, IL, 1993.

5. E.M. Miller, D.J. Andersh, and A.J. Terzzouli Jr., "Facetization Level and the Effect on XPATCH Predictions," Proceedings of the Ninth Annual Review of Progress in Applied Computational Electromagnetics, Naval Postgraduate School, Monterey, CA, March 22 - 26, 1993, pp. 610 - 617.

6. J.T. Kajiya, "Siggraph 83 Tutorial on Ray Tracing," Siggraph 83 State of the Art in Image Synthesis Course Notes, July 1983.

7. G.A. Deschamps, "Ray techniques in electromagnetics," Proc. IEEE, vol. 60, pp. 1022 - 1035, Sept 1972.

THE DEVELOPMENT OF NEW LEARNING ALGORITHMS

James L. Noyes
Professor of Computer Science
Dept. of Mathematics & Computer Science

Wittenberg University
Ward Street and University Avenue
Springfield, OH 45501

Final Report for:
Summer Faculty Research Program
Wright Laboratory
Avionics Directorate
WL/AACF

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

and

Wright Laboratory

August 1996

THE DEVELOPMENT OF NEW LEARNING ALGORITHMS

James L. Noyes
Professor of Computer Science
Dept. Of Mathematics and Computer Science
Wittenberg University

Abstract

Computational learning techniques serve as powerful methods of solving a variety of problems. This report investigates a new wire fitting function that can be used to facilitate computational learning. The wire fitting function is first evaluated and then applied to the development of both memory-based and gradient-based (neural network) algorithms to solve a problem involving optimal control. The memory-based algorithms are implemented in code.

1. Introduction. This project involved two phases. The first phase was to implement and evaluate the accuracy of a new type of approximation function that used a technique called "wire fitting." The second phase was to write a small simulation program, combine it with the wire fitting function implementation and a neural network implementation, and add the additional code necessary to produce a mechanism to learn how to make optimal decisions through experience.

2. Phase 1. This phase involved studying two functions defined by Equations 1 and 2 in Baird and Klopf [Bair93], implementing those functions in one dimension, choosing some test functions, and evaluating the accuracy of Equation 2 which involves the technique of wire fitting.

2.1 Implementation. Two function implementations were done. The first was done primarily to assess the qualitative capability of the functions. It was done in *Mathematica*. Here the behavior of functions and their errors could be readily graphed and investigated. The second implementation was done in *Turbo Pascal*. The purpose here was to determine the quantitative nature of this code. Here the numeric precision could be investigated within an application setting.

2.2 Evaluation Methodology. Five test functions were investigated using *Mathematica* for the qualitative and quantitative parts and *Turbo Pascal* for a check of the quantitative part. The qualitative part investigated the behavior of the following within the domain of interest: f(u), g(u), h(u), f(u) - h(u), g(u) - h(u). Here h(u) is the exact benchmark test function defined in the domain given by $a \leq u \leq b$. The quantitative part examined the maximum absolute and relative errors within the appropriate domain. The following *Mathematica* functions were to evaluate and plot the indicated f(u) and g(u) = f(x,u) functions, respectively defined by Equation 1 and Equation 2 in [Bair93]:

```
f[v_] := Module[{num = 0, den = 0, temp = 0, i = 0,
                 defined = True,
                 maxy = Max[Table[y[k], {k,1,m}]]},
                While[defined && i < m,
                   i = i + 1;
                   temp = Abs[v - u[i]] + maxy - y[i];
                   If[temp == 0,
                      defined = False,
                      num = num + y[i]/temp;
                      den = den + 1/temp]];
                If[defined, num/den, maxy]]
```

```
g[v_] := Module[{num = 0, den = 0, temp = 0, i = 0,
                 defined = True,
                 maxy = Max[Table[y[k], {k,1,m}]]},
                While[defined && i < m,
                   i = i + 1;
                   temp = (v - u[i])^2 + c[i]*(maxy - y[i]);
                   If[temp == 0,
                      defined = False,
                      num = num + y[i]/temp;
                      den = den + 1/temp]];
                If[defined, num/den, y[i]]]
```

The *Turbo Pascal* code is not shown here, due to its significantly greater length.

Note that the **g[v_]** formulation is based upon a one-dimensional version of Equation 2, namely

$$(1) \quad g(v) = \frac{\sum_{i=1}^{n} y_i / [(v-u_i)^2 + (Max_k[y_k] - y_i)]}{\sum_{i=1}^{n} 1 / [(v-u_i)^2 + (Max_k[y_k] - y_i)]}$$

Each test function h(u) below was chosen for its special characteristics:

The function **h(u) = 1** is a constant function with ideal properties, well-normalized and no numerical error potential (underflow, overflow, roundoff) in any domain.

If $h(u) = y_{max} = y_i$ for i=1,2,...,m (m is the number of data points) then it can be easily shown that $f(u) = y_{max}$ and $g(u) = y_{max}$. This means, of course, that no error should result from either f(u) or g(u) when the tests were run based upon the data generated by this function. This was readily verified.

The function **h(u) = 1000u + 5.5, $-3 \leq u \leq 3$** is a linear function with a large slope. The maximum occurs at the right-hand endpoint b = 3 with $y_{max} = 3005.5$. Although for the given domain there is no numerical error potential, this function is not always easy to approximate by f(u) and g(u) because of the steepness.

For the f(u) approximation tests, the maximum (absolute) error always occurs at the left-hand endpoint a = -3. For m = 2, this error was 5994.01 and it

monotonically decreased (for increasing m) so that for m = 50 this error was 5759.03. Notice that this error is at a data point and was larger than the errors at any of the non-data points. The maximum relative error occurs near the midpoint of the interval in all cases (because the absolute value of h(u) is smallest there). It also monotonically decreased from 1201.44 when m = 2 to 1149.59 when m = 50. The function f(u) is not a good approximator even when h(u) is linear.

The errors in the g(u) approximation at the data points when all $c_i = 0$ were all 0 since g(u) then behaves as an interpolating function. For all $c_i = 0$ the maximum error for m = 2 was 900.85 while the maximum absolute error when m = 50 was 58.44. The corresponding maximum relative error was around 2.5 for all values of m investigated. As the $c_i > 0$ values approached 1, the error for a given number of points, m, increased monotonically. Of course, the g(u) approximation is no longer exact at the data points. For $c_i = 1/4$ the maximum error was 5859.38 when m = 2 and was only down to 4693.71 when m = 50. For $c_i = 1$, the corresponding values were 5964.22 and 5295.40. In general, these errors got smaller as m increased. The same is true for the relative error which went from around 1200 when m = 2 down to around 900 to 1050 when m = 50.

For f(u) and for all g(u) when $c_i > 0$, these functions were relatively linear with large errors. For g(u) when $c_i = 0$, the errors tended to be small with the difference function g(u) - h(u) oscillating more frequently as m increased (it looks like a "bent spring" when m = 50).

The function $\mathbf{h(u) = 10\log_{10}(u),\ 1/10 \le u \le 10}$ is a well-behaved nonlinear function whose maximum occurs at the right-hand endpoint point b = 10 with y_{max} = h(b) = 10. The minimum is at the left-hand endpoint with h(a) = -10. Notice that as a approaches zero, h(u) becomes increasingly difficult to approximate for u in the left end of this domain.

For f(u) the maximum absolute error went from 13.378 for m = 2 up to 16.155 for m = 50. For g(u) when $c_i = 0$, there was no error at the data points and the error at the non-data points went from 12.125 from m = 2 down to 0.951 at m = 50. Again, there was a very close fit with a high degree of oscillation. For g(u) when $c_i = 1$, the absolute error went from 7.714 for m = 2 to 13.692 for m = 50.

The function $\mathbf{h(u) = \sin(u)\cos(2u),\ 0 \le u \le 2\pi}$ was taken from [Smag94]. It has an interesting oscillation pattern with three local minima and three local maxima in the above domain. The global maximum is $y_{max} = +1$ at approximately u = 4.71239, while the global minimum is -1 at approximately u = 1.57080. Oscillatory functions tend to be more difficult to approximate. However this function is always bounded within any domain, so there is no overflow potential although roundoff can occur.

For f(u) the maximum absolute error slightly increased as m went from 2 to 50, with this error always less than 1.0. For g(u) when $c_i = 0$, it got closer and closer to h(u) by oscillating around it with zero error at the data points (as usual) and going from around 1.0 at m = 2 to around 0.071 at m = 50. When $c_i = 1$, the maximum error at m = 50 was 0.840 at one of the data points and 0.844 over the interval.

The function $\mathbf{h(u) = \tan(u),\ 0 \le u \le \pi}$ was also taken from [Smag94]. It has a discontinuity at $\pi/2$ where approaching from the left yields positive infinity, while approaching from the right yields negative infinity. Because of this discontinuity it violates the assumptions for f(u) and g(u), but nevertheless is a good test function. The left segment has a minimum of 0 at a = 0 with an undefined maximum at $\pi/2$, while the right segment has an undefined minimum at $\pi/2$ with a maximum of 0 at b = at π. This function becomes quite difficult to approximate at the discontinuity and also can easily produce roundoff error and overflow there. The error approaches infinity as u approaches $\pi/2$.

The functions f(u) and g(u) essentially behaved in the same manner as with the previous test function, except around the finite discontinuity as expected. For m = 50, the maximum error of f(u) was under 40 as long as u was not closer than 0.5 from the discontinuity, while the maximum error of g(u) when $c_i = 0$ was under 0.3 as long as u was not closer than 0.5 from the discontinuity.

For each function h(u) above, six different equally-spaced data sets were generated (m = 2, m = 3, m = 5, m = 10, m = 20, m = 50). For each of the six cases, f(u) from Equation 1 of [Bair93] was studied along with $\mathbf{f(x,u)}$ from Equation 2 of [Bair93] with n = 1, this function is called g(u) here. Five different values for all of the m c_i values were used ($c_i = 0$, $c_i = 1/4$, $c_i = 1/2$, $c_i = 3/4$, $c_i = 1$). In all, 180 cases were investigated. There was no time to investigate non-equally spaced points.

2.3 Findings. The maximum of the rational f(u) function was always equal to the maximum y_i value as specified by the designers. However f(u) was *not* found to be a good approximator for any case other than h(u) = 1. The maximum of the rational g(u) function was also equal to the maximum y_i value as specified for *any* values of c_i. The function g(u) was found to be a good approximator only when all c_i = 0. It can be easily shown that in this case g(u) is always interpolatory. The approximation accuracy improved as the number m of data points increased, with g(u) oscillating around h(u) as expected. For a fixed m, when all $c_i > 0$, the maximum errors of g(u) became smaller as c_i approaches 0, but the oscillation of g(u) increases.

3. Phase 2. This phase was intended to place the wire fitting function as described in Figure 2 of [Bair93] within a practical setting that involves unsupervised learning. Two types of learning algorithms may be developed which utilize this function: memory-based, employing statistical estimating equations (Appendix A), and gradient-based, employing neural networks (Appendix B).

This phase involved writing two (training and application) algorithms together with corresponding prototype simulation codes. Both of these used wire fitting. One used a neural network to feed it while the other used memory-based learning equations. In both cases the purpose was to produce a mechanism that learns how to make optimal decisions through experience.

The remaining discussion in this paper uses the following notation:

n - the number of wires used
p - the total number of state and control vectors used
q - the number of components of the control (action) vector
s - the number of components of the state vector
t - the current trajectory "time" indicator
$\mathbf{u}$ - the current control (action) vector (also written $\mathbf{u}^t$)
u_k - the kth component of the control (action) vector
$\mathbf{x}$ - the current state vector (also written $\mathbf{x}^t$)
x_L - the Lth component of the state vector

To evaluate the function $f(\mathbf{x},\mathbf{u})$ defined by Equation 2, Equation 3 is also needed. These are shown below. Equation 2 is the wire fitting function.

$$(2)\quad f(\boldsymbol{x},\boldsymbol{u}) = \lim_{\epsilon\to 0^+} \frac{\sum_{i=1}^{n} y_i(\boldsymbol{x})\,[\|\boldsymbol{u}-\hat{\boldsymbol{u}}_i(\boldsymbol{x})\|^2 + c_i(y_{max}(\boldsymbol{x})-y_i(\boldsymbol{x})) + \epsilon]^{-1}}{\sum_{i=1}^{n} [\|\boldsymbol{u}-\hat{\boldsymbol{u}}_i(\boldsymbol{x})\|^2 + c_i(y_{max}(\boldsymbol{x})-y_i(\boldsymbol{x})) + \epsilon]^{-1}}$$

$$(3)\quad y_{max}(\boldsymbol{x}) \equiv \max_i\{y_i(\boldsymbol{x})\} = \max_u f(\boldsymbol{x},\boldsymbol{u})$$

3.1 The Memory-Based Method. For a memory-based function approximation system, the stored information will consist of a set of triplets $(\mathbf{x}^t,\mathbf{u}^t,E^t)$. Each of the p data points $(\mathbf{x}^t,\mathbf{u}^t,E^t)$ is projected into the hyperplane of the current state $\mathbf{x}$, to give a projected point $(\mathbf{x},\mathbf{u}^t,E^t)$. All of the data points, not just the closest ones have an effect on the wire associated with each projected point.

At time t, the "trajectory" is in vector state $\mathbf{x}^t$. If the control action vector $\mathbf{u}^t$ is performed, the system will output the scalar function $f(\mathbf{x}^t,\mathbf{u}^t)$. The reinforcement learning algorithm then calculates a scalar estimate E^t of what $f(\mathbf{x}^t,\mathbf{u}^t)$ should have been, based upon the results of performing action $\mathbf{u}^t$ in state $\mathbf{x}^t$. Once the estimate has been calculated, the triplet can be stored. In a reinforcement learning system, the function $f(\mathbf{x},\mathbf{u})$ typically represents the utility of performing action $\mathbf{u}$ in state $\mathbf{x}$, so the $\mathbf{u}$ that maximizes $f(\mathbf{x},\mathbf{u})$ is the optimal action to perform in state $\mathbf{x}$. Wire fitting is a function approximation method designed to facilitate finding the maximum of a function $f(\mathbf{x},\mathbf{u})$ for a given $\mathbf{x}$.

The vector and scalar functions $\hat{\mathbf{u}}_i(\mathbf{x})$ and $y_i(\mathbf{x})$ can be calculated from the set of stored memories by using Equations 4 - 10 of [Bair93]. These are given as follows:

$$(4)\quad d_{it} = \frac{1}{\sum_{L=1}^{s}(x_L-x_{tL})^2 + \sum_{k=1}^{a}(u_{ik}-u_{tk})^2 + \beta} \qquad i=1,2,\ldots,n;\ t=1,2,\ldots,p$$

$$(5)\quad \overline{y_i} = \frac{\sum_{t=1}^{p} y_t d_{it}}{\sum_{t=1}^{p} d_{it}} \qquad i=1,2,\ldots n$$

$$(6)\quad \overline{u_{ik}} = \frac{\sum_{t=1}^{p} u_{tk} d_{it}}{\sum_{t=1}^{p} d_{it}} \quad i=1,2,\ldots,n;\ k=1,2,\ldots,q$$

$$(7)\quad \overline{yu_{ik}} = \frac{\sum_{t=1}^{p} y_t u_{tk} d_{it}}{\sum_{t=1}^{p} d_{it}} \quad i=1,2,\ldots,n;\ k=1,2,\ldots,q$$

$$(8)\quad m_{ik} = \frac{\overline{yu_{ik}} - (\overline{y_i})(\overline{u_{ik}})}{\overline{u_{ik}} - (\overline{yu_{ik}})^2} \quad i=1,2,\ldots,n;\ k=1,2,\ldots,q$$

$$(9)\quad \hat{u}_{ik}(x) = \overline{u_{ik}} + \alpha m_{ik} \quad i=1,2,\ldots,n;\ k=1,2,\ldots,q$$

$$(10)\quad y_i(x) = \overline{y_i} + \alpha \sum_{k=1}^{q} m_{ik}^2 \quad i=1,2,\ldots,n$$

If old memories are eventually lost, perhaps because of a finite-sized memory set, then the $\hat{\mathbf{u}}_i(\mathbf{x})$ and $y_i(\mathbf{x})$ functions would be expected to improve with experience, yielding memory-based learning. If there are n functions $\hat{\mathbf{u}}_i(\mathbf{x})$, $y_i(\mathbf{x})$, for i=1,2,...,n then every state will intersect all n of the wires.

Equations 2 and 3 are in the *Interpolation Function* block. Equations 4 - 10 are in the *Learned Function* block, which has a total of (s+1)n outputs and are combined with an external action control **u**. As inputs to the *Interpolation Function* block. For a given state **x** the set of vector pairs $(\hat{\mathbf{u}}_i(\mathbf{x}), y_i(\mathbf{x}))$ are interpolated to give $f(\mathbf{x},\mathbf{u})$. Equation 2 defines $f(\mathbf{x},\mathbf{u})$ for a particular **u** to be a weighted average of the $y_i(\mathbf{x})$ values. If **u** is "near" a particular $\hat{\mathbf{u}}_i(\mathbf{x})$, then the corresponding $y_i(\mathbf{x})$ is given more weight.

From this one can compute the estimate E^t from:

$$(11)\quad E^t \leftarrow E^t + r[(R+\delta E^{t+1}) - E^t]$$

Specifically, this entire computation starts by generating a set of random vectors $\mathbf{x}^t$ and $\mathbf{u}^t$ together with scalars $E^t = y_t$ for $t=1,2,\ldots,p_0$ (where p_0 is input). Then set $p = p_0$. This allows Equations 4 - 10 to work. (After $p = 2p_0$ of these have been generated, the first p_0 can be replaced by the second, again resetting $p = p_0$.)

First, compute for n wires and p *stored* data points (where p increases at each iteration), Equation 4:

$$d_{it} = 1/[\sum(x_L - x_{tL})^2 + \sum(u_{ik} - u_{tk})^2 + \beta] \quad \text{for } i=1,\ldots,n; \quad t=1,\ldots p$$

where the first sum ranges from L=1 to L=s (the number of state variables), the second sum ranges from k=1 to k=q (the number of control variables), and $\beta > 0$ prevents division by zero. This means that the effect of the t<u>th</u> data point on the i<u>th</u> wire is inversely proportional to its distance from the projection. In general, the denominator can be expanded as:

$$(x_1-x_{t1})^2 + (x_2-x_{t2})^2 + \ldots + (x_s-x_{ts})^2 + (u_{i1}-u_{t1})^2 + (u_{i2}-u_{t2})^2 + \ldots + (u_{iq}-u_{tq})^2 + \beta$$

Here at time t, the u_{ik} values for $k=1,2,\ldots,q$ are produced as the i<u>th</u> wire $\mathbf{u}^{out}$ components of the NN output (along with $\mathbf{y}^{out}$) and the u_{tk} values for $k=1,2,\ldots,q$ are produced as the corresponding random $\mathbf{u}^t$ components (notice that in d_{it}, the first index is different on the action variables). Both of these are saved. Specifically, for s = 2 and q = 2, the above equation becomes:

$$d_{it} = 1/[(x_1-x_{t1})^2 + (x_2-x_{t2})^2 + (u_{i1}-u_{t1})^2 + (u_{i2}-u_{t2})^2 + \beta] \quad \text{for } i=1,\ldots,n; \; t=1,\ldots p.$$

Equations 5 - 8 perform a weighted linear regression. Equations 9 and 10 place the i<u>th</u> wire $(\hat{\mathbf{u}}_i(\mathbf{x}), y_i(\mathbf{x}))$ near the projected point, slightly uphill in the direction found by the weighted linear regression. That is, the output consists of $y_i(\mathbf{x})$ elements with $i=1,2,\ldots,n$, together with n control vectors $\hat{\mathbf{u}}_i(\mathbf{x})$.

For a given state $\mathbf{x}$, the functions $\hat{\mathbf{u}}_i(\mathbf{x})$ and $y_i(\mathbf{x})$ are defined by Equations 4 - 10. If there are n *wires*, then there will be a wire associated with each of the *data points* nearest to state $\mathbf{x}$ in Euclidean distance. The i<u>th</u> wire will not necessarily go through the i<u>th</u> data point, but $\hat{\mathbf{u}}_i(\mathbf{x})$ will typically be fairly close to the $\mathbf{u}$ vector of the associated data point.

3.2 The Gradient-Based Method. The *Learned Function* block in Figure 2 was implemented with a feed-forward neural network with two hidden layers containing h units and one unit (the output unit associated with f(**x**,**u**)). This is a multi-layer feed-forward network with the topology: s-h-(q+1)n-1.

In order to determine the weight improvement, the (partial) derivatives of Equation 2 are needed. These may be computed by the following equations [Bair96]:

$$(12) \quad d_i = \|\mathbf{u} - \hat{\mathbf{u}}_i\|^2 + c_i(y_{max} - y_i) + \epsilon \quad i=1,2,\ldots,n$$

$$(13) \quad S = \frac{1}{\sum_{i=1}^{n} \frac{1}{d_i}}$$

$$(14) \quad f(\mathbf{x},\mathbf{u}) = S\sum_{i=1}^{n} \frac{y_i}{d_i}$$

$$(15) \quad \frac{\partial y_{max}}{\partial y_j} = \begin{matrix} 1 \ \ if \ \ y_{max} = y_j \\ 0 \ \ otherwise \end{matrix}$$

$$(16) \quad \frac{\partial y_i}{\partial y_j} = \begin{matrix} 1 \ \ if \ \ i = j \\ 0 \ \ otherwise \end{matrix}$$

$$(17) \quad \frac{\partial d_i}{\partial y_j} = c_i[\partial y_{max}/\partial y_j - \partial y_i/\partial y_j]$$

$$(18)\quad \frac{\partial d_i}{\partial \hat{u}_{jk}} = -2(u_k - \hat{u}_{jk})$$

$$(19)\quad \frac{\partial f(\mathbf{x},\mathbf{u})}{\partial \hat{u}_{jk}} = -S\sum_{i=1}^{n} y_i(\partial d_i/\partial \hat{u}_k)/d_i^2 + S^2[\sum_{i=1}^{n}(\partial d_i/\partial \hat{u}_{jk})/d_i^2][\sum_{i=1}^{n} y_i/d_i]$$

$$(20)\quad \frac{\partial f(\mathbf{x},\mathbf{u})}{\partial y_j} = -S\sum_{i=1}^{n}[y_i(\partial d_i/\partial y_j)/d_i^2 - (\partial y_i/\partial y_j)/d_i] + S^2[\sum_{i=1}^{n}(\partial d_i/\partial y_j)/d_i^2][\sum_{i=1}^{n} y_i/d_i]$$

$$(21)\quad \Delta w = r[(R+\delta y_{max}) - f^t]\{\sum_{j=1}^{n}[(\partial f/\partial y_j)(\partial y_j/\partial w) + \sum_{k=1}^{q}(\partial f/\partial \hat{u}_{jk})(\partial \hat{u}_{jk}/\partial w)]\}$$

In Equations 12 - 21 above, u_k is the kth element of the current control vector $\mathbf{u}$. The scalar y_j is the jth wire output, and $\hat{u}_{jk}$ is the kth element of the jth wire vector, all $q + (q+1)n$ of these are output from the next-to-last layer of the neural net. These values, plus q additional values, u_k for $k=1,2,\ldots,q$ are inputs into the last layer of the net. In Equation 21, r is the learning rate, R is the reward, δ is the discount factor, $y_{max} = \text{Max}\{ y_j, j=1,2,\ldots,n \}$ and $f^t = f(\mathbf{x}^t,\mathbf{u}^t)$.

3.3 Simulation Scenario. The practical setting chosen was called a "golf course scenario" because the application involved a simple, but important situation whereby a 2-dimensional state vector $\mathbf{x} = (x_1,x_2)^T$ representing the position of a ball on a 2-D golf course could be controlled or hit in an "optimal" direction (which can be thought of as an angle and distance) where the control vector is also 2-dimensional in this case and can be denoted as $\mathbf{u} = (u_1,u_2)^T$ where $u_1 = \theta$ and $u_2 = \rho$. Hence $s = 2$ and $q = 2$. Note that in both cases the superscript denotes transpose since unadorned vectors denote column vectors. All distance units will be in yards.

For the memory-based approach, Equations 4 - 10 are used. The values are saved at each time point t as x_{tL} for L=1,2,...,s and u_{tk} for k=1,2,...,q. After training is complete, the triplets ($\mathbf{x}^t$, $\mathbf{u}^t$, E^t) for t=1,2,...,p are saved in a file for later use.

For the gradient-based approach, this is a multi-layer feed-forward network with the topology 2-h-(q+1)n-1 which becomes: 2-h-3n-1. The $\mathbf{x}$ input to the network is the 2-dimensional vector representing the coordinates of a point on the golf course. The output of the network consists of y_i elements for i=1,2,...,n, together with the control vector values u_{ik}, each of the form (θ,ρ) for k=1,2 representing the angle (in radians) and the distance (in yards) that a golf ball is hit. After training is complete, the weights and biases are saved in a file for later use.

For both approaches, it will be assumed that the number of wires $n \leq 5$.

The idea is to generate a random position on the golf course. The goal of the game is to get the golf ball onto the green (of a given size) without getting into one of the sand traps (of a given number and size) or go out of bounds. The goal of the process is to play enough games so that the system, through the use of either a neural network or memory-based equations, learns the optimal angle and distance to hit the ball from this random position so that it will reach the green without going into a sand trap. The idea is to do this even if the ball has never been in that position before.

3.4 Implementation. The golf course simulation was done with two main procedures.

The first one defines the course (the size, placement and dimension of the green, together with the placement and dimensions of each sand trap). No green or sand trap will overlap. This procedure will be called once.

The second procedure will be called many times. Each time a state vector $\mathbf{x}$ will be accepted, and an appropriate value will be returned. The reward will be +W if $\mathbf{x}$ is on the green and it will be -W if $\mathbf{x}$ is in a sand trap. Otherwise it will be $f(\mathbf{x},\mathbf{u})$.

There was only time to code one of the two main methods, so the memory-based method was selected, because the equations were available for it first. Some code had already been written for the gradient-based neural network as well.

The memory-based code of approximately 30 pages in length is complete and written in Turbo Pascal. It contains in-code documentation (comment statements) for each procedure, including purpose, precondition, and postcondition, together with documenting comments for each constant, variable, and procedure parameter. It allows many training options to be chosen and permits the storage of up to 1500 unique triplets of information. It also contains many trace output statements so that the user can carefully investigate how Equations 2 - 11 interact during the training process. It also allows the triplets to be stored in a user-named disk file so that the game can be played at a later time. Although the code has been debugged through the use of many types of inputs, there was not enough time to do a thorough evaluation. However, this program should serve as a useful tool in the investigation of this type of memory-based learning.

3.5 Evaluation. The code is ready to allow a user to experiment with the parameters associated with the training algorithm. These parameters include: the number of wires to use, the c-coefficients of Equation 2, the learning rate, α and β, the number of triplets to keep before over-writing, the convergence value, and the discount factor. Based upon the preliminary evaluation done thus far, the number of triplets to keep and the learning rate are both significant. It should be noted that the use of double precision should be considered in order to facilitate the training process.

Here are two possible evaluation strategies that can be used with any training methodology that is applied to the 2-D golf course scenario:

a. One test is that starting at any user-chosen point, the trajectory should never cause the ball to go into a sand trap.

b. A more methodical type of evaluation is to generate a grid of golf course points and to execute a calculation (e.g., by using the memory-based equations or by using a forward pass of a neural network) to determine the next point (distance and direction of each) and display these.

4. Acknowledgments. Thanks go to Leemon Baird, who developed nearly all of the equations in this report and who took the time to answer many questions about them. Special thanks go to Jim Morgan, my advisor, who spent countless hours assisting me and working closely to develop and refine the algorithms in this report.

==========

References

Bair93 Leemon C. Baird III and A. Harry Klopf, *Reinforcement Learning with High-Dimensional, Continuous Actions*, Wright Laboratory Avionics Directorate Report WL-TR-93-1147, Wright Patterson AFB, OH, 45433-7409, November 1993.

Bair96 Personal communication from Leemon Baird to Jim Morgan, July 1996.

Smag94 P. Patrick van der Smagt, "Minimization Methods for Training Feedforward Neural Networks," *Neural Networks*, Vol. 7, No. 1, pp. 1-11, 1994.

==========

Appendix A

Memory-Based Learning Algorithms

The general description of a memory-based *training process* is as follows (see the disk file *mwiresim.pas* for detailed Pascal source code for both the training and application):

Step 0: Initialize the process.

Initialize the golf course parameters (size of the course, placement and size of the green, number, placement and size of sand traps, etc.).

Initialize the learning algorithm. Set the number of state variables $s = 2$ and $q = 2$ as the number of control variables. Initialize the $\mathbf{c}$ vector for computing $f(\mathbf{x},\mathbf{u})$. Initialize $\alpha \geq 0$ and $\beta > 0$. Initialize the discount factor δ. Initialize $\mu > 0$, the stopping criterion and decide on maxcase, the maximum

number of training cases to try. Initialize the aveerr = 1. Select n, the number of wires. Initialize the pseudo-random number PRN generator. Either read a file of triplets $(\mathbf{x}^t,\mathbf{u}^t,E^t)$ for t=1,2,...,p or else set the number of position vectors tried to p = p_0 and generate this many of the triplets to start the memory-based learning. Select the largest value of p to use before over-writing the triplets in storage.

Step 1: Generate a (feasible) random position $\mathbf{x}$ on the golf course.

Step 2: There are three situations:

Situation A. If $\mathbf{x}$ is in one of the sand traps, the reward R for this point is -W (a loss). Don't generate a corresponding $\mathbf{u}$ (since there will be no successor). This is the end of the current trajectory. Go to Step 1.

Situation B. If $\mathbf{x}$ is on the green, the reward R for this point is +W (a win). Don't generate a corresponding $\mathbf{u}$ (since again there is no successor). This is the end of the current trajectory. Go to Step 1.

Situation C. If $\mathbf{x}$ is neither on the green or in a sand trap (the usual situation), set t ← t + 1 with $\mathbf{x}^t \leftarrow \mathbf{x}$.

Step 3: Over-write the randomly generated triplets. Compute the wire quantities. Each of the $y_i = y_i(\mathbf{x})$ values for i=1,2,...,n is associated with a wire in one-dimensional space with the corresponding control vectors $\mathbf{u}^{\wedge}_i = \mathbf{u}^{\wedge}_i(\mathbf{x})$ for i=1,2,...,n.

Step 4: Randomly generate a control vector $\mathbf{u}$ and set $\mathbf{u}^t \leftarrow \mathbf{u}$ and then set $\mathbf{u}^{sav} \leftarrow \mathbf{u}^t$.

Step 5: Set $E^t \leftarrow f(\mathbf{x}^t,\mathbf{u}^t)$. This uses the "original" $\mathbf{u}^t$ (even if $\mathbf{u}^t$ is reset later due to an infeasible $\mathbf{x}^{t+1}$).

Step 6: Compute a feasible successor state (where the next golf ball will land with the current state and control) and the discounted value of the successor state. First compute $\mathbf{x}^{t+1}$ (the successor state) using $\mathbf{x}^t$ and $\mathbf{u}^t$. If $\mathbf{x}^{t+1}$ is on the course, continue; otherwise recompute $\mathbf{u}^t$ and $\mathbf{x}^{t+1}$ so that the successor state is

feasible. If $\mathbf{x}^{t+1}$ is on the green, then the reward R ← +W. Compute the error by using cerror ← E^t - R and go on to Step 8. If $\mathbf{x}^{t+1}$ is in a sand trap, then the reward R ← -W. Compute the error by using cerror ← E^t - R and go on to Step 8. Otherwise continue.

Step 7: The ball associated with the successor state $\mathbf{x}^{t+1}$ is still playable. Compute the wires for this and the discounted value of the successor state $\mathbf{x}^{t+1}$. Compute $y_{max} = Max\{ y_i(\mathbf{x}^{t+1}) \}$ for i=1,2,...,n. Then compute cerror ← $E^t - y_{max} \cdot \delta$.

Step 8: Update the current estimate: $E^t \leftarrow E^t + r \cdot cerror$. Update the triplets (over-write the oldest with the newest, if necessary).

Step 9: Test if it is time to stop the training by computing aveerr ← $0.999 \cdot aveerr + 0.001 \cdot cerror^2$. If aveerr $\leq \mu$ or the maximum number of training cases has been exceeded, go on to Step 10. If the reward R = $\pm$W, go back to Step 1, otherwise go back to Step 3.

Step 10: Stop the training process with the $\mathbf{x}^t$, $\mathbf{u}^t$ and E^t values ready to use. Possibly write these to an indicated data file.

The general description of the memory-based *application process* is as follows:

Step 0: Initialize the process.

Initialize the golf course parameters (placement of the green, sand traps, etc.) exactly as in the training process.

Initialize the memory by reading the $\mathbf{x}^t$, $\mathbf{u}^t$ and E^t values from a file and determining p.

Initialize the parameters α (used in Equations 9 and 10) and β (used in Equation 4). Set the number of state variables and control variables to 2.

Step 1: Ask the user if a state vector is to be evaluated. If no vector is to be evaluated, go to Step 3.

Step 2: Have the user enter $\mathbf{x}$. If it is out of bounds, in a sand trap, or on the green, so indicate and go back to Step 1. If it is playable, do an evaluation of Equations 4 - 10. From this, compute y_{max} and set the optimal $\mathbf{u}$ to the

corresponding $\mathbf{u}_{max}$ by using Equations 9 and 10. Indicate the "optimal" control, compute the next point, display it together with an indication of its position. Return to Step 1.

Step 3: Stop the application process.

==========

Appendix B

Gradient-Based Learning Algorithms

The general description of a gradient-based *training process* is as follows:

Step 0: Initialize the process.

Initialize the golf course parameters (size of the course, placement and size of the green, number, placement and size of sand traps, etc.).

Initialize the learning algorithm. Set the number of position vectors tried to 0. The number of state variables s = 2 and q = 2 as the number of control variables. Initialize the $\mathbf{c}$ vector for computing $f(\mathbf{x},\mathbf{u})$. Initialize $\alpha \geq 0$ and $\beta > 0$. Initialize the discount factor δ. Initialize $\mu > 0$, the stopping criterion. Initialize the aveerr to a large positive number. Select n, the number of wires.

Initialize the neural network. The first layer has two (state variable) inputs. The second layer, which is hidden, has a user-specified number of units. The third layer, also hidden, contains (q+1)n controls and wires denoted as $(\mathbf{u}_1,y_1)$, $(\mathbf{u}_2,y_2),\ldots,(\mathbf{u}_n\ y_n)$. Since q = 2, this layer outputs up to $3n \leq 15$ values. Either read the weights and biases from a file or initialize them to small pseudo-random values.

Step 1: Generate a (feasible) random position $\mathbf{x}$ on the golf course, set the trajectory counter $t \leftarrow 0$. Initialize the number of trajectory points kept p = 0.

Step 2: There are three situations:

Situation A. If $\mathbf{x}$ is in one of the sand traps, no forward propagation is necessary since the reward R for this point is -W (a loss). Don't generate a corresponding $\mathbf{u}$ (since there will be no successor). This is the end of the current trajectory. Go to Step 1.

Situation B. If $\mathbf{x}$ is on the green, no forward propagation is necessary since the reward R for this point is +W (a win). Don't generate a corresponding $\mathbf{u}$ (since again there is no successor). This is the end of the current trajectory. Go to Step 1.

Situation C. If $\mathbf{x}$ is neither on the green or in a sand trap (the usual situation), set $t \leftarrow t + 1$, $p \leftarrow p + 1$, with $\mathbf{x}^t \leftarrow \mathbf{x}$.

Step 3: Use $\mathbf{x}^{inp} \leftarrow \mathbf{x}^t$ and do a <u>forward propagation</u> through the neural network, producing $(\mathbf{u}^{out}, \mathbf{y}^{out})$. Each of the y_i values is associated with a wire in one-dimensional space with the two corresponding elements $\mathbf{u}^{out}$ as the corresponding independent vector.

Step 4: Randomly generate a control vector $\mathbf{u}$ and set $\mathbf{u}^t \leftarrow \mathbf{u}$ and then set $\mathbf{u}^{sav} \leftarrow \mathbf{u}^t$. Use this to compute $f(\mathbf{x},\mathbf{u})$. This will be input to the last layer of the net, together with the $\mathbf{u}_i$ any y_i values for $i=1,2,\ldots,n$.

Step 5: Set $f^t \leftarrow f(\mathbf{x}^t,\mathbf{u}^t)$. This uses the "original" $\mathbf{u}^t$ (even if $\mathbf{u}^t$ is reset later due to an infeasible $\mathbf{x}^{t+1}$.

Step 6: Compute a feasible successor state (where the next golf ball will land with the current state and control) and the discounted value of the successor state. First compute $\mathbf{x}^{t+1}$ (the successor state) using $\mathbf{x}^t$ and $\mathbf{u}^t$. If $\mathbf{x}^{t+1}$ is on the course, continue; otherwise recompute $\mathbf{u}^t$ and $\mathbf{x}^{t+1}$ so that the successor state is feasible. If $\mathbf{x}^{t+1}$ is on the green, then the reward $R \leftarrow +W$. Compute the error by using $cerror \leftarrow R - f^t$. Go to Step 8. If $\mathbf{x}^{t+1}$ is in a sand trap, then the reward $R \leftarrow -W$. Compute the error by using $cerror \leftarrow R - f^t$. Go to Step 8. Otherwise continue.

Step 7: The ball associated with the successor state $\mathbf{x}^{t+1}$ is still playable. Compute the discounted value of the successor state $\mathbf{x}^{t+1}$ by setting $\mathbf{x}^{inp} \leftarrow \mathbf{x}^{t+1}$ and performing a <u>forward propagation</u> through the neural network. Compute $y_{max} = Max\{$

y_i } and the corresponding $\mathbf{u}^{out}$ control vector called $\mathbf{u}_{max}$ associated with the index of y_i which gave the maximum. Then compute cerror ← $y_{max} \cdot \delta - f^t$.

Step 8: Using cerror above, together with Equations 12 - 21 above, do a backward propagation through the neural network adjusting the weights and biases. This is exemplar updating, not epoch updating.

Step 9: Test if it is time to stop the training by computing aveerr ← $0.999 \cdot$aveerr + $0.001 \cdot$cerror^2. If aveerr $\leq \mu$, go to Step 10. If the reward R = $\pm$W, go to Step 1, otherwise go to Step 3.

Step 10: Stop the training process with the weights and biases now trained. Possibly write these to an indicated data file.

The general description of the gradient-based *application process* is as follows:

Step 0: Initialize the process.

Initialize the golf course parameters (placement of the green, sand traps, etc.) exactly as in the training process.

Initialize the neural network exactly as in the training process, except that the weights and bias values will be the same as their final values in the training process (e.g., read these from a data file).

Set the number of state variables and control variables (i.e., s = 2, q = 2). Set the number of wires. Initialize the $\mathbf{c}$ vector.

Step 1: Ask the user if a state vector is to be evaluated. If no vector is to be evaluated, go to Step 3.

Step 2: Have the user enter $\mathbf{x} = \mathbf{x}_{inp}$ and do a forward propagation. From this, compute y_{max} and set the optimal $\mathbf{u}$ to the corresponding $\mathbf{u}_{max}$. Return to Step 1.

Step 3: Stop the application process.

ASSESSMENT OF DEVELOPMENTS IN MACHINE TOOL TECHNOLOGY

Anthony C. Okafor
Associate Professor
Department of Mechanical and Aerospace Engineering
and Engineering Mechanics

Univerisity of Missouri-Rolla
Rolla, Missour, 65409-0050

Final Report for:
Summer Faculty Research Program
Wright Laboratory
Wright Patterson Air Force Base
Dayton, Ohio 45433

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

and

Wright Laboratory

September 30, 1996

ASSESSMENT OF DEVELOPMENTS IN MACHINE TOOL TECHNOLOGY

Anthony C. Okafor
Associate Professor
Department of Mechanical and Aerospace Engineering and Engineering Mechanics
University of Missouri-Rolla

Abstract

The Objectives of this research is to assess developments in machine tool technology and their implementations, and to recommend to Wright Lab Man Tech areas they should focus for future funding. The review covers historical background on U.S machine tool industry, identified the factors that caused the decline in U.S. machine tool industry in the early 1980, identifies on-going and planned developments in key machine tool technologies, trends and issues that affect the competitiveness of U. S. machine tool industry and manufacturing in general. These technologies are relevant to both defence and commercial manufacturing. The assessment is made in comparison with developments in Japan and Europe where appropriate.

This report is written by Dr. Anthony C. Okafor, Associate Professor of Mechanical Engineering at the University of Missouri-Rolla, who spent two months (May -July 1996) at the U. S. Air Force Wright Lab's Manufacturing Technology Directorate, Wright Patterson Air Force Base, Dayton, Ohio, as a 1996 Summer Faculty Research Associate. His research and stay was supported by the U. S. Air Force Office of Scientific Research (AFORS) . This report is written for the Chief of the Processing and Fabrication Division, Mr. Bruce Rasmussen. Questions regarding this report should be addressed to Dr. Okafor at Okafor@shuttle.cc.umr.edu, voice mail: (573)-341-4695.

TABLE OF CONTENTS

	Page
I: ABSTRACT	2
II. INTRODUCTION	4
Methodology and sources of information Historical Background	
III. ASSESSMENT OF DEVELOPMENT IN MACHINE TOOL TECHNOLOGY AND THEIR IMPLEMENTATIONS	7
Machine Tool Structure	7
• Hexapod Machine Tool Structure	
Machining	7
• High Speed Spindles	
• High Speed Machining	
• Machining of Advanced Composites	
• Environmentally Conscious Machining	
Machine Tool Metrology and Precision Machining	10
• High Performance Machine Tools	
• Kinematic and Thermal Error Characterization and Compensation	
Monitoring/Sensing/Control	13
CNC Controls	13
• PC-based Open System Architecture	
• Linear Motors/Linear Axis Drives	
IV. References	15
V. Acknowledgements	16

ASSESSMENT OF DEVELOPMENTS IN MACHINE TOOL TECHNOLOGY

Anthony C. Okafor

II. INTRODUCTION

Methodology and Sources of Information

The assessment presented in this report is designed to cover key machine tool technologies and issues that affect the competitiveness of U. S. machine tool industry and manufacturing in general. The review started with a review of literature and reports on machine tool industry, review of machine tool related project reports managed by Man Tech, visits and discussions with technical personnel at the following organizations: The Institute of Advanced Manufacturing Sciences in Cincinnati Ohio (Dr. E. Eugene Merchant, Senior Consultant, and Mr. John Pfiefer, Director of Machining Technology); Cincinnati Milacron Inc. - a machine tool builder (Mr. Joe Kamman, Director, Corporate Research and Development Tech Services, Mr. Richard A. Curless, Manager Advanced Technology, Mr. Eric R. Kline, Manager, Precision Engineering Services, and Mr. John B. Kohls, Manager, Government Systems Operations), National Center for Manufacturing Sciences - a consortium of users and manufacturers of machine tools (Mr. Daniel J. Maas, Executive Director, Manufacturing Processes & Materials, Production Equipment & Systems; Mr. John T. McCabe, Program Manager, Manufacturing Processes and Equipment, Mr. Kerry J. Barnett, Program Manager, Advanced Manufacturing Technology, Mr. T. L. (Tony) Haynes, Program Manager, and Mr. Chuck Ryan, Ph.D., Program Manager), and communications with Mr. Charles F. Carter, Jr., Vice President - Technology, AMT-The Association for Manufacturing Technology, U.S.A; and my personal background in machine tool technology.

The machine tool technologies reviewed are relevant to both defense and commercial manufacturing. Three components of defense manufacturing are: air frame, jet engines and electronics, each with separate concerns. With the shrinking of defense industrial base, defense manufacturing has to depend more on commercial manufacturing.

HISTORICAL BACKGROUND ON U. S. MACHINE TOOL INDUSTRY

Machine tools are a small but vital sector of manufacturing. They are the backbone of any industrialized nation's economy. They are essential for producing the technologies required in an industrial economy. Consequently, improvements in machine tool technology can have dramatic effects on the overall manufacturing performance. Without a healthy domestic machine tool industry, there is risk that U. S. manufacturers will not have access to the latest technology for one to two years. This is because machine tool builders worldwide tend to first sell their most current product lines close to home to ensure that any problems are fixed [1994 Rand Report].

Entering the 1980s, the United States was the largest producer of machine tools, with 20 percent of the world market (see figure 1). Through the 1970s, U. S. leadership in this market was supported by a large domestic demand for machine tools coupled with a steady flow of technological innovations developed for military needs. It pioneered the development of the technologies - NC (Numerical Control) and CNC (Computer Numerical Control) - that was to transform the machine tool industry. The U. S. Air Force provided the funding that made this possible. But within a decade, Japanese and German firms took

the technology that the United States had invented and found ways to produce it more effectively. By the end of the decade, the U. S. machine tool industry was less than half the size of those of Japan and Germany (see figure 1). It fell from the position of leadership to that of second-tier producer. Japan and Germany became the world's leading machine tool producers. The U. S. share of world markets dropped from 20 percent to just over 7 percent a decade later, at the same level as Italy, whose economy is less than one quarter the size of the U. S. economy. More than two-thirds of U. S. machine tool firms had closed, and the U.S. government resorted to protectionism to defend portions of its domestic market [1994 Rand Report]. The loss of U. S. machine tool competitiveness was most evident in the domestic market, where imports surged from 24 to 54 percent of total U. S. sales in just seven years. This as a consequence, resulted in the evaporation of the historical U. S. trade surplus in machine tool, and the concomitant trade deficit of $1.7 billions in 1986.

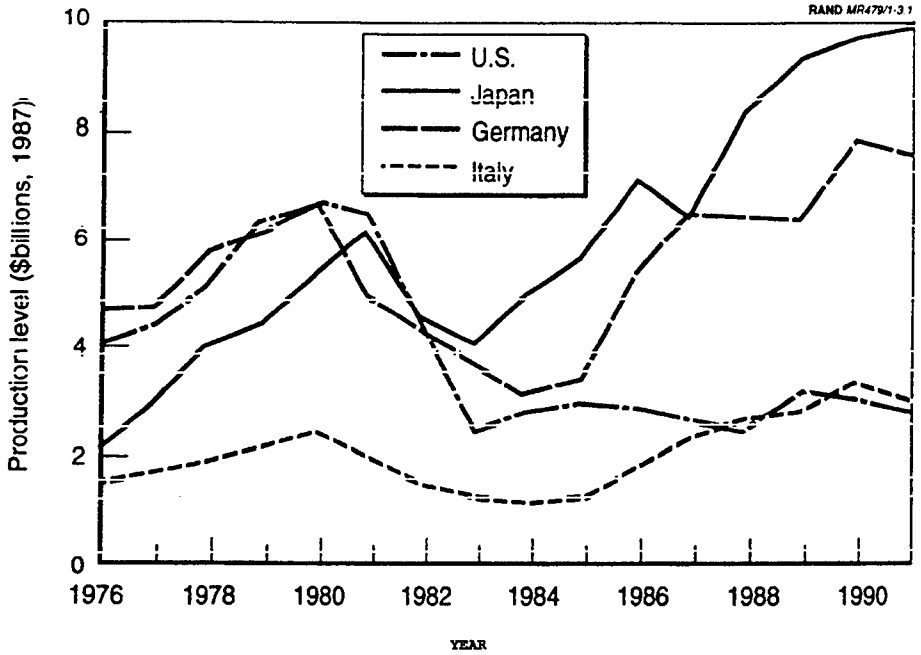


fig 1. Machine Tool Production

Factors Identified as the Causes of the Decline in U.S. Machine Tool Industry in Early 1980 (1994 Rand Report)

The 1994 Rand report identified the following four factors as the causes of the U.S. machine tool industry's decline in the early 1980.

- The machine tool industry overdependent on a ***declining U. S. market.***
- The ***growing international competition*** - the U. S. machine tool maker's strategy for coping with cyclical demand-accumulating orders during booms and working them off during recessions - backfired, as imports gained U. S. market share by delivering in weeks rather than months

- ***Japan became the world productivity leader*** by combining a new product technology (reliable, standardized CNC tools) with a major process innovation (modular production).
- ***The high value of the dollar to other major currencies*** during this period, thus hurting the price competitiveness of the U. S. industry both at home and in export markets.

Additional Factors (Discussions with Mr. Richard Curless, Cincinnati Milacron)

Mr. Richard Curless of Cincinnati Milacron noted that major machine tool technologies (NC, CNC, Tool Changing, and FMS) were developed in the 1960s, and pointed out the following additional factors that influenced U. S. machine tool decline in the early 1980s:

- Research in the U. S. were done by individual machine tool builders in isolation, while Japanese were more collaborative.
- Developments of new technology from concept to prototype took long 2 - 5 years in the U. S., while it took Japanese 9-12 months.
- Japan have their technical and sales people from the field and engineers in house.
- They were good in taking good ideas from technology and implementing it.
- They do this with their cameras, taking pictures of machines, etc. All these started in mid 70s.
- They avoided basic research, taking the technology that the U. S. has developed and commercializing it. American machine tool companies died.

U. S. Approach Today

Today U. .S. approach is different from what it used to be. They have taken the approach that anything they do today must include commercialization. They can take pictures, go to Europe and Japan and shoot pictures like the Japanese did, or they can collaborate. This is the direction that I was told that Cincinnati Milacron has taking now for the past 7 - 8 years. Everyone is moving towards team approach, concurrent engineering. They try to get their customers involved. This new team approach at Cincinnati Milacron, borrowed from Japan, is reported to have yielded positive results. Cincinnati Milacron call their design team approach strategy "WOLFPACKS" - because there are teams meant to be killers of the competition. It was used in the development of their Viking centerless grinding machine which they claim is a major success in terms of accuracy, reliability, cost and market. It is also being used in their current project called HYPERMACH (high performance, high speed machine) to be discussed later. Ingersoll , a major supplier of machine tools and manufacturing systems to the aerospace, automotive, farm, truck, and tractor industries, designs machines in collaboration with customers, even building factories and service centers close to major customers. It is able to respond to customer needs by building state-of-the-arts manufacturing capability that concentrates on flexibility, and always quick to implement the latest in technology.

Government support varies in form, from giving federal laboratories (NIST, NASA, DOE, and DOD) authority to sign cooperative agreements (CRADAs) with industry, universities, and consortia to spend laboratory funds to participate in collaborative research endeavors (technology transfer from labs to firms), to funding for industrial consortia, such National Center for Manufacturing Sciences, (NCMS); and grant programs such as NIST's Advanced Technology Program (ATP), DOD Technology Reinvestment Program (TRP), and establishment of Teaching Factories through NCMS, etc.

III. ASSESSMENT OF DEVELOPMENTS IN MACHINE TOOL TECHNOLOGY

This broad assessment of developments in machine tool technology is divided into several sections. The key technology trends seen impacting the industry are briefly discussed here under each of those discrete headings.

MACHINE TOOL STRUCTURE

Hexapod Machine Tool Structure

A new six legged machine tool structure based on the stewart platform has now been developed. This new machine tool is called "Hexapod". The Stewart platform is a platform or object such as a spindle supported on six expandable columns or cylinders that can assume many positions and attitudes based on the position (length) of each column. It was first described in 1962, but the coordination of each leg was relatively slow, so the idea had limited application. The configuration is now more practical because of the availability of high-speed and multi-axis controls.

The ststed advantages are - low mass and high rigidity. The high rigidity comes from the fact that the arrengement is made so that the supporting and controlling columns or legs never undergo bending. They are always in tension or compression.

An example of vaible hexapod prototype is the Variax from Giddens & Lewis. The Giddens & Lewis version of these six legged machines and a smaller version from Geodetics Inc., a Swiss firm, were first shown at the 1994 International Machine Tool Show (IMTS) in Chicago (American Machinist, May, 1996).

This technology is still in its infancy so it is hard to see any strong trend at this time. The strongest interest in Variax hexapod is reported to be from the aerospace industry. Due to the machine's rigidity, speed, contouring, and accuracy, other industries like tool and die, automotive and compressor manufacturers are also reported to be interested.

Issues:

There is a need to evaluate the implications and extent of the capabilities of hexapod.

MACHINING

HIGH SPEED SPINDLES

Manufacturers in the USA, Europe, and Japan are looking for the most efficient way to remove more metal faster. For this to happen, all of the machine's characteristics need to be rethought, from the basic configuration to the tool holder. There is an increasing trend towards high-speed machining and a concommitant need for high- speed spindles. This section reviews the developements of high speed spindles in the USA and Japan.

High speed Spindle: Developments in Japan

Improved cutting tools and increase in productivity require higher spindle speed, and higher spindle speeds mean more heat and failures unless changes are made to bearings. According to 1996 NSF report (Jay Lee), Mazak claimed that customers complain that spindle bearings needed to be re-adjusted and replaced every year. It was also reported that Hitachi-Seiki indicated that 16,000 - 20,000 rpm is their standard design

for their products, and that 40,000 rpm is being requested by some customers for special applications. The kind of special applications for 40,000 rpm is not indicated. They believe that spindle under 20,000 rpm is more realistic for manufacturing industry in Japan.

High Speed Spindle: Development in USA

The Alliance for Strategic Machine Tool Technologies Initiatives (NCMS, Ford, Ingersoll Rand, GM, etc.) are developing three spindles using Roller bearings, ball bearings, and hydrostatic (fluid film) bearings. Another company, Precision Magnetic Bearing Systems (PREMAG) is developing High speed magnetic spindles.

1. **High Speed Spindle for Ford Power Train** This is a 12,000 rpm spindle with 75 motor horse power. It uses 2 taper roller bearings, a motor drive, and a 94 mm front bearing bore. The machine selected for this development is Ingersoll high velocity milling machine.

   Issues: The issues to be concerned with here are: good heat transfer, sustaining accuracy over long period, due to roller bearing wear.

2. **High Speed Spindle for Ingersoll Rand in Mainfield Mchigan.** This is a 20,000 rpm spindle with 35 motor horse power. It uses ball bearings .

   Issues: The Issues to be concerned here are high run-out of the bearings which is reported to be between 15-25 micro inch. Run-out is caused by out-of-roundness of the balls due to wear and races distortions under varying loads.

3. **High Speed Spindle for GM Power Train in Toledo.** This is initially a 3000 rpm spindle using hydrostatic (fluid film) bearing to test proof of concept. A higher speed version up to 100, 000 rpm is to be developed later. The spindle has a through the shaft cooling. The lubricant is the cutting fluid. The technology will be used in three places: a) high speed machining of components (Boeing is requesting a 100,000 rpm spindle from Cincinnati Milacron), b) a high performance tooling, and c) high precision grinder using Milacron's Viking grinder platform.

Issues: The Issues here are: It requires high pressure fluid to support the shaft which needs high pressure pumps, sealant, and are prone to fluid borne contaminants.

High Speed Magnetic Bearing Spindles for Grinding Machines by PREMAG

The development of the magnetic spindle is conducted in three phases.

Phase I: Concept Spindle Development Incorporating Passive Magnetic Bearing: This is a 8,000 rpm shaft speed with .83 motor horse power, 6,000 Lb/in radial stiffness, and 56,000 Lb/in thrust bearing stiffness.

Issues: The issues here are: spindle cooling technique to be explored; significant heating at the startor motor which increases with speed; and dynamic balancing due to unalignment of magnetic center.

Phase II: High Precision Magnetic Spindle Design. This is a 26,000 rpm spindle speed with a 5 Hp maximum motor power, 1.5 in/Lb maximum spindle torque, and 85,000 Lb/in spindle stiffness.

Issues: The issues here are none.

Phase III: Ultra High Precision Spindle: This is a 86,000 rpm shaft speed with a 3 HP motor power, 5 micro-in spindle run-out, 100,000 Lb/in spindle nose stiffness, and 1.5 in-Lb spindle torque.

Issues:

The Issues here are: explore vacuum and liquid cooling for all coils; study of rotor dynamics to evaluate a method to control angular instability.

Conclusions on High Speed Spindle

- U. S. has been demanding higher and higher spindle speeds up to 120,000 rpm.
- Japan and Europe are not concerned with very high spindle speed but spindle speeds up to 20,000 rpm with uniform torque.
- Higher spindle speed and constant torque throughout the speed range is important
- Torque should not be sacrified in order to obtain higher speed.

HIGH SPEED MACHINING

High speed machining is an emerging technology that offers significant potential for fabricating structural components - faster and more accurately. It results in lower cutting forces and greater metal removal rates and thus enables the production of thin walled monolithic air craft parts since the force exerted against walls of the part are reduced. Such parts are much less expensive and more accurate than their riveted counterparts and are usually lighter. It offers an order of magnitude improvement in machine tool productivity and precision. High speed machine tool spindles are being developed that offers the above numerated advantages, but development of advanced machine tools and controls required to capitalize on this technology has not kept pace.

Increased productivity demands high speed, accurate workpiece positioning systems plus controls with high speed servo loops and integration of new sensors. U. S. aircraft manufacturers estimate they must cut production costs by 30 - 40 % in 3 to 5 years to be globally competitive. Furthermore, advanced machining capabilities are needed to improve the affordability of producing new military aircraft, such as JAST. One obvious possibility to cut production costs is by using unitized (monolithic) machined parts, thereby dramatically reducing part counts and the related high assembly costs.

Issues: with high speed machining the following issues need to be considered or investigated - tool holding and balancing; active vibration control (at low speed machining we have few (5-10) cycles of tooth impact, while at high speed machining we have more (up to 400) cycles, and thus we must worry about machine panel resonance; speed of control; axis drive (there is a trend towards linear motors); mechanics of high speed machining (chip formation, tool wear mechanism, cutting forces, and temperature); servo system; following error (previously has a gain of one, now the gain has increased); optimization of the cutting process, etc.

MACHINING OF ADVANCED COMPOSITES

Advanced composites are increasingly being used in aerospace, naval, space, and automotive vehicles, due to their high specific strength and stiffness, light weight, tailorable coefficient of thermal expansion, and in particular, the ability to generate near zero thermal expansion structures. There is no handbook of knowledge available on

machining conditions for these materials unlike for their metal counterparts that has machining handbook by METCUT. Advanced composites require high speed machining along with reduced feeds. Standard tools are suitable only for short production run due to rapid tool wear rate. Presently used tools are tungsten carbide (WC), polycrystalline diamond (PCD) or diamond tipped or coated.

Issues:

The Issues here are: it is necessary to acquire sufficient knowledge and data on machining of these materials to determine optimum machining conditions. Other issues include delamination, quality of machined components, high temperature (for drilling of thermoplastic composites, high temperature results in chip packing, overheating and melting of workpiece; therefore low helix angles on the order of 10 to 50, clearances of 9 - 20, and low point angles of 60 to 120, with sharp drills are generally used), tool forces and tool wear.

ENVIRONMENTALLY CONSCIOUS MACHINING

Manufacturers in Japan are searching for cutting fluids that are useful for a wide variety of applications and that can be easily recycled[Lee, Jay]. It is also reported that more companies are also interested in installing in-house coolant and lubricant recycling facilities, while many others are studying methods to eliminate lubricants entirely through the use of dry processing. One such method is the use of a vacuum suction head, with compliance controlled by an air cylinder, which encloses the cutting tools. The pressure difference results in the suction force to collect the chips to a container for recycling. OKK machine works is reported to be testing the system on the production floor.

In the USA, environmental restrictions will require much more secure guarding in order to contain mist created by the action of the coolant and the process [Carter, C. F]. Higher cutting speeds and higher coolant pressures will exacerbate the situation. Since the cost of guarding for fluid containment is already significant, we can expect to see innovation in guard design and material selection. Also a new joint research initiative between USA National Science Foundation (NSF) and Environmental Protection Agency (EPA) issued in April, 1996, requests for proposals for research on elimination of cutting fluids in machining. Researchers at the University of Missouri-Rolla, which the author is involved, are planning to conduct research on intelligent and environmentally conscious machining with lubricooling by high pressure waterjet. The proposed research will investigate the use of high pressure waterjet , which is not detrimental to the operator and environment, to replace cutting fluids in machining. Preliminary research conducted at Rolla, show that water under high pressure can serve both as coolant and lubricant, thereby reducing the cutting force and prolonging tool life. Other environmentally conscious machining research at US universities (notably Michigan Tech), are concerned with: understanding/modeling the role of cutting fluids in machining, determining environmentally conscious machining process conditions, Classification of cutting fluids and establishment of analytical fluid selection criteria, and classification of the chip formation process. The introduction of more stringent standards in Europe will create well-defined procedures for documentation and warning signs which will become international in application.

Issues here include: classification of cutting fluids and environmental effects, fixturing for vacuum suction head for the research done in Japan, effects of all the proposed substitutes for cutting fluids on the quality of machined components, tool wear and tool life.

- ***High Performance Machine Tool***

Cincinnati Milacron in partnership with United Technologies, Remmele Engineering, Boeing, McDonnell Douglas, NCMS, is investigating the development of advanced high performance machine tool which exploit the emerging technology of high-speed spindles. The objective of Cincinnati and user partners is to develop a new high speed Alpha machine test platform that encompasses a full complement of enabling technologies needed to achieve ultra-fast manufacturing of aircraft components accurately.

The high performance machine program, referred to as "**HYPERMACH**" is a three (3) phase effort: **Phase I is a State-of-the-Art study project** funded by NCMS, Lawrence Livermore National Labs, and Industrial partners including Cincinnati Milacron, Boeing, United Technologies Research Center, McDonnell Douglas, and Remmele Engineering. The study project's objective was to identify the gap that exists between the High Performance Machine (HPM) equipment needs of the manufacturing user community and the HPM technology that are available on equipment in the commercial marketplace and in the Beta test sites. This phase I project was completed in the first quarter of 1996. **Phase II is a full scale engineering design of an Alpha machine** using a model based approach that can incorporate the dynamic performance characteristics of machine, control, and process. The design will include use of new high performance components necessary for achieving the fast process capability described in the program objective. The required machine parameters include:

- Spindle rpm greater than 60,000, capable of reliable integration into a 5-axis platform
- Spindle power greater than 60 HP
- Contouring feed rates greater than 2,500 IPM, approaching 3-4,000 IPM with advanced linear-motor technology
- Axes acceleration/deceleration in the range of 1.5-2 G
- Dynamic stiffness appropriate for achieving 0,002" tolerance
- Open Architecture high speed control capable of processing sensor data.

This phase II project has been approved by NCMS and is pending sign-off by U.S. Air Force.

Phase III is the final phase of the alpha program. It involves the manufacture of the Alpha machine by Cincinnati Milacron, and the installation and test of this machine at Remmele Engineering. Remmele plans to undertake a comprehensive series of tests to debug and refine various machining processes, using aircraft parts provided by the end-users. Machine performance tests including both function tests and reliability tests, technology and process evaluations, and utilization procedures will be accomplished during this phase. Feedback and comprehensive analysis based upon this series of machining tests would enable Cincinnati Milacron to design, build, and deliver **beta machines** to end-users, McDonnell Douglas and Boeing, Once these **beta machines** are fully tested in production environments at end-user facilities, the results would be available for Milacron refinement of machine design and development of reliable production-hardened equipment for sale in the commercial marketplace.

- ***Kinematic and Thermal Error Characterization and Compensation***

The greatest contributor to loss of accuracy in machine tools is thermal distortion. Machines have always been designed to minimize the influence of temperature (heat) on machine structures, but thermal distortion remains a major problem. This becomes a more significant problem with high speed machining, as machine tool users are demanding higher rotating speeds and higher feed rates to increase productivity without sacrificing accuracy. This requires the control of heat generated by the bearings and ball screws. Two approaches are generally considered when trying to control the thermal deformation: error avoidance approach and error compensation approach. In Japan, error avoidance techniques are normally implemented at the design stage of a machine, such as minimizing generated heat, symmetrical structural design, thermally stable materials, cooling for spindles and ball screws, and so on [Lee,J, 1996]. Controlling the thermal behavior by using circulated coolant through spindle and ball screws are becoming standard practices. For example, Lee reported that Okamoto Machine Tool Works Ltd. uses a water-cooling system on the wheelhead for all precision grinders, and that Hitachi-Seiki uses an oil circulated system on the spindles for all Vertical Machining Centers (VMCs). NSK produces a range of sizes of standard hollow ball screws. Most of machine tools in Japan, according to the report, are equipped with manual functions for operators to offset the cutting conditions through constant measurement of dimensional variation from parts. It is reported that the typical oil feed rate for forced cooling is about 1 gal./min. It is claimed that thermal displacement can be reduced by 30 to 50 percent by using this approach. According to Lee, few companies in Japan control thermal deformation through software compensation. Hitachi-Seiki, it is reported, has developed a ATAC10 method (Artificial Intelligence Thermal Accuracy Control within 10 = B5m range)to off-set the machining condition to compensate for thermal errors. Basically it linearizes experimental errors from spindle head, spindle nose, and linear errors. A prediction can be made based on the linearized chart. Hitachi-Seiki, it is reported, claimed that the result is very promising.

Similar developments in error avoidance and error compensation approaches are going on in the USA. For example, under the strategic machine tool technology initiative, NCMS, Ford, Ingersoll Rand, and GM are developing high speed spindle with hydrostatic (fluid film) bearing and through the shaft cooling as reported in the section under high speed spindle. Lubricant is the cutting fluid. The same through the shaft cooling is being used to develop high performance tooling.

Also within the last ten years machine tool builders have incorporated the use of measuring probes to compensate for thermal error. On a tool changing machine, this is accomplished by loading a measuring probe into the spindle and probing a gauge position on the part or fixture. Any deviation from the expected or desired position is fed into the control as a correction [Carter, C. F].

More recently, builders are developing a method whereby thermocouples are placed at strategic points feed into a mathematical model of the machine tool, and corrections are made at frequent intervals [Carter, C,. F]. This method is believed to hold great promise and early results show compensation for thermal error in the order of 10:1.

Other current error compensation developments in the USA include the following:

- ***Error Compensation Using Neural Network (Tetra Precision Inc. U.S.A)***

Tetra Precision Inc., through the Department of Defense (DOD) Small Business Innovation Research (SBIR) Research grant proposed to develop a system for improvement of machining precision in conventional machine tool (Machining Center) using artificial neural network technology. They plan to measure the temperature at critical locations of the slides, structure, and spindle of a three axis Machining Center, and

measure the geometric and thermal errors in the machine tool. The measured errors will be correlated with temperature readings using Fuzzy Logic based neural network. The measurement of axis parametric errors and spindle drift errors are to be made using Laser Ball Bar (LBB).

Issues: The Issues here are there is no plan for actual implementation of error compensation. There is a need to integrate the developed thermal error models with machine tool controller for real-time error compensation.

- ***Spindle Thermal Error Characterization and Compensation Sensor System for Machining Center Accuracy Enhancement (Automated Precision Inc., USA)***

Automated Precision Inc. reported that it has, through a contract with Wright Laboratory's Manufacturing Technology Directorate, developed a low cost, real-time , sensor/computer based, high performance spindle thermal modeling and compensation system for CNC Machining Centers. The system used non-contact sensors and thermal sensors. There was no actual compensation.

Issues: There is need to investigate the implementation of actual error compensation. Though Automated Precision stated that actual compensation was beyond the scope of the research. There is also a need to investigate the effect of cutting conditions, environmental effects, sensor selection and optimum sensor location.

MONITORING/SENSING/CONTROL

In general, most manufacturers use the cutting time or the number of holes drilled in the case of drilling as the criteria of tool life in machining, both in the USA and Japan. This is not an efficient method since tool wear and tool life are random and stochastic. Even tools from the same batch will exhibit different wear and tool life. A more realistic and more accurate method to monitor tool wear and tool life is indirect monitoring and sensing of the cutting process signatures like the cutting force, torque, temperature and acoustic emission. A lot of research has been conducted on tool condition monitoring and sensing mostly in academic and research institutions, but the research results have not been widely implemented on the shop floor. A recent review on tool condition monitoring literature data base listed a total of 520 publications on tool condition monitoring[Teti, R]. A 1996 NSF report [Lee,J], noted that Mitsubishi is jointly undertaking prototype testing of a non-contact magnetostrictive monitoring of the milling cutter with Professor Inasaki at Keio University. Professor Inasaki developed a non-contact tool cutting force measurement technique by using magnetostrictive sensors mounted on the tool holder. A long and slender film which is composed of strong magnetic materials is reported to be deposited on the cutting tool shank with an angle of -45 and +45 degree to the center axis of the tool shank. The system detects radial force, thrust force, and torque by applying the magnetostrictive effects. According to the report, the results have shown better performance than the Kistler dynamometer. Professor Inasaki is also reported to have developed a sensing system for monitoring grinding process. A CBN wheel with a built-in AE sensing system, including batteries, FM transmitter, and signal processing circuits, has been developed and is claimed to be an excellent method for monitoring conditioning process when the super abrasive wheels are used.

CNC CONTROLS

Several manufacturers now produce controls based on personal computer (PC) microprocessors and architecture. Those who use proprietary architecture are providing inputs or front ends based on the PC MS DOS and/or Windows applications. In either case, the user (or builder/integrator) has more capability to add proprietary macros or other operator aids to the control. One of the aids is software which allows the operator to perform some maintenance procedures. Given the operator a broader job function is an increasing user trend. Since it is possible to combine a PC motherboard with available motion control boards, one might ask [Cater]: Why aren't more store-front operations making CNC controls ? The answer , Carter explained, is that established controls are supported by documentation, training, spare parts. As a result, most builders and users are not willing to trade that support structure for limited cost savings on a total system.

Programmable logic controllers (PLC) continue to grow in capability to the point where PLCs combined with motion control begin to look like CNCs.

On **Open architecture**, first it must be recognize that there is no agreed-upon definition for open architecture. To some users, it means being able to accept any communication protocol locally in use. To others, it means having the operator interface on all controllers look exactly the same - an unlikely event [Carter]. To the machine tool application engineers it means having a standard interface for slide drives, and standard input/output for sensors and logic controls. To the engineers in large companies and universities, it means all of the above made from off-the shelf building blocks.

Due to the pressure from both users and integrators, an opening of architecture is occurring and will continue. The MS DOS and Windows front end is an example. There is pressure for a standard slide drive interface, and one, called SERCOS , has emerged from the international standard group (IEC). It is being utilized by builders and users in the U.S. A European consortium of builders and users is working to agree on the use of established international standards for internal bus structure and input/output protocols. That consortium stand the best chance to define open architecture.

In the U.S. a Next Generation Controller (NGC) program was sponsored by the Office of the Secretary of Defense, Air Force Man Tech, and the National Center for Manufacturing Sciences (NCMS). The objective of NGC is to provide standards for the next generation machine controllers based on open system architecture. NGC will develop and validate a specification for an Open System Architecture Standard (SOSAS) for process controllers. Vendors discouraged that because of the investments each company has made. Industry could not agree on anything, as a result, only report came out . General Motors (GM) got frustrated and took it over (NCMS). The new effort now being led by GM is called Open Modular Architecture Controller (OMAC). GM views it as an evolutionary step and not an end by itself. Another effort in open system architecture called Manufacturing Operating System (MOS) is a $52M CRADA project funded by DOE. It is PC based. DOE lab is providing $26M and industry is providing $26M. This CRADA was formed about two years ago. There is no report yet. A DOE led group called TEAM is trying to draw these various group together. There is no new money for this effort.

Issues: there is no need for another controller program. There is a need for a comprehensive study of all these open architecture programs. There an expressed need on the development of Next Generation CNC Language (NGL) that will take the solid modeling feature to the CNC environment. in other words, how do we use the high level mathematics of solid modeling to drive the CNC machine.

In Japan, open architecture controllers are now being implemented by some manufacturers. Fanuc's CNC 150,160, 180, and 2000 series are open architecture CNCs. Fanuc is still the largest producer of CNC. In addition, Fanuc is implementing artificial

intelligence (A.I.) technologies into its Robocut EDM machines, the Robodrill CNCs, and the Roboshot plastic injection molding machines. Independently from Fanuc, a working group named OSEC (Open System Environment for Controller) was formed in 1994 to realize an open architecture in all factory automation equipment. The OSEC consists of three machine tool builders - Toshiba Machines Co., Toyoda Machines Works, Ltd., and Yamazaki Mazak Corp., and IBM Japan, Mitsubishi Electric Corp., and SML Corp. The intent of the OSEC is to develop protocols and specifications for PC-based open architecture controllers in the factory automation industry. A report entitled "Open System Environment for Controller Architecture" was released in September, 1995 to serve as a developmental guide for participants in OSEC [Lee,J.]. It is reported that Mitsubishi is currently developing a new generation process controllers (i.e. PLCs and NCs) under MS-DOS. A prototype PC-based NC controller has been demonstrated and is under testing. Today, Mitsubishi is reported to have 45 percent of the market share on its PLC products in Japan. In addition, Mitsubishi has developed the first fuzzy controller for the V-series die-sinking electrical discharge machines (EDM) in Japan.

LINEAR MOTOR DRIVES

The major dynamic inertia in a machine tool is in the ball screw and motor. For this reason, in the USA, linear motors have been used occasionally, and are now applied to a limited number of machine models (Carter, C. F). Ten years ago, 400 ipm was typical feeding speed (feed rates) on Machining Centers and Turning Centers. Today, these rates are 1000, 1200, and even 1500 or 2000 ipm (Curless). Traditional way systems are being replaced with high capacity, low friction, low cost commercial systems allowing high rates, but still accurate positioning and reliability.

In Japan linear motor are becoming popular in many applications. Now the feeding speed of NC machine tools is generally more than 20 m/min (787.4 ipm), and 60 m/min (2,362.2 ipm) has appeared (Lee, J).

Issues: Heat is one of the problems associated with the application of linear motors. Nevertheless, work continues on linear motors and we may expect wider applications in the future. With increase in speed, vibration and noise have been causing problems in using ball screws. In addition, NSK has indicated that mechatronics customers are constantly requesting linear motor type positioning tables for semiconductor manufacturing instead of the traditional ball screw due to its fast wear out rate.

RECOMMENDATION

From the assessment of developments in machine tool technologies the following technology areas are recommended to Wright Lab Man Tech for funding support:

- High speed magnetic bearing spindles for grinding, milling, and turning
- High performance machine tool phase II & III
- Advanced machining of composites
- High speed machining
- Automatic machine tool error characterization and compensation

REFERENCES

Rand, "The Decline of U. S. Machine Tool Industry and Prospects for Its Sustainable Recovery ", Vol 1, David Finegold etal., 1994

Rand, "The Decline of U. S. Machine Tool Industry and Prospects for Its Sustainable Recovery ", Vol 2 ,Appendices, Edited by David Finegold, 1994

Lee, J. , "Int-96-02 - Status on Research and Production Practices in Industrial Machinery and Machine Tools Industry in Japan ", NSF report : INT 96-02, February 5, 1996

Pritschow, G., and Daniel, C, "Open Control System - A Future-Oriented Concept ", Proceedings of the 27th CIRP International Seminar on Manufacturing Systems, Design, Control and analysis of Manufacturing Systems, May 21-23, 1995, pp 5-17.

Teti, R, "Review of Tool Condition Monitoring Literature Data Base", Report from 44th CIRP Seminar on Computer Aided Tolerancing, Annals of the CIRP, Vol 44/2/1995.

Lau, K. C. , "Development of a Spindle Thermal Error Characterization and Compensation Sensor System for Machining Center Accuracy Enhancement", Final Technical report, WL-TR-93-8034, December, 1993.

Rao, D. T, "Development of Magnetic Spindle for Machine Tool Application", Interim Project Report No. PREMAG-95-04, for ARPA Contract No. SP0900-94-C-0007, September 1, 1995

Carter, C. F, Personal and written Communication, July 3, 1995

Kamman, J, Curless et.al, Round table discussion with four divisional managers at Cincinnati Milacron, June 14 and 21, 1996

Maas, D etal. Discussions with five divisional managers at National Center for Manufacturing Sciences, Ann Arbor, Michigan, July 1, 1996

Merchant, E., and Pfiefer, J., Personal Discussions with Dr. Merchant and Mr. Pfiefer, June 12, 1996

ACKNOWLEDGEMENT

The author would like to thank the following individuals: My Lab point of contact, Mr. G. Shumaker, chief of Manufacturing and Engineering Systems Division, Wright Lab Manufacturing Technonolgy Directorate, for his support and advice; Mr. Bruce Rasmussen, Chief of the Processing and Fabrication Division, Wright Lab Manufacturing Technology Directorate, for his support and suggesting the subject reviewed; Ms Debbie Kennedy, of Processing and Fabrication Division, who shared her office with me during my eight weeks stay at Wright Lab, for her professional and moral support, and for being a good sister; Mr . Vincent Johnson, for making my stay at Wright Lab a pleasant one.

I would also like to thank the following individuals for providing useful information and suggestions during my visit to their respective organizations. The Institute of Advanced Manufacturing Sciences in Cincinnati Ohio (Dr. E. Eugene Merchant, Senior Consultant, and Mr. John Pfiefer, Director of Machining Technology); Cincinnati Milacron Inc., Cincinnati Ohio - a machine tool builder (Mr. Joe Kamman, Director, Corporate Research and Development Tech Services, Mr. Richard A. Curless, Manager Advanced Technology, Mr. Eric R. Kline, Manager, Precision Engineering Services, and Mr. John B. Kohls, Manager, Government Systems Operations), National Center for Manufacturing Sciences - a consortium of users and manufacturers of machine tools (Mr. Daniel J. Maas, Executive Director, Manufacturing Processes & Materials, Production

Equipment & Systems; Mr. John T. McCabe, Program Manager, Manufacturing Processes and Equipment, Mr. Kerry J. Barnett, Program Manager, Advanced Manufacturing Technology, Mr. T. L. (Tony) Haynes, Program Manager, and Mr. Chuck Ryan, Ph.D., Program Manager), and communications with Mr. Charles F. Carter, Jr., Vice President - Technology, AMT-The Association for Manufacturing Technology, U.S.A;

ASSESSING THE SUITABILITY OF THE CFD++ ALGORITHM FOR ADVANCED PROPULSION CONCEPT SIMULATIONS

Dr. Paul D. Orkwis
Assistant Professor
Department of Aerospace Engineering and Engineering Mechanics

University of Cincinnati
Mail Location 70
Cincinnati, Ohio 45221-0070

Final Report for:
Summer Faculty Research Program
Wright Laboratory

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

and

Wright Laboratory

September 1996

ASSESSING THE SUITABILITY OF THE CFD++ ALGORITHM FOR ADVANCED PROPULSION CONCEPT SIMULATIONS

Dr. Paul D. Orkwis
Assistant Professor
Department of Aerospace Engineering and Engineering Mechanics
University of Cincinnati

Abstract

The unified grid CFD++ algorithm was assessed for its suitability as an advanced propulsion concept flow field simulation tool. Six test cases were computed including a compressible flat plate, an unsteady inviscid shock tube, two supersonic backstep configurations, a supersonic open cavity studied previously by the author, and a three-dimensional inlet configuration. Structured grids were used in all cases for comparison against similar solvers. The code was run on several serial workstations and on two parallel machines. It was found that the code operates seamlessly in parallel and preliminary results demonstrated reasonable solution accuracy for the problems considered. This report also details the basic strategies to be followed when applying this code.

ASSESSING THE SUITABILITY OF THE CFD++ ALGORITHM FOR ADVANCED PROPULSION CONCEPT SIMULATIONS

Dr. Paul D. Orkwis

Introduction

The future of computational fluid dynamics (CFD) rests on researchers producing codes that are at the same time accurate, efficient, user-friendly and flexible. A major step in that direction is the CFD++ algorithm of Chakravarthy at Metacomp Technologies [1, 2, 3]. This code wraps several highly accurate cell centered finite volume flux difference split solvers into a "unified" gridding strategy, which allows the user to choose either or both of the major gridding approaches; structured and unstructured. In addition, the code utilizes the latest major advance in computing architecture, the massively parallel supercomputer, by incorporating the most recent "pointwise" turbulence models. These ideas combined hold the potential for providing the "next generation" of CFD solvers, hence the name CFD++ is indeed apt.

This final report will describe the author's work in evaluating the CFD++ code for the purposes of performing advanced propulsion concept simulations. Test cases were computed to determine the accuracy, efficiency and robustness of the algorithm for both steady and unsteady turbulent viscous flows. This work sets the stage for evaluating the code as a more complete propulsion system analysis tool when reaction rate chemistry is eventually included. The following sections discuss briefly some of the details of the algorithm, the run-time requirements, results obtained for six pertinent test cases, and is summarized by conclusions and future recommendations.

Numerical Method

This section describes the numerical models employed in the CFD++ code. The reader is referred to the CFD++ Users Manual [1] and the CFD++ Technical Reference [2] for further details.

Governing Equations

The CFD++ algorithm allows the user to choose one of several equation sets based primarily on the number of turbulence equations solved. One can perform inviscid or laminar viscous calculations using 5 equations, one-equation turbulence model (such as Baldwin-Barth) simulations with 6 equations and two-equation model (such as k-e) computations with 7 equations. The inviscid case solves the three-dimensional Euler equations. The laminar case is solved by the three-dimensional Navier-Stokes equations. The turbulent cases are solved by the three-dimensional Navier-Stokes equations together with one or two turbulence equations. Details of the equations can be found in the CFD++ Technical Manual. It should be

noted that the "full" Navier-Stokes equations are solved rather than some simplified subset.

Of potential interest is the solved for conserved variable vector.

$$q = [e, \rho, \rho u, \rho v, \rho w, \rho\sigma_1, ..., \rho\sigma_n]^T$$

Where e is the total energy, ρ the density, u, v and w the standard Cartesian velocity components, and σ_i are the turbulence quantities. The advantage of the conserved variable approach together with a finite volume formulation is a demonstrated ability to conserve the correct physical variables. This ability has been tested on a simple case.

Grid Cells

The major advancement of CFD++ is its ability to combine structured and unstructured cells. That is, in 2D one can use quadrilateral and/or triangular cells and in 3D hexahedral, triangular prisms and/or tetrahedral cells. In this way, the most appropriate grid can be utilized for each region. An example of this might be structured grids wrapping the near wall region of a body with unstructured grids used farther away. It further allows considerable gridding flexibility since it avoids the use of multiple zone or block strategies in which the iterative solver is sometimes different across zone/block boundaries. Future plans call for the addition of moving grid cells and noncontinuous grid lines, however, the current code is already very capable with respect to nonmoving grids. The code is not quite as flexible as an overset type grid approach, but it does offer conservation properties that the overset approaches do not. However, an overset grid option is planned.

Of course, this additional flexibility comes at the price of increased storage and startup complexity. However, tools are available to port Plot3D type grids into the appropriate CFD++ read-in format. Unfortunately, the typical Plot3D grid does not take advantage of the dual structured/unstructured gridding strategy. Hence, the user will likely need to understand the read-in protocols to correctly set up his/her problem. This requires the user to input the node types and locations, the cell types and defining nodes and the boundary condition defining nodes. Further, if one wishes to set an initial condition or read-in some other solution, that solution can be imposed. It might be helpful to point out that proper variable length (i.e. single/double precision, C/FORTRAN outputs) is important to this process, hence, coding the defining routines in the C programming language might be more appropriate than attempting a FORTRAN solution. These ideas are outlined in the CFD++ User Manual.

Boundary Conditions

The CFD++ code offers a variety of boundary condition options. An advantage of the unified grid structure is its ability to allow one to specify rather arbitrary boundary shapes. Boundary conditions can be imposed both within and on the boundaries of the mesh. A surface tangency condition can be imposed as well as no-slip, with either adiabatic or fixed temperature for solid walls. Supersonic inflows can be completely prescribed from freestream

values or from a set inflow profile (included at the author's suggestion). Supersonic outflows can be handled currently by extrapolation or characteristic variable conditions. A more generic inflow/outflow condition is also available that handles both subsonic and supersonic flows. Symmetry planes can be processed as well. No block to block condition is available since it is unnecessary in the unified accounting system.

Algorithms

The generic setup of the code is as a unified grid finite volume solver. There exist both an explicit Runge-Kutta (RK) and a point implicit (PI) time integration scheme. Both schemes can drive any of the "right-hand-side" (or residual) options. Two differencing schemes are currently available for the residual, a Roe type Riemann solver and a more diffusive Rusanov type differencing. Viscous terms are differenced centrally. Both schemes can be run with either first, second or third order accuracy based on the choice of extrapolation routine. Two flux limiters are available; the minmod limiter and the Van Leer continuous limiter. Second order time accuracy can be achieved with either a fixed timestep (included at the authors suggestion) or fixed maximum CFL switch. CFL is based on the maximum eigenvalue of the flow, hence, computation to a specific time level is best achieved with the fixed timestep option. Steady state solutions can be obtained without regard to time accuracy, hence, first order temporal accuracy is available as well as a fixed CFL condition (thereby producing a timestep that is different at different grid locations.)

Turbulence Models

CFD++ was directed from the beginning to be a parallel algorithm, hence, the developers wanted to make the code as parallel as possible. This is why the explicit and point implicit schemes were chosen as integrators in time, since they do not change with each domain decomposition. To that end the turbulence models were also chosen as pointwise procedures, such that they do not depend upon distances to walls or on some vertical flow profile. The Baldwin and Barth [4] one equation model and the Goldberg [5] two-equation $k-\epsilon$ model were chosen for this purpose. In the former the turbulent eddy viscosity is the dependent variable and in the latter the turbulent kinetic energy, k, and the turbulent dissipation rate, ϵ, are employed. Users must input either field initial conditions or fixed freestream conditions for these properties when used. No algebraic model is currently available.

Outputs

The code package from Metacomp technologies contains a variety of postprocessing options for the data obtained from CFD++. In general, the result is stored in a file (typically called cdepsout.bin) which has the same format as the input file cdepsin.bin. This file can be rewritten in either a Tecplot or Plot3D format, using routines called "quatec", "tritec", "hextec" and "mixtec" or "topl3d". The names of the Tecplot related files correspond to

the type of cell which is in the output. The Tecplot format will output files with any combination of the variables x, y, z, p, ρ, u, v, w, ω_x, ω_y, ω_z, q6, q7, M, T and $|\nabla\rho|$. Where q6 and q7 represent the turbulence quantities, ω_i are the vorticity components and $|\nabla\rho|$ is the magnitude of the density gradient (similar to a Schlieren image). The "topl3d" file produces the standard Plot3D format conserved variables, which can be read by either Plot3D or Tecplot. The Tecplot output from the .tec files can be output in any combination of the above variables by changing the switches in the file "tecplot.inp".

The original code was capable of outputting "restart" files at a user selected interval. A modification was made at the author's suggestion to allow a sequentially ordered series of restart files for plotting unsteady solutions and making movies. These sequential files have the same name as the original restart file but with a .iteration# extension. These files can again be operated on by the Tecplot and Plot3D output conversion routines.

Also included in the set of outputs is a screen-dump of statistics appropriate to each iteration. The data includes the time, timestep, maximum and minimum eigenvalues, maximum right hand side value (rhs) and the sum of the rhs values. This file must be postprocessed to extract the timeseries of residual values. Several small conversion routines were written by the author for this purpose.

In addition to the above, the author suggested the creation of a file to output surface pressure properties each iteration. This is useful for unsteady calculations where one might be interested in a certain cyclic solution behavior which is reached after some initial transient period. It is also useful in computing statistical quantities like the sound pressure level at specific surface locations. The user calls this routine by specifying an information set sequence number in the probe section of mcfd.inp. A specific set of boundary conditions sets is input by this information sequence. Surface pressure values are output to a file named mcfd.info2.bcsN and heat transfer data to a file mcfd.info3.bcsM, where M and N correspond to the boundary condition sequence number. These quantities can be output at select iterations through the variables "ntout2" and "ntout3". The outputs can be read in ASCII form by use of the tools "infout2" and "infout3".

Other postprocessing options have been suggested, including output of the near wall y^+ values.

Parallelization

An important and useful aspect of the CFD++ algorithm is the ease with which it can be run in parallel. The code developers currently provide compilation routines for the Intel Paragon, SGI Power Challenge array, IBM SP2, and the Cray C-90 and T-3D. Apart from machine specific runtime instructions the code requires only the addition of a node-to-processor map contained in the file mcpusin.bin.#. Where # corresponds to the number of processors which will be run. Clearly, the information in that file will be different for each machine partition. A routine named "hexdec" is available which writes the node to cpu map given the desired splitting. Care must be taken to ensure that a clean splitting occurs and little wasted effort is expended, which is commonly known as load balancing. Domain decomposition with

"hexdec" distributes each zone of the Plot3D style input onto all processors. A second tool, "tometis", is available which treats the grid as a single zone, in a manner similar to the general unified grid approach.

Results

In this section the results of applying the CFD++ code to several simple test cases related to advanced propulsion concepts are presented. The first is a supersonic compressible flat plate flow, the second is an unsteady shock tube problem, the third and fourth are two conditions for a backstep, the fifth is an unsteady supersonic open cavity and the sixth is a three-dimensional inlet configuration.

Compressible Flat Plate

The laminar compressible flat plate test case was computed first to demonstrate the ability of the code to compute viscous compressible flows. The simple flat plate was run with "slug-flow" initial condition defined by $M_\infty = 2$ and $Re \approx 1.82 \cdot 10^6$ at a temperature of $T_\infty = 282.353K$. Velocity was defined for this case as $U_\infty = 674.82m/sec$, density was determined from the perfect gas equation of state and input as $\rho_\infty = 0.0463472kg/m^3$, and the ratio of specific heats was given as $\gamma = 1.4$. The CFD++ algorithm uses pressure, density, length, temperature, molecular viscosity and coefficient of thermal conductivity as reference quantities. The reference pressure was then computed from the equation

$$P_{Ref} = \rho_{Ref} \cdot {U_{Ref}}^2$$

This value was computed to be $P_{Ref} = 3880.14Pa$.

The nondimensional x direction velocity results are shown in figure 1 together with the exact results found in the White text [6]for the above case. Note that the agreement is excellent, although a slight discrepancy can be noted at the boundary layer edge. This is likely due to the very weak shock wave which forms at the leading edge of the plate. A wall to freestream temperature ratio of approximately 1.648 was computed and should be compared to the analytical value of 1.679. This represents a less than 2% error.

This test case took approximately 1500 iterations to reduce the residual by 3 orders of magnitude. Greater reductions are possible with different combinations of solvers/limiters. The code required approximately 1.05 seconds per iteration on an SGI R8000 Power Indigo 2. A clustered grid dimensioned 50 cells by 50 cells was employed. The point implicit solver was used in this case. It required CFL to be set to 1.0 initially, but could be increased to 200 after the solution had begun to develop. The low CFL was kept for the first 600 iterations. This is a common approach for using CFD++ and is used again in later sections.

Shock Tube

A shock tube was chosen next to test the time accuracy of the code. This case was run as two-dimensional, although CFD++ permits calculations in a one-dimensional mode. This case also offered an opportunity to learn how to format the solution file cdepsin.bin file for defining an initial guess. The test case was defined by a pressure ratio of 2:1, with the low pressure side set at $1atm$ and density set to $\rho = 1.24204kg/m^3$ throughout. Initial temperature was defined through the perfect gas equation of state. The governing equations were solved for the unsteady shock and expansion relations as done in Anderson's [7] text. This provided the basis for an "exact" solution with which to compare. A 100 node fixed spacing grid was employed.

Figure 2 shows the results obtained versus the exact solution at nondimensional time 0.521278 with the Runge-Kutta algorithm run at $CFL = 1.0$ for both flux limiters and no flux limiter. The compression parameter for the minmod limiter was set to 4. These results are representative of the general results obtained with a variety of CFL values and at various time levels. The results indicated that the minmod limiter with a compression parameter of 1 was indistinguishable from the Van Leer continuous limiter, hence, these results were not shown. Increasing the compression parameter steepens the slope of the flow features. A slightly leading phase error is apparent for the shock wave in all three cases, since the midpoint of the shock is forward of the exact result. This might also be expected because the initial condition is somewhat smeared due to the discrete nature of the solution. Phase results for the contact discontinuity appear to be predicted more accurately. The expansion is damped somewhat in all cases and the phase error again appears to be slightly leading.

The necessity for flux limiting is apparent from the non-limited results, where property overshoots were found as expected. It is suggested from the results that good inviscid time accuracy can be expected. This property is tested further via the cavity case.

Backsteps

Next the ability of the code to compute more complicated flow physics was tested by simulating the flow about two laminar backsteps studied experimentally by Brown and Jakubowski [8]. These flows are both nominally above Mach 4 but represent relatively low values of Reynolds number. Surface heat transfer rates are available for the first and surface pressure data for the second. The former corresponds to the experimental case 180-1-21 and the latter to 200-1-27. In these cases the backstep model is placed into the flow stream of an arc-jet type windtunnel, hence, it is appropriate to assume zero boundary layer thickness at the leading edge of the plate. The length of the forward plate is $12.5cm$, the step is $1.02cm$, and the aft plate is $11.4cm$. The experimentalists suggest that only a portion of the aft plate data is free of three dimensional influence. Case one is defined via $Re_L = 2.65 \cdot 10^3$ and $M_\infty = 4.23$, while case 2 is defined via $Re_L = 2.496 \cdot 10^4$ and $M_\infty = 4.07$. Test freestream values of pressure, density, temperature and velocity, stagnation pressure, temperature and enthalpy, and wall temperature are given from frozen flow calculations. The freestream values were used to define the reference properties in CFD++.

A 40x20 grid was employed in the forward region and a 50x40 grid was used in the aft region. The grid was clustered in the wall region of the forward and aft plates, and the forward region clustering was continued directly across the aft wall region. The code was again run on an SGI R8000 Power Indigo 2 and took approximately 1.08 seconds per iteration. A more complicated strategy was employed to raise the CFL and lower the surface temperature in this case, since both of these values lead to code instability due to transient growth.

In the first case the wall temperature was held at the freestream value and the CFL was set initially to 0.001. The code was run first order accurate in space with mindis=1 and itsync=0, the former modifies the eigenvalue computations and the latter causes the code to run with a fixed CFL. The code was run 200 iterations in this mode and the CFL was then raised to 0.1 with the same wall temperature and other conditions. The wall temperature was then lowered to one half of freestream with other quantities held fixed and again 200 iterations were run. The wall temperature was then lowered to 35% of freestream and the code run 550 iterations. The wall temperature was then lowered to its final value of approximately 23% of freestream and run for 20000 iterations. The solution was then able to support reduced dissipation differencing, therefore, the code was switched to second order spatial accuracy and normal eigenvalue computations. It was run at CFL=0.1 for 2000 iterations, then CFL=0.2 for 2000 iterations, CFL=0.3 for 2000 iterations and finally CFL=0.4 for another 20000 iterations. This was sufficient to drop the residual some 10 orders of magnitude, although accuracy sufficient for plotting purposes was obtained well before that point. Figure 3a shows pressure contours and streamlines in the backstep region. Surface heat transfer outputs were still in the debugging process at time of this writing.

The second case required a similar CFL and temperature variation approach, albeit with different specific details. These will not be included for brevity. It should be noted that in this case the residual was able to be reduced only 5 orders of magnitude. This is likely due to an increased Reynolds number and identical grid combination as compared to the first case.

Nondimensional pressure results are shown for this case in figure 3b. Note that the aft plate pressure is computed fairly accurately but the near step region is rather poor. This again is considered the result of poor grid resolution in the region aft of the step. Grid clustering near the reattachment region is somewhat relaxed, hence the separation length is underpredicted. A more thorough grid resolution analysis is currently being conducted.

Supersonic Open Cavity

The next test case was designed to test the time accuracy of the code when the flow is viscous. The author and colleagues had previously computed this test case [9, 10, 11] with another code and with various algebraic turbulence models, as well as in laminar mode. The test is a simple two-dimensional cavity with a length to depth ratio of 2. A corresponding experimental dataset also exists. In that case the cavity was resident in the side wall of the wind tunnel annd hence was subjected to an oncoming wall boundary layer. The core flow

conditions were $M = 2$ and $Re/m = 37.7 \cdot 10^6/m$. The identical inflow boundary condition was used as in the previously tested cases. The grid was also identical to that case and consisted of three blocks. The forward and aft blocks were made up of 66x55 nodes and the center of 66x120 nodes. Grid line continuity was preserved across zonal interfaces and points were clustered to all walls. A maximum CFL value of approximately 1.0 was employed by using a nondimensional timestep of 0.002.

This was the first test of the code to be run in parallel and the ASC MSRC Paragon machine was the platform of choice. Splittings of 16 and 30 nodes were employed. The "hexdec" tool from Metacomp was used for the splitting. This tool splits each zone so that every processor will have a portion of all zones. The splitting for the 30 node case was nearly as efficient as the 16 node case in spite of the added communication overhead losses because of lower load balancing losses. The CFD++ code performed seamlessly in parallel. Apart from understanding the idiosyncracies of each machine, the only added input was the node map formed by the "hexdec" routine. The code required approximately 3.32 seconds per iteration when run with 16 processors and approximately 1.85 seconds per iteration when run with 30 processors. Further increases in total processors were considered, but it was clear that the grid per processor would be reduced considerably and processor to processor communication would increase greatly, thereby leading to load balancing inefficiency.

The code was initialized by setting the flow above the cavity to that of the inflow profile and flow in the cavity with pressure, temperature and density identical to the wall values and velocities to half their edge values. Transients are typically thought to be purged by approximately 6 characteristic times for these problems, but recent work by the author's student suggests that as many as 30 to 40 characteristic times may be necessary. The code was therefore run for 42 characteristic times and sampled throughout.

Results for this test case were quite promising as shown in figures 4 and 5. Both show a full oscillation cycle of the cavity flow. Figure 4 shows the vorticity contours and can be read from any point in a counter-clockwise manner. Likewise, figure 5 is a cyclic representation of the magnitude of density gradient, which is much like a Schlieren image. Qualitatively, the results demonstrate the same flow physics found previously by the author and his colleagues. A more in depth analysis of the flow field can be found in those references.

The dominant frequency was computed for this flow field time history by analyzing the spectral content of the surface pressure time history through use of the FFT technique. The dominant frequency was found to be 20.12 kHz when data from the last 42 characteristic times (T_c) was used. This should be compared to a value of 26.6kHz found in the previous work and approximately 23 from the experiment. The result is consistent with the previous turbulence model findings in which diffusion was shown to lower dominant frequency. It is reasonable because the code when run with a higher CFL should be expected to be more diffusive than with a low CFL.

The higher CFL also clouds a second comparison which can be made with dominant frequency. That is, how the dominant frequency changes as the characteristic time is increased. As stated earlier, the author's student has found that many cases require a considerable period to purge transients. FFT data is affected by the length of this process, hence only

the later data should be used. This is a problem with higher CFL values because a lower sampling rate is produced. At 42 T_c the discrete FFT results are in approximately 1kHz increments, at 20 T_c the increment is roughly twice that and so on. Results for 20 T_c show the dominant frequency moving to 19.06kHz, however, the increment is now 2kHz, which leaves the issue unresolved. Increased run times will be performed for this case in the near future to resolve this issue conclusively. Other future tests for this problem will involve use of the available turbulence models found in CFD++ and various timesteps.

3-D Inlet

The final test case was a fully three-dimensional supersonic inlet configuration. The inlet was tested experimentally in a wind tunnel and was made up of two symmetric wedge type structures. These structures are placed perpendicular to the tunnel walls as shown in figure 6a. The flow before the structures is a boundary layer profile varying in the z-direction. The computation therefore sees an inflow profile as a boundary condition. This boundary condition was supplied by the code developers at the author's request. The flow then travels in the x-direction and sees a compression corner formed by symmetric 15^o wedges varying in the y-direction. The boundaries then become again parallel to the x-direction. The structures are then closed via symmetric 45^o expansion corners. The latter feature proved to be difficult with regard to initializing the CFD++ starting solution.

Flow conditions were $M_{core} = 8.3$ and oncoming stream Reynolds number per meter of $Re/m = 5.6 \cdot 10^6/m$. Wall temperature was held fixed at $300K$ (although the initial computations shown herein assume an adiabatic wall.) Oncoming boundary layer thickness was found to be $\delta = 3.25cm$. Reference conditions were chosen to reflect these conditions.

The grid for this case was contained 106x80x79 nodes for a total of 640,848 cells. It was used previously to compute the same test case by the author's focal point. The grid was clustered to all solid walls. The size of the grid mandated the use of a parallel computer, hence the Power Challenge array at the U.S. Army Research Laboratory was the platform of choice. The number of nodes is typically limited to an effective maximum of 8 on this cluster. The code required approximately 63 seconds per iteration in that configuration.

The startup of this solution was again quite difficult because of the severe flow field. The flow was started by computing first a two-dimensional constant z plane. Clearly this does not include the effects of the wind tunnel wall boundary layer, hence, once the two dimensional solution was sufficiently converged and the flow features had developed, the solution was projected onto the three-dimensional grid. The flow was then further altered by scaling the velocity components so that they would be similar to the inflow profile. Note that the flow undergoes several acceleration/decelerations because of the wedge structures, so simply using the inflow profile would not be appropriate. The above provided a sufficiently close initial condition to allow the three-dimensional case to be computed.

In spite of the dimensional sequencing applied to this problem, starting even the two-dimensional solution proved difficult. Once again the flow was initialized by starting with more dissipative flux differencing and lower order spatial accuracy. The expansion proved

quite difficult to compute since the initial condition produces a Mach number flow that cannot be turned by 45° in an inviscid sense. That is, the post expansion Mach number becomes infinite. This problem is further exacerbated by relatively poor gridding in that region. A more diffusive flux limiting approach helps this problem.

The results for this case are not currently complete. Shown herein are the grid in figure 6a and pressure contours along two planar slices (y and z slices.) It is apparent from figure 6b that the y constant slice shows an effect from the lower wall on the two-dimensional profile. This profile is more apparent in the constant z-plane slice where multiple shock and expansion reflections are present. It is clearly too early in this computation to provide a reasonable guess as to the quality of the solution, however, it was included to demonstrate the capabilities of the code.

In summary, the above cases will continue to be run until accurate converged solutions are obtained. The current report should therefore be thought of as an update on work in progress. The chosen cases were included as a means to illustrate certain aspects of the code for advanced propulsion concept simulations.

Summary

The CFD++ algorithm was evalulated for use as an advanced propulsion concept simulation tool. In that regard the algorithm was found to be accurate and robust when applied in the appropriate manner. Six test cases were computed to test various aspects of the code for steady and unsteady, viscous flows. Further evaluatory studies will continue to be conducted and suggestions passed on to the code developers.

Future Plans

Work will continue on the test cases that have not yet been converged. Two papers are anticipated to be submitted to the 1997 Joint Propulsion Meeting in Seattle, Washington, with Drs. Sekar, Chakravarthy and Peroomian as co-authors. The papers will focus separately on the steady state and unsteady applications.

Acknowledgements

The author would like to thank the POPS branch at Wright Laboratory for their financial support of this effort. Further thanks goes to Dr. Balu Sekar, the author's focal point while on station, and Drs. Sukumar Chakravarthy and Oshin Peroomian of Metacomp Technologies, Inc. for their guidance and patience. Computational resources were supplied by the POPS and the ASC and ARL DOD Major Shared Resource Centers.

Bibliography

[1] Chakravarthy, S. and Peroomian, O., *CFD++ User Manual, Version 96.07,* Metacomp Technologies, Westlake Village, CA, July 1996.

[2] Chakravarthy, S. and Peroomian, O., *CFD++ Technical Reference,* Metacomp Technologies, Westlake Village, CA, January 1996.

[3] Peroomian, O., Chakravarthy, S., and Sekar, B., "Some Internal Flow Applications of a Unified-Grid CFD Methodology," AIAA Paper 96-2926, 32nd AIAA/ASME/ISAE/ASEE Joint Propulsion Conference and Exhibit, Lake Buena Vista, FL, July, 1996.

[4] Baldwin, B.S. and Barth, T.J., "A One-Equation Turbulence Model for High Reynolds Number Wall-Bounded Flows," AIAA Paper 91-0610, Reno, NV, 1991.

[5] Goldberg, U.C., "Toward a Pointwise Turbulence Model for Wall- Bounded and Free Shear Flows," *ASME Journal of Fluids Engineering,* Vol. 116, No. 1, March 1994, pp. 72-76.

[6] White, Frank M., it Viscous Fluid Flow, McGraw Hill, New York, 1974.

[7] Anderson, J.D. Jr., *Modern Compressible Flow: with Historical Perspective,* McGraw Hill, NY, 1982.

[8] Brown, R.D. and Jakubowski, A.K., "Heat-Transfer and Pressure Distributions for Laminar Separated Flows Downstream of Rearward-Facing Steps with and Without Mass Suction," NASA TN D-7430, August 1974.

[9] Orkwis, P.D., Tam, C.-J., and Disimile, P.J., "Observations on Using Experimental Data as Boundary Conditions for Calculations," *AIAA Journal,* Vol. 33, No. 1, January 1995, pp. 176-178.

[10] Tam, C.-J., Orkwis, P.D., and Disimile, P.J., "A Comparison of Baldwin-Lomax Turbulence Models for 2-D Open Cavity Computations," *AIAA Journal,* Vol. 34, No. 3, March 1996, pp. 629-631.

[11] Tam, C.-J., Orkwis, P.D., and Disimile, P.J., "Supersonic Open Cavity Flow Physics Ascertained from Algebraic Turbulence Model Simulation," accepted *AIAA Journal.*

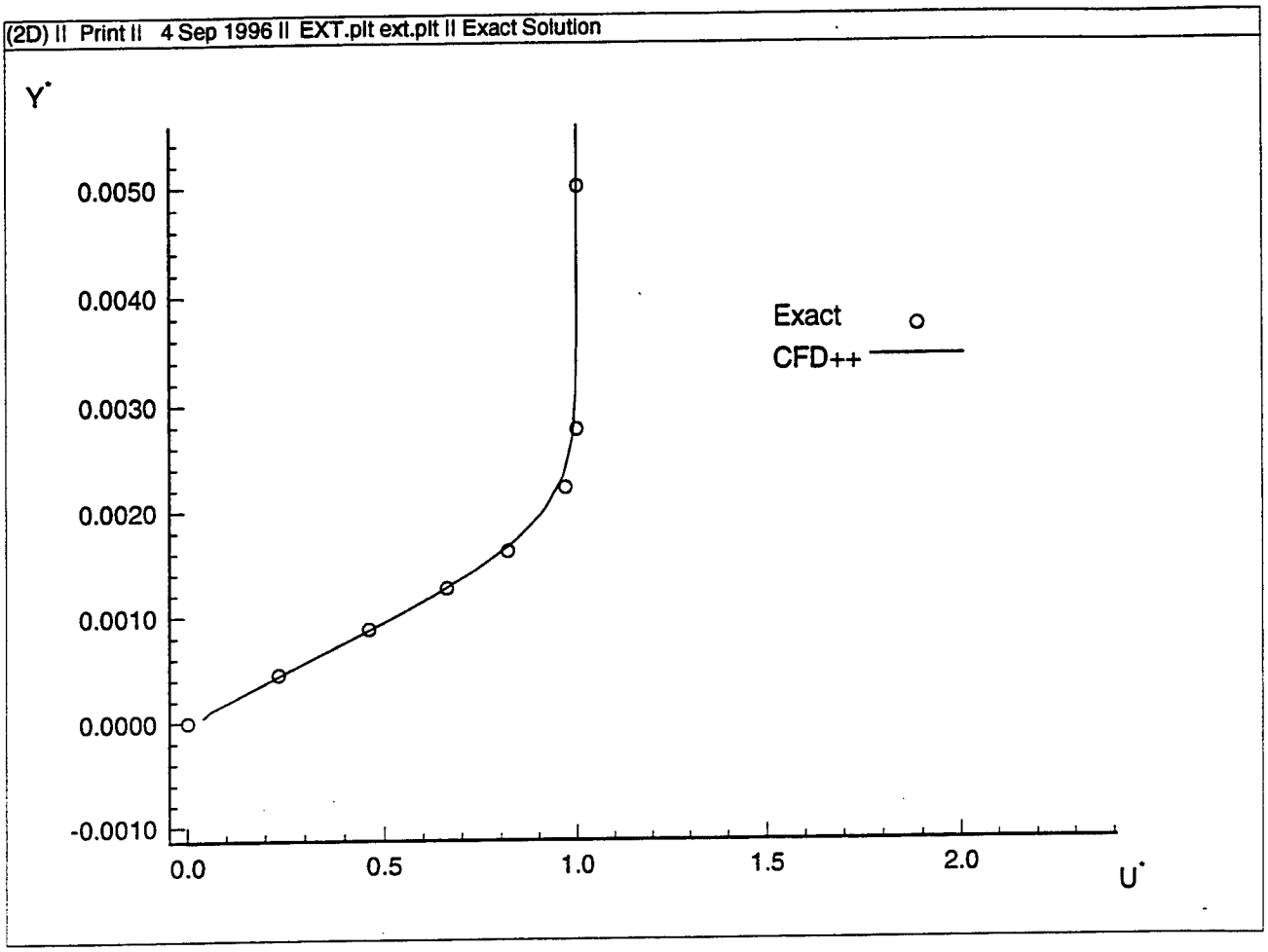


Figure 1: Nondimensional Velocity Profile Comparison with Experiment for Compressible Flat Plate

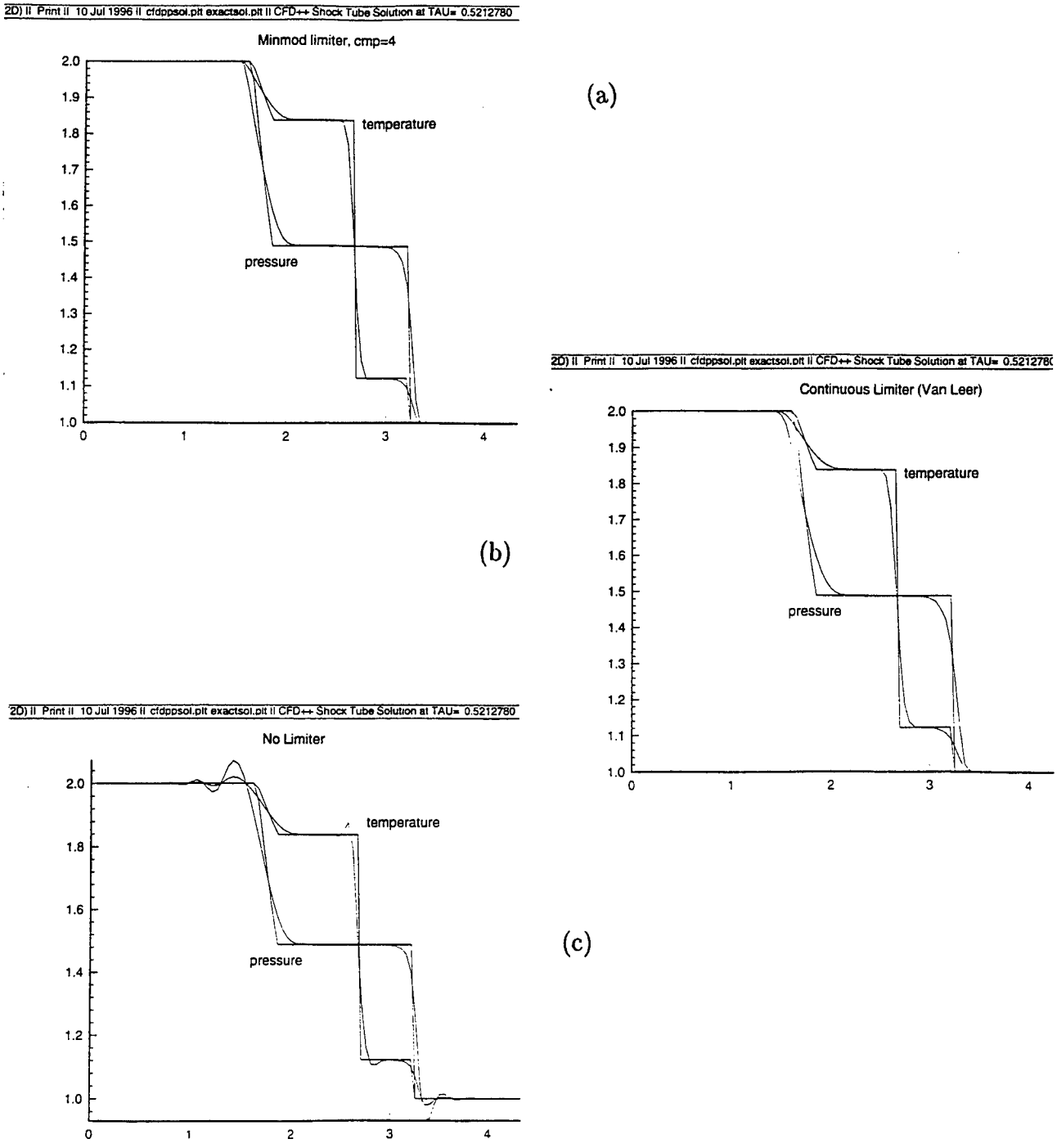
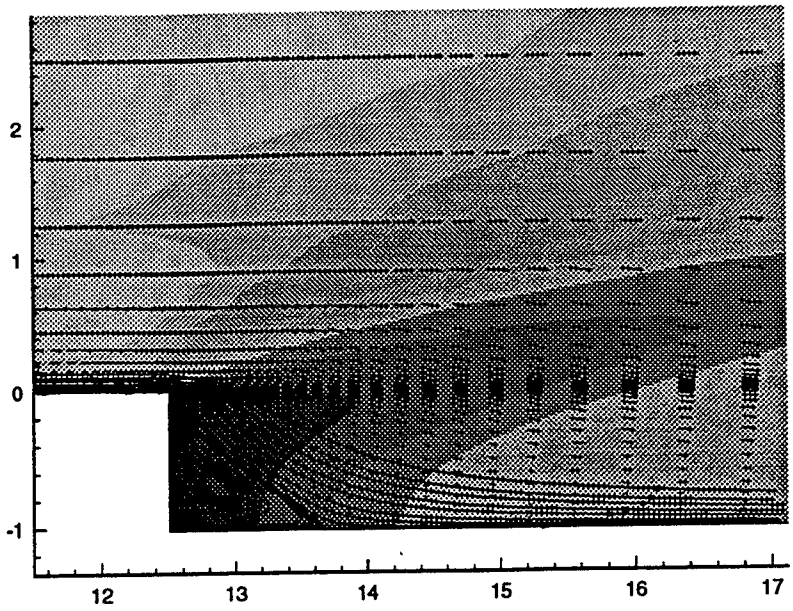
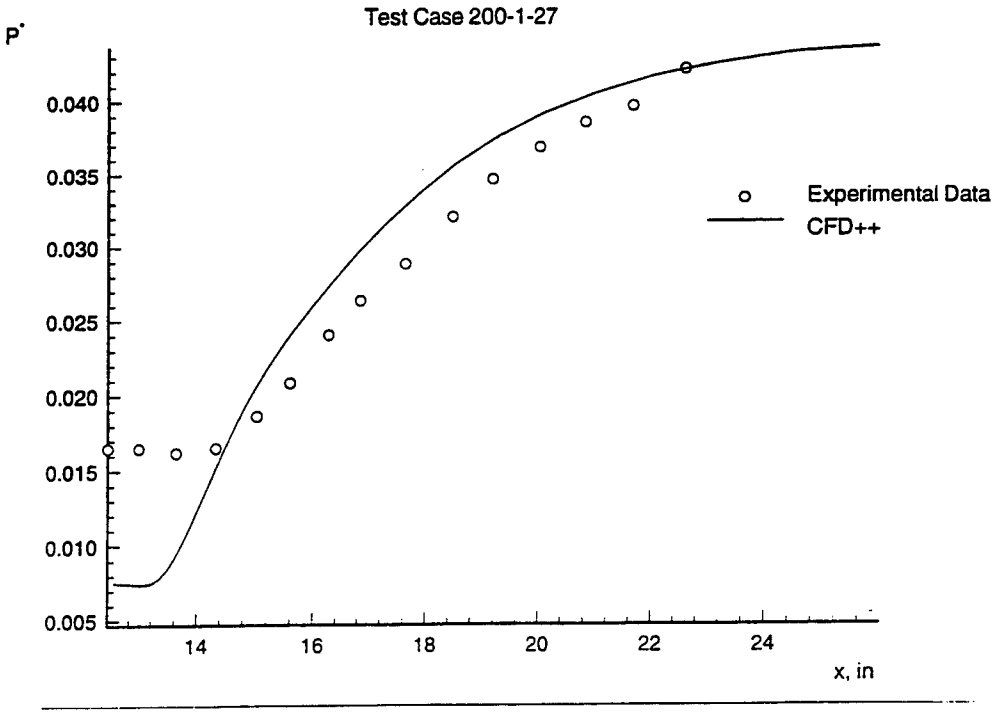


Figure 2: Nondimensional Pressure and Temperature Comparisons with Exact Solution for 2:1 Pressure Ratio Shock Tube at $\tau = 0.521278$. a) Mimmod limiter Compression Parameter = 4, b) Van Leer Continuous Limiter, c) No Limiter.



(a)



(b)

Figure 3: Backstep Cases of Brown and Jakubowski. a) Nondimensional Pressure Contours, Velocity Vectors and Streamlines - case 180-1-21, b) Nondimensional Pressure versus Experimental Data - case 200-1-27

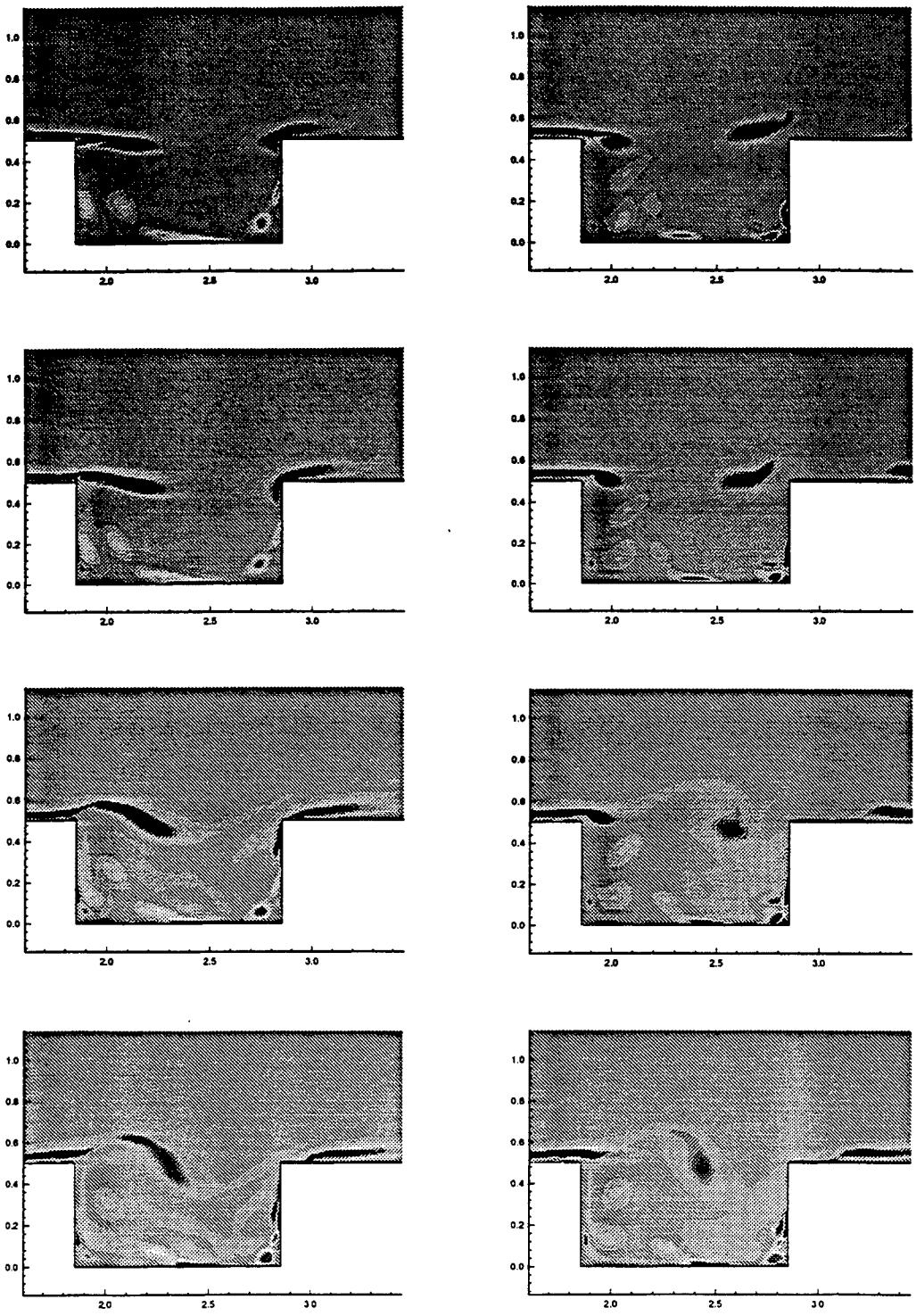


Figure 4: Supersonic Open Cavity Vorticity Contour Cycle

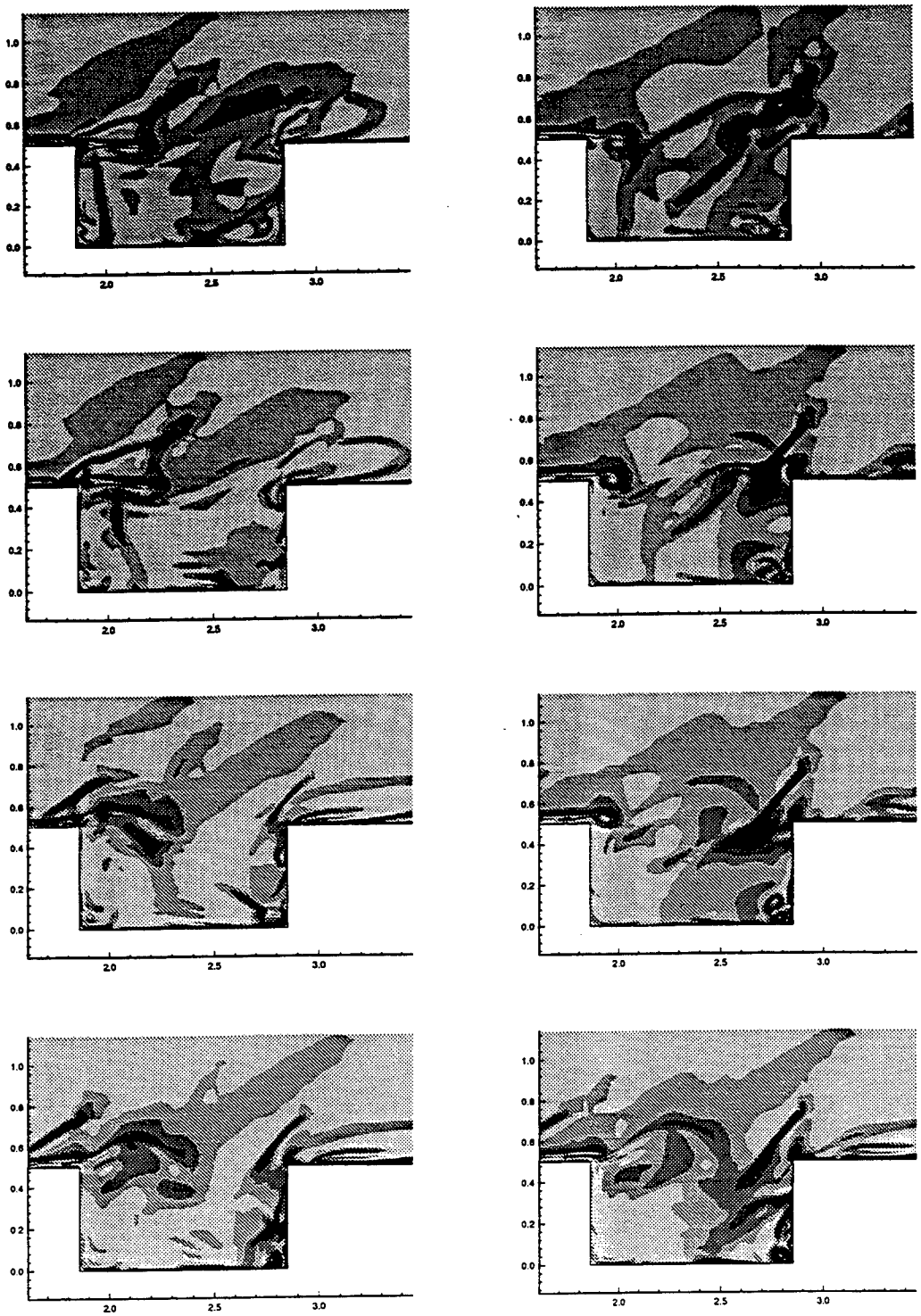


Figure 5: Supersonic Open Cavity Magnitude of Density Gradient Contour Cycle (Simulated Schlieren Images)

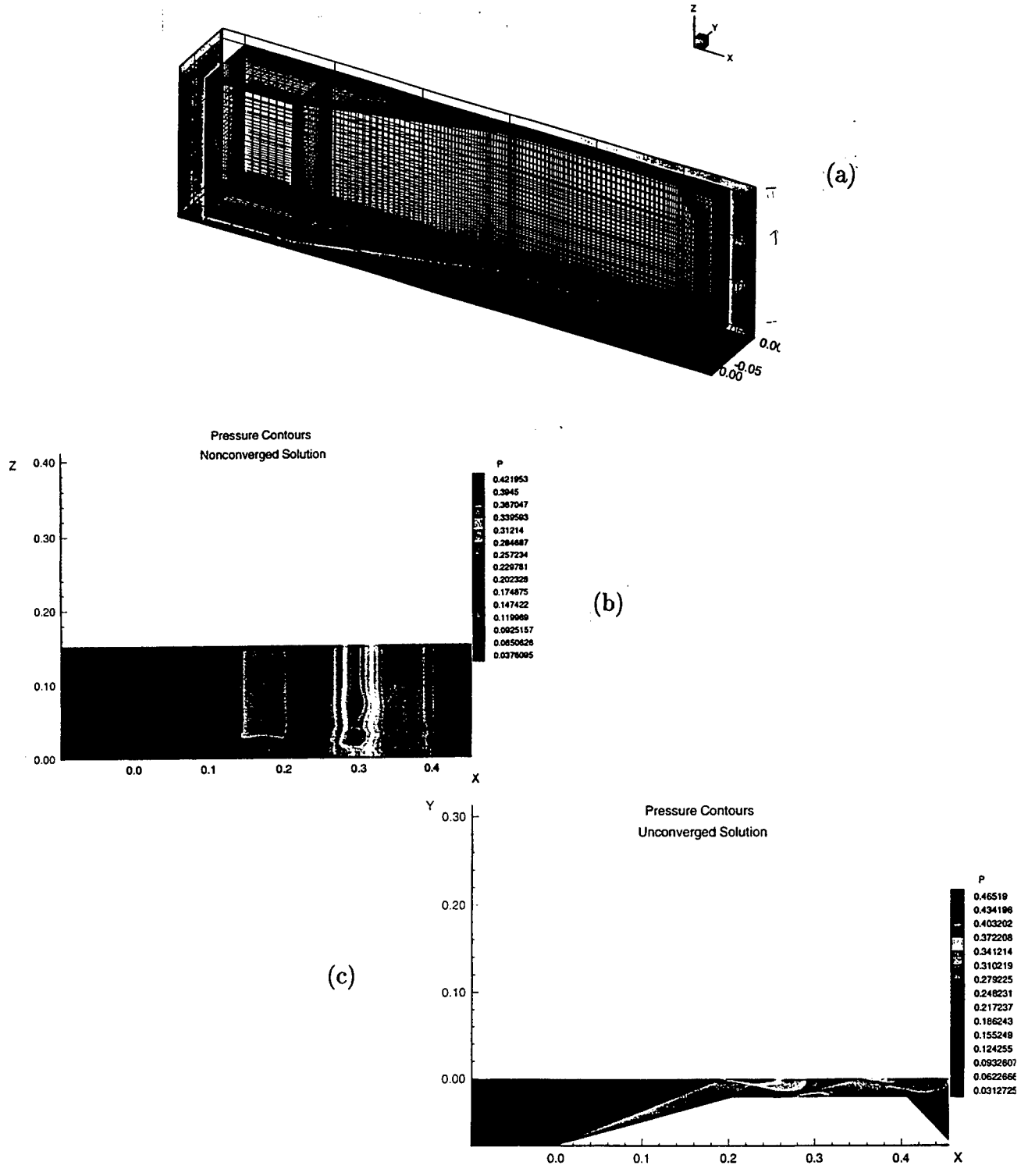


Figure 6: Three-Dimensional Inlet Configuration. a) Grid, b) Nondimensional Pressure Contours on Constant Y Plane, c) Nondimensional Pressure Contours on Constant Z Plane