UNITED STATES AIR FORCE

SUMMER FACULTY RESEARCH PROGRAM
1990
PROGRAM TECHNICAL REPORT
UNIVERSAL ENERGY SYSTEMS, INC.
VOLUME 3 OF 4

PROGRAM DIRECTOR, U. E. S.
RODNEY C. DARRAH
PROGRAM ADMINISTRATOR, U. E. S.
SUSAN K. ESPY

PROGRAM MANAGER, A. F. O. S. R.
LT. COL. CLAUDE CAVENDER

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	5 June 1991	Annual/Final 1 Sep 89-31 Aug 90

4. TITLE AND SUBTITLE
United States Air Force Summer Faculty Research Program 1990 *Volume 3 of 4*
Program Technical Report (w/atch 1 & 2)

5. FUNDING NUMBERS
F49620-88-C-0053

6. AUTHOR(S)
Mr Rodney Darrah

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Universal Energy Systems (UES)
4401 Dayton -Xenia
Dayton OH 45432

AFOSR-TR·

8. PERFORMING ORGANIZATION REPORT NUMBER
01 0960

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NI
Bldg 410
Bolling AFB DC 20332-6448

Lt Col V. Claude Cavender

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The United States Air Force Summer Faculty Research Program (USAF-SFRP) is designed to introduce university, college, and technical institute faculty members to Air Force research. This is accomplished by the faculty members being selected on a nationally advertised competitive basis for a ten-week assignment during the summer intersession period to perform research at Air Force laboratories/centers. Each assignment is in a subject area and at an Air Force facility mutually agreed upon by the faculty members and the Air Force. In addition to compensation, travel and cost of living allowances are also paid. The USAF-SFRP is sponsored by the Air Force Office of Scientific Research.

91-18963

14. SUBJECT TERMS

91 1223 175

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

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)

UNITED STATES AIR FORCE

SUMMER FACULTY RESEARCH PROGRAM

1990

PROGRAM TECHNICAL REPORT

UNIVERSAL ENERGY SYSTEMS, INC.

VOLUME III of IV

Program Director, UES
Rodney C. Darrah

Program Manager, AFOSR
Lt. Col. Claude Cavender

Program Administrator, UES
Susan K. Espy

DTIC
COPY
INSPECTED
6

Submitted to

Air Force Office of Scientific Research

Bolling Air Force Base

Washington, DC

December 1990

TABLE OF CONTENTS

PREFACE

The United States Air Force Summer Faculty Research Program (USAF-SFRP) is designed to introduce university, college, and technical institute faculty members to Air Force research. This is accomplished by the faculty members being selected on a nationally advertised competitive basis for a ten-week assignment during the summer intersession period to perform research at Air Force laboratories/centers. Each assignment is in a subject area and at an Air Force facility mutually agreed upon by the faculty members and the Air Force. In addition to compensation, travel and cost of living allowances are also paid. The USAF-SFRP is sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, United States Air Force, and is conducted by Universal Energy Systems, Inc.

The specific objectives of the 1990 USAF-SFRP are:

(1)   To provide a productive means for U.S. faculty members to participate in research at Air Force Laboratories/Centers;

(2)   To stimulate continuing professional association among the faculty and their professional peers in the Air Force;

(3)   To further the research objectives of the United States Air Force;

(4)   To enhance the research productivity and capabilities of the faculty especially as these relate to Air Force technical interests.

During the summer of 1990, 165-faculty members participated. These researchers were assigned to 23 USAF laboratories/centers across the country. This four volume document is a compilation of the final reports written by the assigned faculty members about their summer research efforts.

LIST OF 1990 PARTICIPANTS

NAME / ADDRESS

DEGREE, SPECIALTY, LABORATORY ASSIGNED

Thomas Abraham
Instructor
Saint Paul's College
Dept. of Science and Math
Lawrenceville, VA 23868
(804) 848-3111

Degree: MS
Specialty: Mathematics
Assigned: Avionics Laboratory

Charles Alajajian
Assistant Professor
West Virginia University
PO Box 6101
Morgantown, WV 26506
(304) 293-6371

Degree: PhD
Specialty: Electrical Engineering
Assigned: Rome Air Development Center

Theodore Aufdemberge
Professor
Concordia College
4090 Geddes Road
Ann Arbor, MI 48105
(313) 985-7349

Degree: PhD
Specialty: Physical Chemistry
Assigned: Geophysics Laboratory

Richard Backs
Assistant Professor
Wright State University
Dept. of Psychology
Dayton, OH 45435
(513) 873-2656

Degree: PhD
Specialty: Psychology
Assigned: Aerospace Medical Research Lab.

William Bannister
Professor
Lowell, University of
Dept. of Chemistry
Lowell, MA 01854
(508) 934-3682

Degree: PhD
Specialty: Organic Chemistry
Assigned: Engineering & Services Center

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED		

Margaret Batschelet
Assistant Professor
Texas-San Antonio, Univ. of
Division of English
San Antonio, TX 78285
(513) 691-5357

Degree: PhD
Specialty: English
Assigned: Human Resources Laboratory
Training Systems

Frank Battles
Professor
Massachusetts Maritime Acad.
Basic Science Dept.
Buzzards Bay, MA 02532
(508) 759-5761

Degree: PhD
Specialty: Physics
Assigned: Geophysics Laboratory

John Bay
Assistant Professor
Virginia Polytech Institute
Dept. of Electrical Eng.
Blacksburg, VA 24061
(703) 231-5114

Degree: PhD
Specialty: Electrical Engineering
Assigned: Flight Dynamics Laboratory

Reuben Benumof
Professor
Staten Island, College of
130 Stuyvesant Pl.
Staten Island, NY 10301
(718) 390-7973

Degree: PhD
Specialty: Physics
Assigned: Geophysics Laboratory

Phillip Bishop
Assistant Professor
Alabama, University of
PO Box 870312
Tuscaloosa, AL 35487
(205) 348-8370

Degree: PhD
Specialty: Exercise Physiology
Assigned: School of Aerospace Medicine

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

Robert Blystone
Professor
Trinity University
715 Stadium Dr.
San Antonio, TX 78212
(512) 736-7243

Degree: PhD
Specialty: Zoology
Assigned: School of Aerospace Medicine

Michael Breen
Assistant Professor
Alfred University
Myers Hall
Alfred, NY 14802
(607) 871-2258

Degree: PhD
Specialty: Mathematics
Assigned: Avionics Laboratory

Bruno Breitmeyer
Professor
Houston, University of
Dept. of Psychology
Houston, TX 77204
(713) 749-6108

Degree: PhD
Specialty: Experimental Psychology
Assigned: School of Aerospace Medicine

Mark Brusseau
Assistant Professor
Arizona, University of
429 Shantz Bldg. #38
Tucson, AZ 85721
(602) 621-3244

Degree: PhD
Specialty: Environmental Chemistry
Assigned: Engineering & Services Center

David Buckalew
Assistant Professor
Xavier University
7325 Palmetto St.
New Orleans, LA 70125
(504) 483-7527

Degree: PhD
Specialty: Biology
Assigned: Occupational & Environmental
Health Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED		

Theodore Burkey
Assistant Professor
Memphis State University
Chemistry Dept.
Memphis, TN 38152
(901) 678-2634

Degree: PhD
Specialty: Chemistry
Assigned: Frank J. Seiler Research Lab.

Larry Byrd
Assistant Professor
Arkansas State University
PO Box 1740
State University, AR 72467
(501) 972-2088

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Aerospace Medical Research Lab.

Charles Camp
Assistant Professor
Memphis State University
Civil Engineering Dept.
Memphis, TN 38152
(901) 678-3169

Degree: PhD
Specialty: Civil Engineering
Assigned: Armament Laboratory

William Campbell
Associate Professor
Talladega College
Math Dept.
Talladega, AL 35160
(205) 362-0206

Degree: PhD
Specialty: Mathematics
Assigned: Weapons Laboratory

Arnold Carden
Professor
Alabama, University of
PO Box 870278
Tuscaloosa, AL 35487
(205) 348-1619

Degree: PhD
Specialty: Metallurgy
Assigned: Armament Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

Richard Carlin
Assistant Professor
Alabama, University of
Dept. of Chemistry
Tuscaloosa, AL 35487
(205) 348-8443

Degree: PhD
Specialty: Chemistry
Assigned: Frank J. Seiler Research Lab.

Gene Carlisle
Professor
West Texas State University
Dept. of Chemistry & Physics
Canyon, TX 79016
(806) 656-2282

Degree: PhD
Specialty: Inorganic Chemistry
Assigned: Weapons Laboratory

Chia-Bo Chang
Associate Professor
Texas Tech. Univ.
PO Box 4320
Lubbock, TX 79409
(806) 742-3143

Degree: PhD
Specialty: Meteorology
Assigned: Geophysics Laboratory

Wayne Charlie
Associate Professor
Colorado State University
Dept. of Civil Engineering
Fort Collins, CO 80523
(303) 491-5048

Degree: PhD
Specialty: Civil Engineering
Assigned: Engineering & Services Center

Chih-Fan Chen
Professor
Boston University
755 Commonwealth Ave.
Boston, MA 02215
(617) 353-2566

Degree: PhD
Specialty: Engineering
Assigned: Electronic Systems Division

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED	

Pinyuen Chen
Associate Professor
Syracuse University
Dept. of Mathematics
Syracuse, NY 13244
(315) 443-1573

Degree: PhD
Specialty: Statistics
Assigned: Human Resources Laboratory
Manpower and Personnel

Muhammad Choudhry
Associate Professor
West Virginia University
PO Box 6101
Morgantown, WV 26506
(304) 293-6375

Degree: PhD
Specialty: Electrical Engineering
Assigned: Aero Propulsion Laboratory

Donald Chung
Associate Professor
San Jose State University
Dept. of Materials Eng.
San Jose, CA 95192
(408) 924-3873

Degree: PhD
Specialty: Materials Science
Assigned: Materials Laboratory

Mingking Chyu
Assistant Professor
Carnegie Mellon University
Dept. of Mechanical Eng.
Pittsburgh, PA 15213
(412) 268-3658

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Aero Propulsion Laboratory

R. H. Cofer
Associate Professor
Florida Instit. of Tech.
150 W. University Blvd.
Melbourne, FL 32901
(407) 768-8000

Degree: PhD
Specialty: Electrical Engineering
Assigned: Avionics Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED	

William Cofer
Assistant Professor
Washington State University
Dept. of Civil & Environ. Eng.
Pullman, WA 99164
(509) 335-3232

Degree: PhD
Specialty: Civil Engineering
Assigned: Weapons Laboratory

John Connolly
Professor
Missouri-Kansas City, Univ. of
Dept. of Chemistry
Kansas City, MO 64110
(816) 276-2286

Degree: PhD
Specialty: Chemistry
Assigned: Materials Laboratory

Gary Craig
Assistant Professor
Syracuse University
Link Hall
Syracuse, NY 13244
(315) 443-4389

Degree: PhD
Specialty: Electrical Engineering
Assigned: Rome Air Development Center

Donald Dareing
Professor
Florida, University of
237 MEB
Gainesville, FL 32611
(904) 392-0827

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Aero Propulsion Laboratory

Vito DelVecchio
Professor
Scranton, University of
Dept. of Biology
Scranton, PA 18510
(717) 961-6117

Degree: PhD
Specialty: Biochemistry
Assigned: School of Aerospace Medicine

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

Paul Dellenback
Assistant Professor
Southern Methodist Univ.
Civil & Mech. Engineering Dept.
Dallas, TX 75275
(214) 692-4172

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Aero Propulsion Laboratory

Eustace Dereniak
Associate Professor
Arizona, University of
McKale Ave.
Tucson, AZ 85721
(602) 621-1019

Degree: PhD
Specialty: Optics
Assigned: Armament Laboratory

Janet Dizinno
Assistant Professor
St. Mary's University
One Camino Santa Maria
San Antonio, TX 78284
(512) 436-3314

Degree: PhD
Specialty: Psychology
Assigned: Wilford Hall Medical Center

Daniel Dolata
Assistant Professor
Arizona, University of
Dept. of Chemistry
Tucson, AZ 85721
(602) 621-6337

Degree: PhD
Specialty: Chemistry
Assigned: Frank J. Seiler Research Lab.

Joseph Dreisbach
Professor
Scranton, University of
Chemistry Dept.
Scranton, PA 18510
(717) 961-7519

Degree: PhD
Specialty: Chemistry
Assigned: Engineering & Services Center

THIS
PAGE
IS
MISSING
IN
ORIGINAL
DOCUMENT

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

Dennis Flentge
Associate Professor
Cedarville College
Box 601
Cedarville, OH 45314
(513) 766-2211

Degree: PhD
Specialty: Physical Chemistry
Assigned: Aero Propulsion Laboratory

Charles Fosha
Associate Professor
Colorado, Univ. of
1867 Austin Bluffs Parkway
Colorado Springs, CO 80918
(719) 548-0602

Degree: PhD
Specialty: Electrical Engineering
Assigned: Armament Laboratory

Lionel Friedman
Professor
Worcester Polytechnic Instit.
100 Institute Rd.
Worcester, MA 01609
(508) 831-5303

Degree: PhD
Specialty: Physics
Assigned: Rome Air Development Center

Daniel Fuller
Department Head
Nicholls State University
Highway 1
Thibodaux, LA 70310
(504) 448-4504

Degree: PhD
Specialty: Chemistry
Assigned: Astronautics Laboratory

Ephrahim Garcia
Assistant Professor
New York-Buffalo, State Univ. of
1012 Furnas Hall
Buffalo, NY 14260
(716) 636-3058

Degree: PhD
Specialty: Aerospace Engineering
Assigned: Frank J. Seiler Research Lab.

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED	

Daniel Garland
Assistant Professor
Embry-Riddle Aeronautical Univ.
Humanities/Social Sci.
Daytona Beach, FL 32114
(904) 239-6641

Degree:	PhD	
Specialty:	Psychology	
Assigned:	Human Resources Laboratory	
	Operations Training Division	

Thomas Gearhart
Associate Professor
Captial University
Science Hall E. Main St.
Columbus, OH 43209
(614) 236-6800

Degree:	PhD
Specialty:	Mathematics
Assigned:	Avionics Laboratory

John George
Professor
Wyoming, University of
Box 3036
Laramie, WY 82071
(307) 766-2383

Degree:	PhD
Specialty:	Applied Mathematics
Assigned:	Armament Laboratory

Frederick Gibson
Instructor
Morehouse College
830 Westview Dr. SW
Atlanta, GA 30312
(404) 681-2800

Degree:	MS
Specialty:	Applied Mathematics
Assigned:	Armament Laboratory

Ashok Goel
Assistant Professor
Michigan Tech. University
Dept. of Electrical Engineering
Houghton, MI 49931
(906) 487-2868

Degree:	PhD
Specialty:	Electrical Engineering
Assigned:	Electronic Technology Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED	

Harold Goldstein
Associate Professor
District of Columbia, Univ. of
4200 Connecticut Ave. N.W.
Washington, DC 20008
(202) 282-7349

Degree: MS
Specialty: Transportation Engineering
Assigned: Human Resources Laboratory
Training Systems

Reinhard Graetzer
Associate Professor
Penn State University
104 Davey Lab.
University Park, PA 16802
(814) 863-0705

Degree: PhD
Specialty: Physics
Assigned: School of Aerospace Medicine

Paul Griffin
Assistant Professor
Georgia Tech.
School of ISYE
Atlanta, GA 30332
(404) 894-2431

Degree: PhD
Specialty: Industrial Engineering
Assigned: School of Aerospace Medicine

William Grissom
Assistant Professor
Morehouse College
830 Westview Dr.
Atlanta, GA 30314
(404) 681-2800

Degree: MS
Specialty: Mechanical Engineering
Assigned: Arnold Engineering Development
Center

David Grossie
Assistant Professor
Wright State University
Dept. of Chemistry
Dayton, OH 45435
(513) 873-2210

Degree: PhD
Specialty: Chemistry
Assigned: Materials Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED	

Pushpa Gupta
Professor
Maine, University of
321 Neville
Orono, ME 04469
(207) 581-3914

Degree: PhD
Specialty: Mathematics
Assigned: School of Aerospace Medicine

Ramesh Gupta
Professor
Maine, University of
Dept. of Mathematics
Orono, ME 04469
(207) 581-3913

Degree: PhD
Specialty: Mathematical Statistics
Assigned: School of Aerospace Medicine

Martin Hagan
Associate Professor
Oklahoma State University
School of Elec. & Comp. Sci.
Stillwater, OK 74078
(405) 744-7340

Degree: PhD
Specialty: Electrical Engineering
Assigned: Aerospace Medical Research Lab.

Lawrence Hall
Assistant Professor
South Florida, Univ. of
Dept. of Computer Sci.
Tampa, FL 33620
(813) 974-4195

Degree: PhD
Specialty: Computer Science
Assigned: Avionics Laboratory

Kevin Hallinan
Assistant Professor
Dayton, Univ. of
Mech. & Aero. Engineering
Dayton, OH 45469
(513) 229-2875

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Aero Propulsion Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED	

Marvin Hamstad
Professor
Denver, Univ. of
Dept. of Engineering
Denver, CO 80208
(303) 871-3191

Degree: PhD
Specialty: Solid Mechanics
Assigned: Flight Dynamics Laboratory

Frances Harackiewicz
Assistant Professor
Southern Illinois Univ.
Technology Bldg. A
Carbondale, IL 62901
(618) 453-7031

Degree: PhD
Specialty: Electrical Engineering
Assigned: Rome Air Development Center

Paul Hedman
Professor
Brigham Young University
Chemical Engineering Dept.
Provo, UT 84602
(801) 378-6238

Degree: PhD
Specialty: Chemical Engineering
Assigned: Aero Propulsion Laboratory

Verlin Hinsz
Assistant Professor
North Dakota State Univ.
115 Minard Hall
Fargo, ND 58105
(701) 237-7082

Degree: PhD
Specialty: Psychology
Assigned: Human Resources Laboratory
Logistics & Human Factors

Chin Hsu
Associate Professor
Washington State Univ.
Dept. of Elec. and Comp. Eng.
Pullman, WA 99164
(509) 335-2342

Degree: PhD
Specialty: Electrical Engineering
Assigned: Flight Dynamics Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED	

Ming-Shu Hsu
Associate Professor
Portland, Univ. of
5000 N. Willamette Blvd.
Portland, OR 97203
(503) 283-7436

Degree:	PhD
Specialty:	Mechanical Engineering
Assigned:	Flight Dynamics Laboratory

Delayne Hudspeth
Associate Professor
Texas-Austin, Univ. of
College of Education
Austin, TX 78712
(512) 471-5211

Degree:	PhD
Specialty:	Education
Assigned:	Human Resources Laboratory Manpower & Personnel Div.

Manuel Huerta
Professor
Miami, Univ. of
PO Box 248046
Coral Gables, FL 33124
(305) 284-2323

Degree:	PhD
Specialty:	Physics
Assigned:	Armament Laboratory

David Hui
Associate Professor
New Orleans, Univ. of
Dept. of Mech. Engineering
New Orleans, LA 70148
(504) 286-6192

Degree:	PhD
Specialty:	Aerospace Engineering
Assigned:	Flight Dynamics Laboratory

George Jumper
Associate Professor
Worcester Poly. Instit.
100 Institute Rd.
Worcester, MA 01609
(508) 831-5368

Degree:	PhD
Specialty:	Mechanical Engineering
Assigned:	Geophysics Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED	

Prasad Kadaba
Professor
Kentucky, Univ. of
Rm. 453 Anderson Hall
Lexington, KY 40506
(606) 257-2943

Degree: PhD
Specialty: Physics
Assigned: Materials Laboratory

Ngozi Kamalu
Lecturer
California Poly. Univ.
Mechancial Engineering Dept.
San Luis Obispo, CA 93407
(805) 756-1336

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Frank J. Seiler Research Lab.

Gillray Kandel
Professor
Rensselaer Poly. Instit.
8th St.
Troy, NY 12180
(518) 276-8269

Degree: PhD
Specialty: Experimental Psychology
Assigned: Human Resources Laboratory
Operations Training Division

Mohammad Karim
Associate Professor
Dayton, University of
300 College Park
Dayton, OH 45469
(513) 229-3611

Degree: PhD
Specialty: Electrical Engineering
Assigned: Avionics Laboratory

Siavash Kassemi
Assistant Professor
Colorado, University of
1867 Austin Bluffs Parkway
Colorado Springs, CO 80918
(719) 593-3326

Degree: PhD
Specialty: Aerospace Engineering
Assigned: Frank J. Seiler Research Lab.

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

Yulian Kin
Associate Professor
Purdue Calumet
Engineering Dept.
Hammond, IN 46323
(219) 989-2684

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Flight Dynamics Laboratory

Kevin Kirby
Assistant Professor
Wright State University
3171 Research Blvd.
Kettering, OH 45420
(513) 259-1373

Degree: PhD
Specialty: Computer Science
Assigned: Avionics Laboratory

David Kirkner
Associate Professor
Notre Dame, Univ. of
Dept. of Civil Engineering
Notre Dame, IN 46556
(219) 239-6518

Degree: PhD
Specialty: Solid Mechanics
Assigned: Engineering & Services Center

Ashok Krishnamurthy
Assistant Professor
Ohio State University
2015 Neil Ave.
Columbus, OH 43210
(614) 292-5604

Degree: PhD
Specialty: Electrical Engineering
Assigned: Aerospace Medical Research Lab.

Paul Kromann
Associate Professor
Fort Valley State College
Campus Box 4821
Fort Valley, GA 31030
(912) 825-6245

Degree: PhD
Specialty: Chemistry
Assigned: Engineering & Services Center

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

Jeffrey Kuhn
Associate Professor
Michigan State University
309 Physics-Astronomy Bldg.
East Lansing, MI 48824
(517) 353-2986

Degree: PhD
Specialty: Physics
Assigned: Geophysics Laboratory

Kyung Kwon
Associate Professor
Tuskege University
Chemistry Dept.
Tuskegee, AL 36088
(205) 727-8089

Degree: PhD
Specialty: Chemical Engineering
Assigned: Engineering & Services Center

Joseph Lambert
Professor
Northwestern University
2145 Sheridan Rd.
Evanston, IL 60208
(708) 491-5437

Degree: PhD
Specialty: Chemistry
Assigned: Materials Laboratory

Gary Leatherman
Assistant Professor
Worcester Polytechnic Instit.
100 Institute Rd.
Worcester, MA 01609
(508) 831-5229

Degree: PhD
Specialty: Materials Science
Assigned: Materials Laboratory

Byung-Lip Lee
Associate Professor
Pennsylvania State University
227 Hammond Bldg.
University Park, PA 16802
(814) 865-7829

Degree: PhD
Specialty: Materials Science
Assigned: Flight Dynamics Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

Tzesan Lee
Associate Professor
Western Illinois Univ.
900 W. Adams St.
Macomb, IL 61455
(309) 298-1485

Degree: PhD
Specialty: Applied Mathematics
Assigned: Aerospace Medical Research Lab.

Won-Kyoo Lee
Associate Professor
Ohio State University
140 W. 19th Ave.
Columbus, OH 43211
(614) 292-6605

Degree: PhD
Specialty: Chemical Engineering
Assigned: Materials Laboratory

Paul Lemke
Professor
Auburn University
131 Funchess Hall
Auburn University, AL 36849
(205) 844-1662

Degree: PhD
Specialty: Molecular Biology
Assigned: School of Aerospace Medicine

Sigmund Lephart
Lecturer
Melbourne Univ. Australia
Parkville 3052
Victoria Australia,
(03) 344-5158

Degree: PhD
Specialty: Biomechanics
Assigned: Aerospace Medical Research Lab.

Shannon Lieb
Associate Professor
Butler University
4600 Sunset Ave.
Indianapolis, IN 46208
(317) 283-9410

Degree: PhD
Specialty: Physical Chemistry
Assigned: Astronautics Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED		

Hao Ling
Assistant Professor
Texas-Austin, Univ. of
Dept. of Elec. & Comp. Eng.
Austin, TX 78712
(512) 471-1710

Degree: PhD
Specialty: Electrical Engineering
Assigned: Rome Air Development Center

C. Randal Lishawa
Assistant Professor
Utica College
Burrstone Rd.
Utica, NY 13502
(315) 792-3139

Degree: PhD
Specialty: Physical Chemistry
Assigned: Geophysics Laboratory

Vernon Matzen
Associate Professor
North Carolina State Univ.
Box 7908
Raleigh, NC 27695
(919) 737-2331

Degree: PhD
Specialty: Structural Mechanics
Assigned: Flight Dynamics Laboratory

Michael McFarland
Assistant Professor
Utah State Univ.
Utah Water Research Lab.
Logan, UT 84322
(801) 750-3196

Degree: PhD
Specialty: Biological Engineering
Assigned: Engineering & Services Center

Perry McNeill
Professor
North Texas, Univ. of
PO Box 13198
Denton, TX 76203
(817) 565-2846

Degree: PhD
Specialty: Education
Assigned: Engineering & Services Center

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED	

Miguel Medina
Associate Professor
Duke University
Dept. of Civil Engineering
Durham, NC 27706
(919) 660-5195

Degree:	PhD
Specialty:	Water Resources
Assigned:	Occupational & Environmental Health Laboratory

Richard Miers
Associate Professor
Indiana Univ. - Purdue Univ.
2101 Coliseum Blvd. E.
Fort Wayne, IN 46805
(219) 481-6154

Degree:	PhD
Specialty:	Physics
Assigned:	Avionics Laboratory

William Moor
Associate Professor
Arizona State Univ.
College of Engineering
Tempe, AZ 85287
(602) 965-4022

Degree:	PhD
Specialty:	Industrial Engineering
Assigned:	Human Resources Laboratory Operations Training Division

Carlyle Moore
Associate Professor
Morehouse College
830 Westview Dr.
Atlanta, GA 30314
(404) 681-2800

Degree:	PhD
Specialty:	Physics
Assigned:	Arnold Engineering Development Center

Kevin Moore
Assistant Professor
Idaho State Univ.
Box 8060
Pocatello, ID 83209
(208) 236-4188

Degree:	PhD
Specialty:	Electrical Engineering
Assigned:	Armament Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

Rex Moyer
Professor
Trinity University
715 Stadium Dr.
San Antonio, TX 78212
(512) 736-7242

Degree: PhD
Specialty: Microbiology
Assigned: School of Aerospace Medicine

Arnold Nelson
Assistant Professor
Louisiana State Univ.
112 Long Field House
Baton Rouge, LA 70803
(504) 388-3114

Degree: MS
Specialty: Physical Education
Assigned: School of Aerospace Medicine

Kirk Nordyke
Instructor
Xavier University
Dept. of Biology
New Orleans, LA 70125
(504) 483-7527

Degree: MS
Specialty: Zoology
Assigned: Occupational & Environmental
Health Laboratory

Olin Norton
Researcher
Mississippi State Univ.
PO Drawer MM
Mississippi State, MS 39762
(601) 325-2105

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Arnold Engineering Development
Center

Muhammad Numan
Assistant Professor
Indiana Univ. of Pennsylvania
45 Weyandt Hall
Indiana, PA 15705
(412) 357-2318

Degree: PhD
Specialty: Physics
Assigned: Electronic Technology Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

Amit Patra
Associate Professor
Puerto Rico, Univ. of
PO Box 5000
Mayaguez, PR 00709
(809) 832-4040

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Aerospace Medical Research Lab.

Shietung Peng
Assistant Professor
Maryland-Baltimore, Univ. of
5401 Wilkens Ave.
Baltimore, MD 21228
(301) 455-3540

Degree: PhD
Specialty: Computer Science
Assigned: Rome Air Development Center

Richard Peters
Assistant Professor
Vanderbilt University
Box 6091 Station B
Nashville, TN 37235
(615) 322-7924

Degree: PhD
Specialty: Electrical Engineering
Assigned: Arnold Engineering Development
Center

Bernard Piersma
Professor
Houghton College
Dept. of Chemistry
Houghton, NY 14744
(716) 567-9301

Degree: PhD
Specialty: Physical Chemistry
Assigned: Frank J. Seiler Research Lab.

Thomas Pollock
Associate Professor
Texas A&M University
Dept. of Aerospace Engineering
College Station, TX 77843
(409) 845-1686

Degree: PhD
Specialty: Materials Science
Assigned: Astronautics Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

Thomas Posbergh
Assistant Professor
Minnesota, Univ. of
107 Akerman Hall
Minneapolis, MN 55455
(612) 625-2871

Degree: PhD
Specialty: Electrical Engineering
Assigned: Frank J. Seiler Research Lab.

James Price
Professor
Iowa, University of
W140 Seashore Hall
Iowa City, IA 52242
(319) 335-2497

Degree: PhD
Specialty: Sociology
Assigned: Human Resources Laboratory
Manpower & Personnel Div.

Gandikota Rao
Professor
St. Louis University
3507 Laclede Ave.
St. Louis, MO 63103
(314) 658-3115

Degree: PhD
Specialty: Meteorology
Assigned: Geophysics Laboratory

K. Sankara Rao
Professor
North Dakota State Univ.
Dept. of Electrical Engineering
Fargo, ND 58105
(701) 237-7217

Degree: PhD
Specialty: Electrical Engineering
Assigned: Aero Propulsion Laboratory

Craig Rasmussen
Assistant Professor
Utah State University
CASS UMC 4405
Logan, UT 84322
(801) 750-2967

Degree: PhD
Specialty: Physics
Assigned: Geophysics Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED	

Joan Rentsch
Assistant Professor
Wright State University
Dept. of Psychology
Dayton, OH 45435
(513) 873-2218

Degree: PhD
Specialty: Industrial Psychology
Assigned: Human Resources Laboratory
Logistics & Human Factors

Michael Resch
Assistant Professor
Nebraska-Lincoln, Univ. of
212 Bancroft Hall
Lincoln, NE 68588
(402) 472-2354

Degree: PhD
Specialty: Materials Science
Assigned: Materials Laboratory

Donald Robinson
Assistant Professor
Xavier University
7325 Palmetto St.
New Orleans, LA 70125
(504) 483-7371

Degree: PhD
Specialty: Chemistry
Assigned: School of Aerospace Medicine

Larry Roe
Assistant Professor
Virginia Poly. Instit. State Univ.
Mechanical Engineering Dept.
Blacksburg, VA 24061
(703) 231-7295

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Aero Propulsion Laboratory

John Russell
Associate Professor
Florida Inst. of Tech.
150 W. University Blvd.
Melbourne, FL 32901
(407) 768-8000

Degree: PhD
Specialty: Aerospace Engineering
Assigned: Arnold Engineering Development
Center

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

Daniel Ryder
Assistant Professor
Tufts University
Chemical Engineering Dept.
Medford, MA 02155
(617) 381-3446

Degree: PhD
Specialty: Chemical Engineering
Assigned: Rome Air Development Center

John Scharf
Chairman
Carroll College
Dept. of Math.
Helena, MT 59625
(406) 442-3450

Degree: MS
Specialty: Civil Engineering
Assigned: Engineering & Services Center

Johanna Schruben
Associate Professor
Houston-Victoria, Univ. of
2302C Red River
Victoria, TX 77901
(512) 576-3151

Degree: PhD
Specialty: Mathematics
Assigned: Weapons Laboratory

Martin Schwartz
Professor
North Texas, Univ. of
PO Box 5068
Denton, TX 76203
(817) 565-3524

Degree: PhD
Specialty: Physical Chemistry
Assigned: Materials Laboratory

David Senseman
Professor
Texas-San Antonio, Univ. of
Div. of Life Sciences
San Antonio, TX 78285
(512) 691-5485

Degree: PhD
Specialty: Biology
Assigned: School of Aerospace Medicine

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

Brian Shelburne
Associate Professor
Wittenberg University
Box 720
Springfield, OH 45501
(513) 327-7862

Degree: PhD
Specialty: Mathematics
Assigned: Avionics Laboratory

Behrooz Shirazi
Assistant Professor
Southern Methodist University
Dept. of Comp. Sci. & Engineering
Dallas, TX 75275
(214) 692-2874

Degree: PhD
Specialty: Computer Science
Assigned: Rome Air Development Center

Leonard Shyles
Associate Professor
Villanova University
Dept. of Communication Arts
Villanova, PA 19085
(215) 645-7923

Degree: PhD
Specialty: Communication
Assigned: Aerospace Medical Research Lab.

William Siuru
Associate Professor
Colorado, Univ. of
1867 Austin Bluffs Parkway
Colorado Springs, CO 80918
(719) 548-0602

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Armament Laboratory

Eleanor Smith
Assistant Professor
Florida A&M University
406 Perry-Paige
Tallahassee, FL 32307
(904) 599-3821

Degree: PhD
Specialty: Sociology
Assigned: Human Resources Laboratory
Training Systems Division

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED	

Wayne Smith
Associate Professor
Mississippi State University
Drawer CS
Mississippi State, MS 39762
(601) 325-2642

Degree: PhD
Specialty: Computer Science
Assigned: Rome Air Development Center

Kenneth Sobel
Associate Professor
New York, City College of
Dept. of Electrical Engineering
New York, NY 10031
(212) 690-4241

Degree: PhD
Specialty: Electrical Engineering
Assigned: Armament Laboratory

Glenn Stark
Assistant Professor
Wellesley College
Dept. of Physics
Wellesley, MA 02181
(617) 235-0320

Degree: PhD
Specialty: Physics
Assigned: Geophysics Laboratory

Stanley Stephenson
Associate Professor
Southwest Texas State Univ.
CIS/ADS
San Marcos, TX 78666
(512) 245-2291

Degree: PhD
Specialty: Psychology
Assigned: Human Resources Laboratory
Training Systems Division

Chun Su
Assistant Professor
Mississippi State Univ.
Dept. of Physics
Mississippi State, MS 39762
(601) 325-2931

Degree: PhD
Specialty: Physics
Assigned: Arnold Engineering Development Center

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED		

Richard Swope
Professor
Trinity University
715 Stadium Dr.
San Antonio, TX 78212
(512) 736-7514

Degree:	PhD
Specialty:	Mechanical Engineering
Assigned:	School of Aerospace Medicine

John Szarek
Assistant Professor
Marshall University
1542 Spring Valley Dr.
Huntington, WV 25755
(304) 696-7314

Degree:	PhD
Specialty:	Pharmaceutical
Assigned:	School of Aerospace Medicine

Kaveh Tagavi
Associate Professor
Kentucky, Univ. of
242 Anderson Hall
Lexington, KY 40506
(606) 257-2739

Degree:	PhD
Specialty:	Mechanical Engineering
Assigned:	Aero Propulsion Laboratory

Devki Talwar
Assistant Professor
Indiana Univ. of Pennsylvania
Dept. of Physics
Indiana, PA 15705
(412) 357-4589

Degree:	PhD
Specialty:	Physics
Assigned:	Electronic Technology Laboratory

Richard Tankin
Professor
Northwestern University
Dept. of Mechanical Engineering
Evanston, IL 60208
(708) 491-3532

Degree:	PhD
Specialty:	Engineering
Assigned:	Aero Propulsion Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

Roger Thompson
Assistant Professor
Pennsylvania State Univ.
233 Hammond Bldg.
University Park, PA 16802
(814) 863-0968

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Astronautics Laboratory

Steven Trogdon
Associate Professor
Minnesota, Univ. of
108 Heller Hall
Duluth, MN 55812
(218) 726-6173

Degree: PhD
Specialty: Mechanics
Assigned: Armament Laboratory

Hai-Lung Tsai
Assistant Professor
Missouri-Rolla, Univ. of
Dept. of Mech. & Aero.
Rolla, MO 65401
(314) 341-4945

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Materials Laboratory

Pamela Tsang
Assistant Professor
Wright State University
309 Oelman
Dayton, OH 45435
(513) 258-2687

Degree: PhD
Specialty: Engineering Psychology
Assigned: Human Resources Laboratory
Logistics & Human Factors

Ronald VanEtten
Associate Professor
Illinois State University
500 W. Gregory
Normal, IL 61761
(309) 438-8346

Degree: MS
Specialty: Computer Science
Assigned: Rome Air Development Center

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED	

George Veyera
Assistant Professor
Rhode Island, Univ. of
Dept. of Civil Engineering
Kingston, RI 02881
(401) 792-2684

Degree: PhD
Specialty: Civil Engineering
Assigned: Engineering & Services Center

Hung Vu
Assistant Professor
California State Univ.
1250 Bellflower Blvd.
Long Beach, CA 90840
(213) 985-1524

Degree: PhD
Specialty: Applied Mechanics
Assigned: Astronautics Laboratory

Bonnie Walker
Assistant Professor
Central State University
Psychology Dept.
Wilberforce, OH 45384
(513) 376-6516

Degree: PhD
Specialty: Experimental Psychology
Assigned: Aerospace Medical Research Lab.

Steven Waller
Associate Professor
South Dakota, Univ. of
Dept. of Physiol. & Pharmacol.
Vermillion, SD 57069
(605) 677-5157

Degree: PhD
Specialty: Pharmacology
Assigned: School of Aerospace Medicine

Peter Walsh
Professor
Fairleigh Dickinson Univ.
Dept. of Electrical Engineering
Teaneck, NJ 07666
(201) 692-2493

Degree: PhD
Specialty: Physics
Assigned: Weapons Laboratory

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED	

Lorin Weber
Professor
Ricks College
Physics Dept.
Rexburg, ID 83460
(208) 356-1907

Degree: MS
Specialty: Science Education
Assigned: Occupational & Environmental
Health Laboratory

Kevin Whitaker
Assistant Professor
Alabama, University of
Box 870280
Tuscaloosa, AL 35487
(205) 348-7366

Degree: PhD
Specialty: Aerospace Engineering
Assigned: Arnold Engineering Development
Center

Trevor Williams
Assistant Professor
Cincinnati, University of
ML 70
Cincinnati, OH 45221
(513) 556-3221

Degree: PhD
Specialty: Control Theory
Assigned: Astronautics Laboratory

John Wills
Professor
Indiana University
Physics Dept.
Bloomington, IN 47405
(812) 855-1479

Degree: PhD
Specialty: Physics
Assigned: Geophysics Laboratory

Martin Wilner
Professor
Lowell, University of
1 University Ave.
Lowell, MA 01854
(508) 934-3786

Degree: PhD
Specialty: Physics
Assigned: Rome Air Development Center

NAME / ADDRESS	DEGREE, SPECIALTY, LABORATORY ASSIGNED

William Wolfe
Associate Professor
Ohio State University
470 Hitchcock Hall
Columbus, OH 43210
(614) 292-0790

Degree: PhD
Specialty: Engineering
Assigned: Flight Dynamics Laboratory

James Wolper
Assistant Professor
Hamilton College
Dept. of Math & Comp. Sci.
Clinton, NY 13323
(315) 859-4417

Degree: PhD
Specialty: Mathematics
Assigned: Rome Air Development Center

Hsien-Yang Yeh
Associate Professor
California State Univ.
1250 Bellflower Blvd.
Long Beach, CA 90840
(213) 985-4611

Degree: PhD
Specialty: Structural Mechanics
Assigned: Astronautics Laboratory

Lawrence Zavodney
Assistant Professor
Ohio State Univ.
209 Boyd Laboratory
Columbus, OH 43210
(614) 292-2209

Degree: PhD
Specialty: Mechanical Engineering
Assigned: Flight Dynamics Laboratory

Wayne Zimmermann
Associate Professor
Texas Woman's University
PO Box 22865
Denton, TX 76204
(817) 898-2166

Degree: PhD
Specialty: Applied Mathematics
Assigned: Weapons Laboratory

PARTICIPANT LABORATORY ASSIGNMENT

C. PARTICIPANT LABORATORY ASSIGNMENT (Page 1)

1990 USAF/UES SUMMER FACULTY RESEARCH PROGRAM

AERO PROPULSION LABORATORY (WRDC/APL)
(Wright-Patterson Air Force Base)

1.	Muhammad Choudhry	7.	Paul Hedman
2.	Mingking Chyu	8.	K. Sankara Rao
3.	Donald Dareing	9.	Larry Roe
4.	Paul Dellenback	10.	Kaveh Tagavi
5.	Dennis Flentge	11.	Richard Tankin
6.	Kevin Hallinan		

ARMAMENT LABORATORY (ATL)
(Eglin Air Force Base)

1.	Charles Camp	7.	Manuel Huerta
2.	Arnold Carden	8.	Kevin Moore
3.	Eustace Dereniak	9.	William Siuru
4.	Charles Fosha	10.	Kenneth Sobel
5.	John George	11.	Steven Trogdon
6.	Frederick Gibson		

HARRY G. ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY (AAMRL)
(Wright-Patterson AFB)

1.	Richard Backs	6.	Tzesan Lee
2.	Larry Byrd	7.	Sigmund Lephart
3.	John Duncan	8.	Amit Patra
4.	Martin Hagan	9.	Leonard Shyles
5.	Ashok Krishnamurthy	10.	Bonnie Walker

ARNOLD ENGINEERING DEVELOPMENT CENTER (AEDC)
(Arnold Air Force Base)

1.	William Grissom	5.	John Russell
2.	Carlyle Moore	6.	Chun Su
3.	Olin Norton	7.	Kevin Whitaker
4.	Richard Peters		

ASTRONAUTICS LABORATORY (AL)
Edwards Air Force Base)

1.	Daniel Fuller	5.	Hung Vu
2.	Shannon Lieb	6.	Trevor Williams
3.	Thomas Pollock	7.	Hsien-Yang Yeh
4.	Roger Thompson		

C. PARTICIPANT LABORATORY ASSIGNMENT (Page 2)

AVIONICS LABORATORY (WRDC/AL)
(Wright-Patterson Air Force Base)

1. Thomas Abraham
2. Michael Breen
3. R. H. Cofer
4. Thomas Gearhart
5. Lawrence Hall
6. Mohammad Karim
7. Kevin Kirby
8. Richard Miers
9. Brian Shelburne

ELECTRONIC SYSTEMS DIVISION (ESD)
(Hanscom Air Force Base)

1. Chih-Fan Chen

ELECTRONIC TECHNOLOGY LABORATORY (WRDC/ETL)
(Wright-Patterson Air Force Base)

1. Ashok Goel
2. Muhammad Numan
3. Devki Talwar

ENGINEERING AND SERVICES CENTER (ESC)
(Tyndall Air Force Base)

1. William Bannister
2. Mark Brusseau
3. Wayne Charlie
4. Joseph Dreisbach
5. David Kirkner
6. Paul Kromann
7. Kyung Kwon
8. Michael McFarland
9. Perry McNeill
10. John Scharf
11. George Veyera

FLIGHT DYNAMICS LABORATORY (WRDC/FDL)
(Wright-Patterson Air Force Base)

1. John Bay
2. Franklin Eastep
3. Marvin Hamstad
4. Chin Hsu
5. Ming-Shu Hsu
6. David Hui
7. Yulian Kin
8. Byung-Lip Lee
9. Vernon Matzen
10. William Wolfe
11. Lawrence Zavodney

FRANK J. SEILER RESEARCH LABORATORY (FJSRL)
(USAF Academy)

1. Theodore Burkey
2. Richard Carlin
3. Daniel Dolata
4. Ephrahim Garcia
5. Ngozi Kamalu
6. Siavash Kassemi
7. Bernard Piersma
8. Thomas Posbergh

C. PARTICIPANT LABORATORY ASSIGNMENT (Page 3)

GEOPHYSICS LABORATORY (AFGL)
(Hanscom Air Force Base)

1. Theodore Aufdemberge
2. Frank Battles
3. Reuben Benumof
4. Chia-Bo Chang
5. George Jumper
6. Jeffrey Kuhn
7. C. Randal Lishawa
8. Gandikota Rao
9. Craig Rasmussen
10. Glenn Stark
11. John Wills

HUMAN RESOURCES LABORATORY (HRL)
(Brooks, Williams, and Wright-Patterson Air Force Bases)

1. Margaret Batschelet
2. Pinyuen Chen
3. James Dykes
4. Daniel Garland
5. Harold Goldstein
6. Verlin Hinsz
7. Delayne Hudspeth
8. Gillray Kandel
9. William Moor
10. James Price
11. Joan Rentsch
12. Eleanor Smith
13. Stanley Stephenson
14. Pamela Tsang

MATERIALS LABORATORY (ML)
(Wright-Patterson Air Force Base)

1. Donald Chung
2. John Connolly
3. Sherif Elwakil
4. David Grossie
5. Prasad Kadaba
6. Joseph Lambert
7. Gary Leatherman
8. Won-Kyoo Lee
9. Michael Resch
10. Martin Schwartz
11. Hai-Lung Tsai

OCCUPATIONAL AND ENVIRONMENTAL HEALTH LABORATORY (OEHL)
(Brooks Air Force Base)

1. David Buckalew
2. Miguel Medina
3. Kirk Nordyke
4. Lorin Weber

ROME AIR DEVELOPMENT CENTER (RADC)
(Griffiss Air Force Base)

1. Charles Alajajian
2. Gary Craig
3. Lionel Friedman
4. Frances Harackiewicz
5. Hao Ling
6. Shietung Peng
7. Daniel Ryder
8. Behrooz Shirazi
9. Wayne Smith
10. Ronald VanEtten
11. Martin Wilner
12. James Wolper

C. PARTICIPANT LABORATORY ASSIGNMENT (Page 4)

SCHOOL OF AEROSPACE MEDICINE (SAM)
(Brooks Air Force Base)

1.	Phillip Bishop	10.	Paul Lemke
2.	Robert Blystone	11.	Rex Moyer
3.	Bruno Breitmeyer	12.	Arnold Nelson
4.	Vito DelVecchio	13.	Donald Robinson
5.	Randall Dupre	14.	David Senseman
6.	Reinhard Graetzer	15.	Richard Swope
7.	Paul Griffin	16.	John Szarek
8.	Pushpa Gupta	17.	Steven Waller
9.	Ramesh Gupta		

WEAPONS LABORATORY (WL)
(Kirtland Air Force Base)

1.	William Campbell	4.	Johanna Schruben
2.	Ggene Carlisle	5.	Peter Walsh
3.	William Cofer	6.	Wayne Zimmerman

WILFORD HALL MEDICAL CENTER (WHMC)
(Lackland Air Force Base)

1. Janet Dizinno

RESEARCH REPORTS

RESEARCH REPORTS

1990 SUMMER FACULTY RESEARCH PROGRAM

Technical Report Number	Title	Professor
Volume I		
Armament Laboratory		
1	Simple Models for Predicting Runway Failure Due to Blast Loading	Dr. Charles Camp
2	Physical Aspects of the Penetration of Reinforced Concrete Slabs	Dr. Arnold Carden
3	Solid-State Imager Replacement for a High-Speed Film Camera	Dr. Eustace Dereniak
4	Evaluation of Weapon Target Allocation Algorithms	Dr. Charles Fosha
5	Methods Which Accelerate Convergence in Iterative CFD Solvers	Dr. John George
6	Designing a Binary Phase Only Filter Via the Genetic Algorithm	Mr. Frederick Gibson
7	Two-Dimensional Simulation of Railgun Plasma Armatures	Dr. Manuel Huerta
8	Neural Networks for Guidance, Navigation, and Control of Exoatmospheric Interceptors *** Not Publishable at this Time ***	Dr. Kevin Moore
9	Gunner Performance in the BSTING Fire Control System	Dr. William Siuru
10	Robust Eigenstructure Assignment with Application to Missile Control	Dr. Kenneth Sobel
11	Multiple Point Detonation Modeling	Dr. Steven Trogdon

Volume I
Arnold Engineering Development Center

Astronautics Laboratory

Volume I

Volume II
Frank J. Seiler Research Laboratory

Geophysics Laboratory

Volume II

Rome Air Development Center

Volume II

Weapons Laboratory

Volume III
(Wright Research Development Center)
Aero Propulsion Laboratory

Avionics Laboratory

Volume III

Volume III

Volume III

Volume IV
Human Systems Division Laboratories
Harry G. Armstrong Aerospace Medical Research Laboratory

Volume IV

Volume IV

School of Aerospace Medicine

Volume IV

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by The

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc

FINAL REPORT

Evaluation of MOS-Controlled Thyristor (MCT) at 270 Volt DC

for Resistive and Inductive Loads

Prepared by: Muhammad A. Choudhry, Ph.D.

Academic Rank: Professor

Department and   Department of Electrical and Computer Engineering

University: West Virginia University

Research Location: Electrical Technology/Aerospace Power Division

                       Wright-Patterson AFB

                       Ohio 45433-6563

USAF Researcher: Joseph A. Weimer

Date: August 15, 1990

Contract No: F49620-88-C-0053

Evaluation of MOS-Controlled Thyristor (MCT) at 270 Volt DC

for Resistive and Inductive Loads

by

Muhammad A. Choudhry

ABSTRACT

The voltage and current characteristics of MOS-Controlled Thyristors of 50 to 150 A rating are obtained at 270 volt dc over a wide range of switching frequencies. Total turn-off time of MCT increases with load current and switching frequency and is less than 2 u sec in most cases. The forward voltage drop across MCT varies from 1 volt to 2 volts for a load current variation of 15 to 90 A. Large voltage transients are observed during switching of inductive loads. The use of snubber circuit across MCT reduces voltage transients and power loss inside the device. However, large currents are observed during turn-on of MCT at high frequency with inductive load.

Acknowledgments

I. INTRODUCTION:

Unlike Gate-Turn- Off Thyristor which can be turned on by applying a small positive pulse to the gate but requires about one fifth of normal current to turn off, MOS-Controlled Thyristor can be turned on and off with a very small amount of current pulse applied between gate and anode.   The Electrical Technology Section of Aerospace Power Division at Wright-Patterson AFB is interested in application of MCT in "More-Electric Aircraft" due to its light-weight, low forward voltage drop and large current carrying capability.

My research interests have been in the areas of control and analysis of ac/dc power system, effect of nonlinear load on ac/dc power system stability and use of power electronics to enhance safety and increase productivity in underground coal mines. My work on the use of GTO to suppress arc in mine dc haulage systems and analysis of ac/dc power system contributed to my assignment to Electrical Technology/Aerospace Power Division.

II. OBJECTIVES OF THE RESEARCH EFFORT:

Power electronic devices can be used for control and efficient utilization of electrical energy. However, these devices may be damaged by the electrical transients due to switching of the device or the load. MOS-Controlled Thyristor (MCT) can be turned on and off with a very small amount of current pulse applied between anode and gate. This device has low forward voltage drop, high current capability and low weight/power ratio.

My assignment as a participant in the 1990 Summer Faculty Research Program (SFRP) was to evaluate voltage and current characteristics of MCT at different switching frequencies with different types of loads at 270 V DC. This information will be utilized in the "More-Electric Aircraft" initiative which will improve thrust to weight ratio, reliability and safety of the aircraft.

III.

a. MOS-Controlled Thyristor has high current density, high dI/dt capability and very small turn-on and turn-off time [1]. This device is expected to displace SCRs and GTO in the future. Future aircrafts are expected to use more electric power at 270 V DC to improve reliability and safety. Figure 1 shows the single line diagram of the circuit used for evaluating the voltage/current characteristics of MCT at 270 VDC. Three-phase ac output of auto-transformer is rectified and a capacitor bank is connected at the rectifier to reduce the ripple in the dc voltage. A large resistor is connected parallel to the capacitor bank to discharge the capacitors when
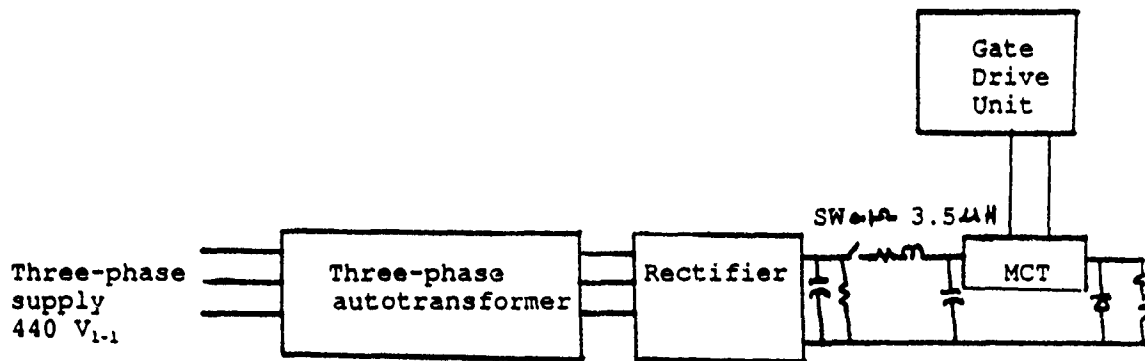
Figure 1. Single-line diagram of experiment to evaluate voltage/current characteristics of MCT at 270 v DC.

switch SW is opened. A filtering capacitor is connected at the MCT to reduce the voltage variation of the anode voltage during turn-on and turn-off of MCT. A free-wheeling diod is connected across the load to reduce switching transient voltage across MCT. Resistive load was modeled by glow bar and water-cooled resistors.

b. Figure 2 shows the variation of on-state forward voltage drop in MCT with anode current. The on-state forward voltage drop decreases with higher junction temperature of MCT. Figure 3 shows the voltage waveform across MCT at a switching frequency of 10 K Hz for a resitive load at 270 VDC. Figure 4 shows the expanded voltage waveform during turn-on and turn-off of MCT. The voltage across MCT decreases from 90% to 10% of dc input voltage in 222.3 u sec.

Figure 5 shows the anode current and gate current waveform at a switching frequency of 10 K Hz for a resistive load at 270 VDC. A gate current of less than 2A is required to turn-on and turn-off 38.2A through the anode of MCT. Figure 6 shows the expanded anode current waveform during turn-on and turn-off of MCT. The current increases exponentially during turn-on due to a small inductance present in the load circuit. The total turn-off time (decreasing of anode current to 10% of its value before turn-off after the gate voltage rises through zero) is 1.92 u sec. Table 1 gives the total turn-off time of MCT as a function of anode current and switching frequency. The turn-off time increases with load current and switching frequency.
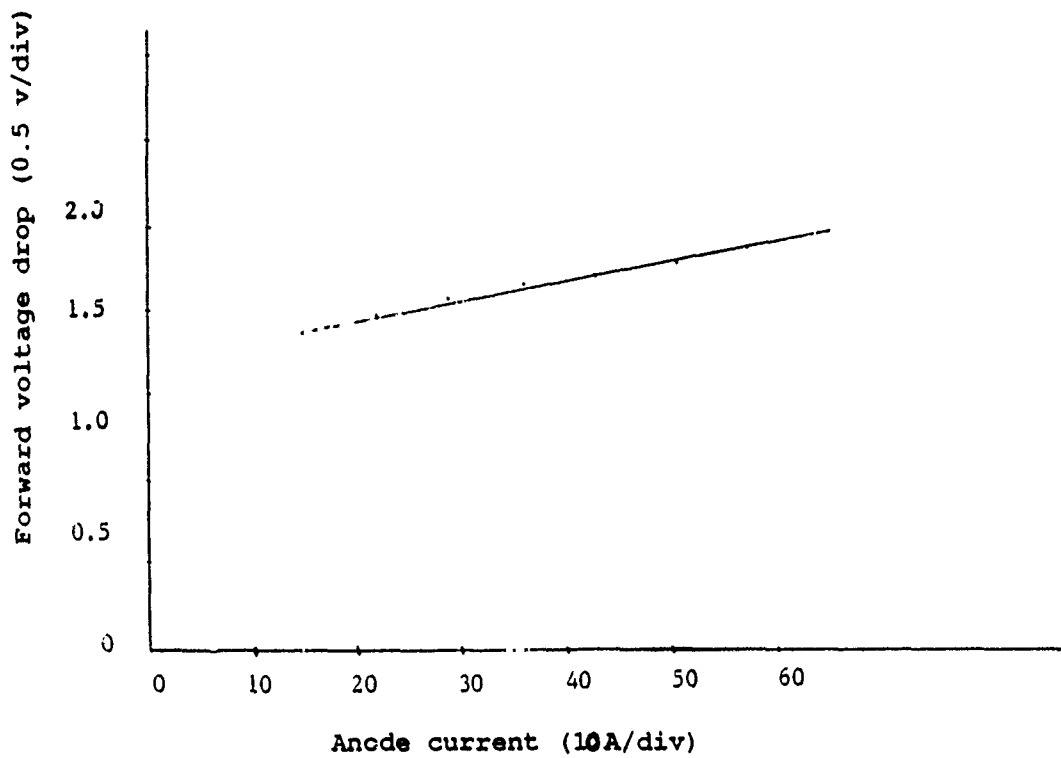
Figure 2. Variation of forward voltage drop with anode current.
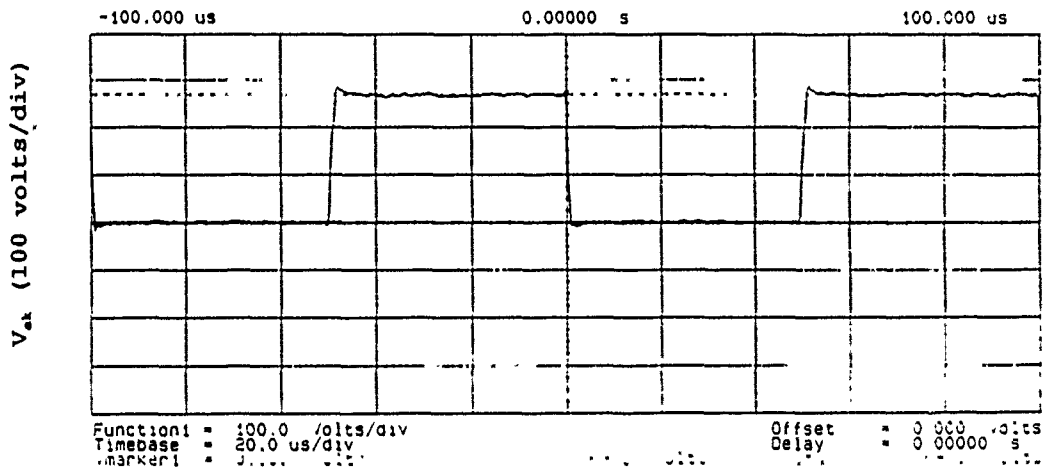
Figure 3. Voltage waveform across MCT at a switching frequency of 10 KHz for a resistive load at 270 v DC.
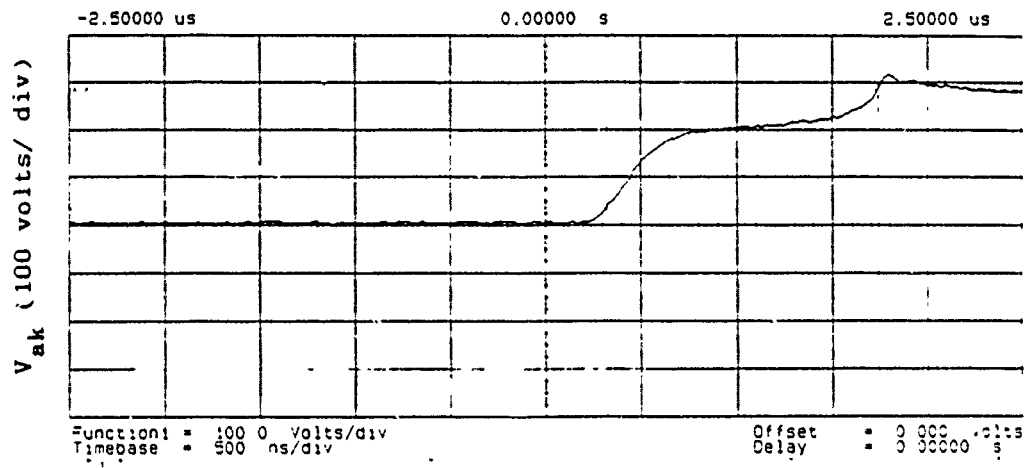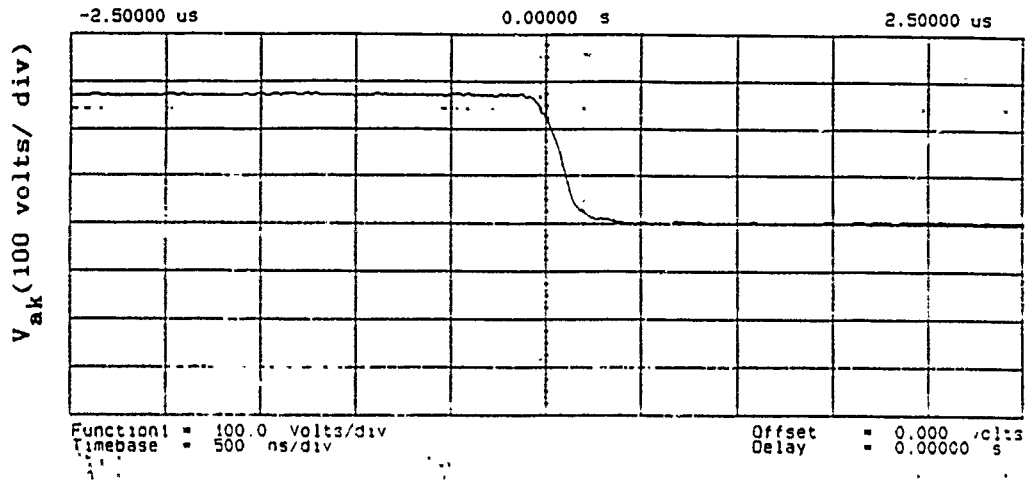
Figure 4. Voltage waveform during turn-on and turn-off of MCT for a resistive load at 270 v DC.
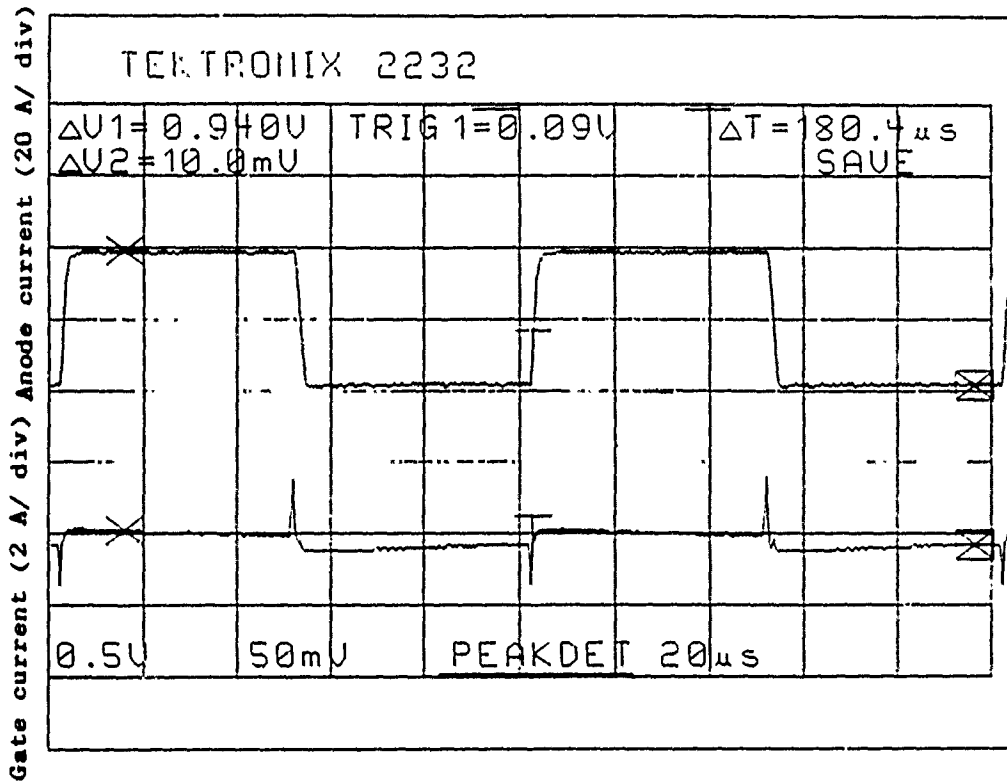
Figure 5. Anode current and gate current waveforms at a switching
frequency of 10 KHz for a resistive load at 270 v DC.

Anode current (20 A/ div)

ΔV1= 0.960V   TRIG 1=0.09V   ΔT=4.510μs
ΔV2=10.0mV   SAVE

0.5V   50mV   SAMPLE   0.5μs

Gate current (2 A/ div)

ΔV1= 1.080V   TRIG 2= 19mV   ΔT=4.510μs
ΔV2=12.0mV   SAVE

0.5V   50mV   '-_E'   0.5μs
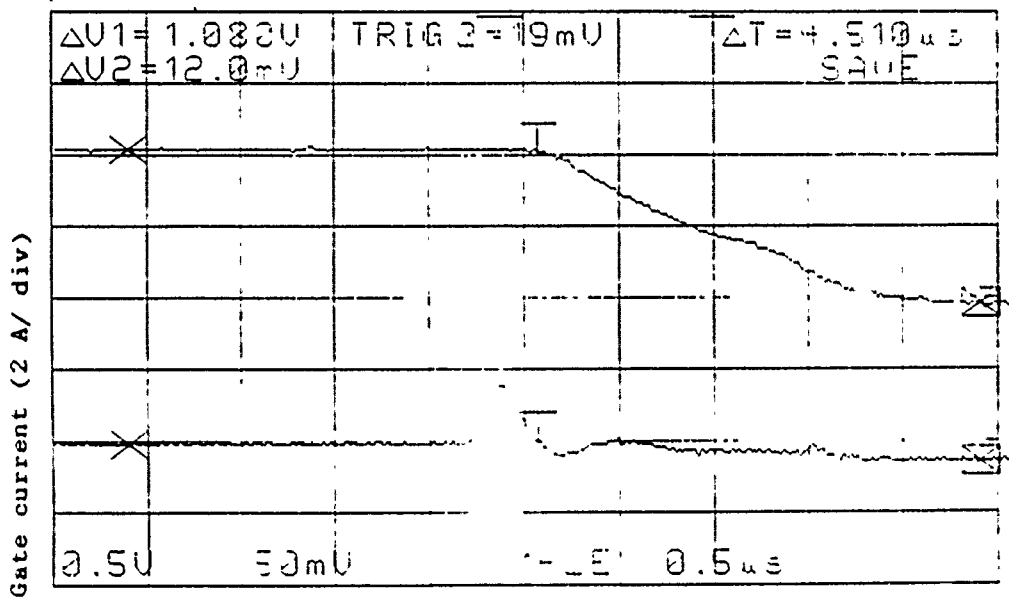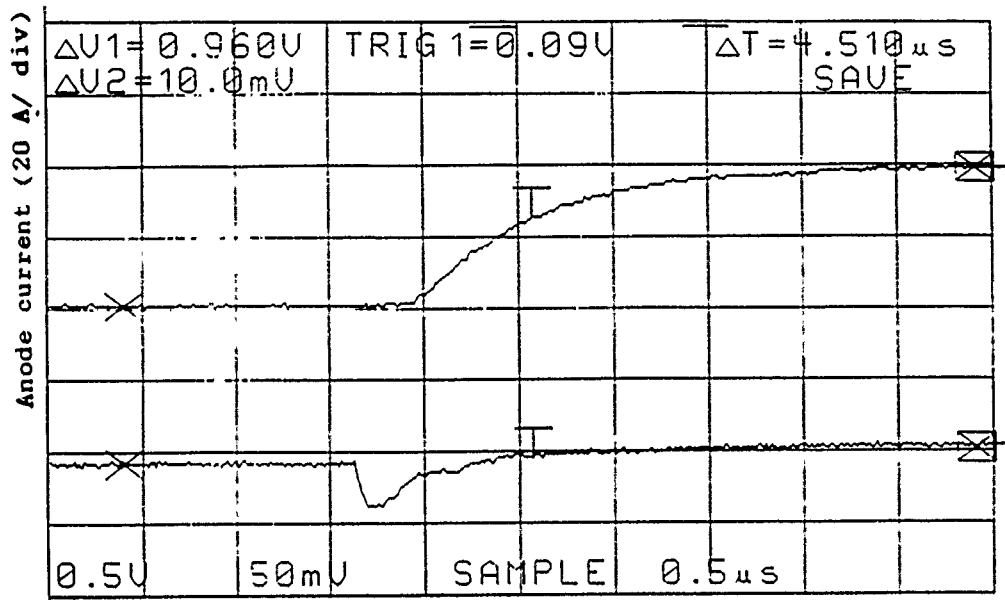
Figure 6.   Anode current waveforms during turn-on and turn-off at
            switching frequency of 10 KHz for a resistive load at
            270 v DC.

Table 1. Total turn-off time of MCT as a function of anode current and switching frequency.

Switching Frequency= 400 Hz		Switching Frequency= 10KHz	
Anode Current	Total turn-off time	Anode Current	Total turn-off time
(A)	(sec)	(A)	(sec)
10.97	1.32	10.74	1.36
20.34	1.39	17.83	1.51
36.47	1.44	42.86	1.98

Figure 7 shows the voltage, gate and anode current waveforms of MCT for an inductive load at a switching frequency of 10 K Hz at 270 VDC. A large transient current flows through MCT during turn-on which causes voltage oscillations across MCT. After the initial current, the current increases exponentially at a time constant determined by the inductance and resistance of the load circuit.

Figure 8 shows the voltage waveform across MCT without the filtering capacitor at the MCT at a switching frequency of 400 Hz. The input voltage is 100 VDC. A transient voltage of 263.6 V appears across MCT. Figure 9 shows the voltage waveform across MCT with a snubber circuit at a switching frequency of 400 Hz. The input voltage is 270 VDC. A transient voltage of 381.8 V appears across the device. The use of snubber circuit reduces the voltage transient across MCT and reduces the switching power loss inside MCT.

IV. RECOMMENDATIONS:

The MOS-Controlled Thyristor is able to turn-off large currents in less than 2 u sec at 270 VDC. However, large voltage oscillations appears during turn-on with inductive load at higher switching frequencies. This research shows that the nature of the load has significant effect on the performance of MCT. Further research will be done to evaluate voltage/current characteristics of MCT in series with a dynamic load. Computer simulation will be developed to optimize the design of snubber circuit of MCT and to determine the cause of high surge current during turn-on at high switching frequency.
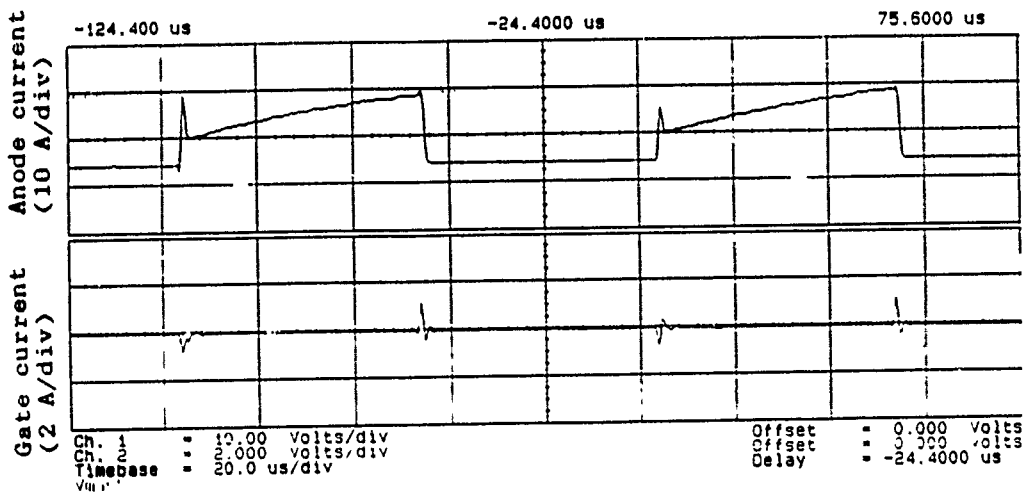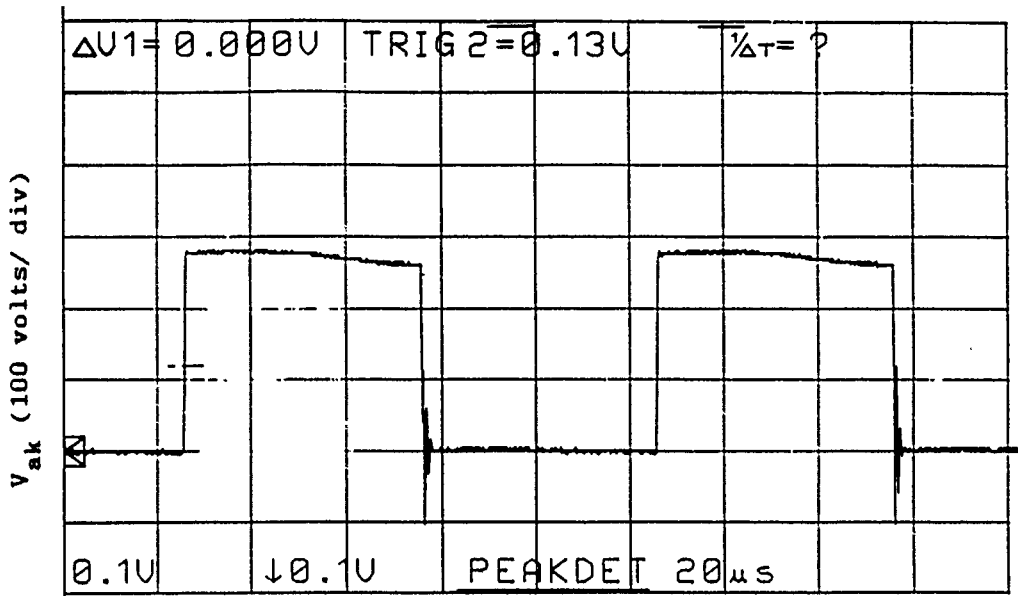
Firgure 7. Voltage and current waveforms for an inductive load at
a switching frequency of 10 KHz at 270 v DC.
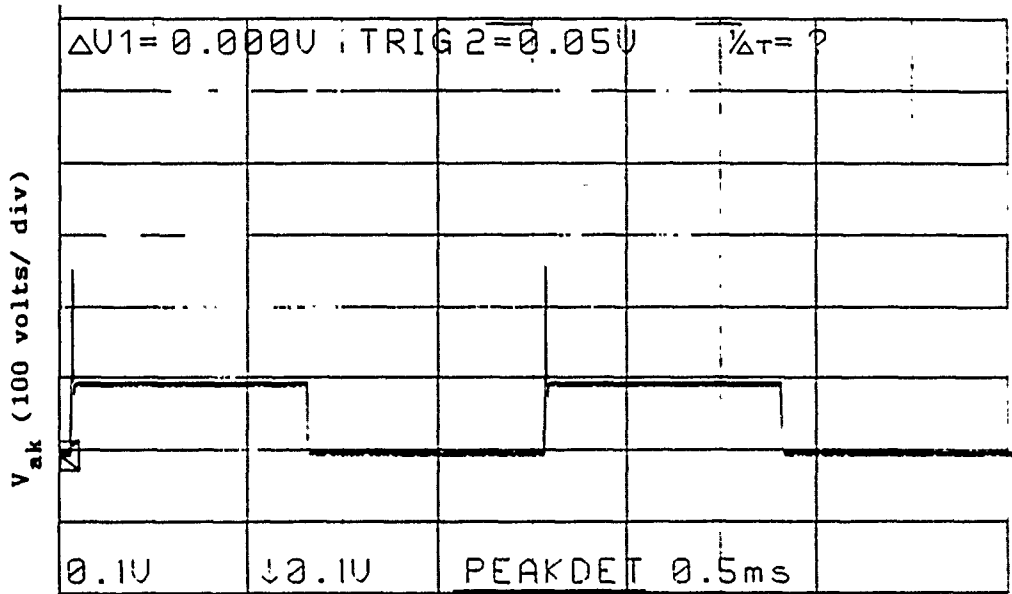
Figure 8. Voltage across MCT without a filtering capacitor at MCT
for an inductive load at 100 v DC.
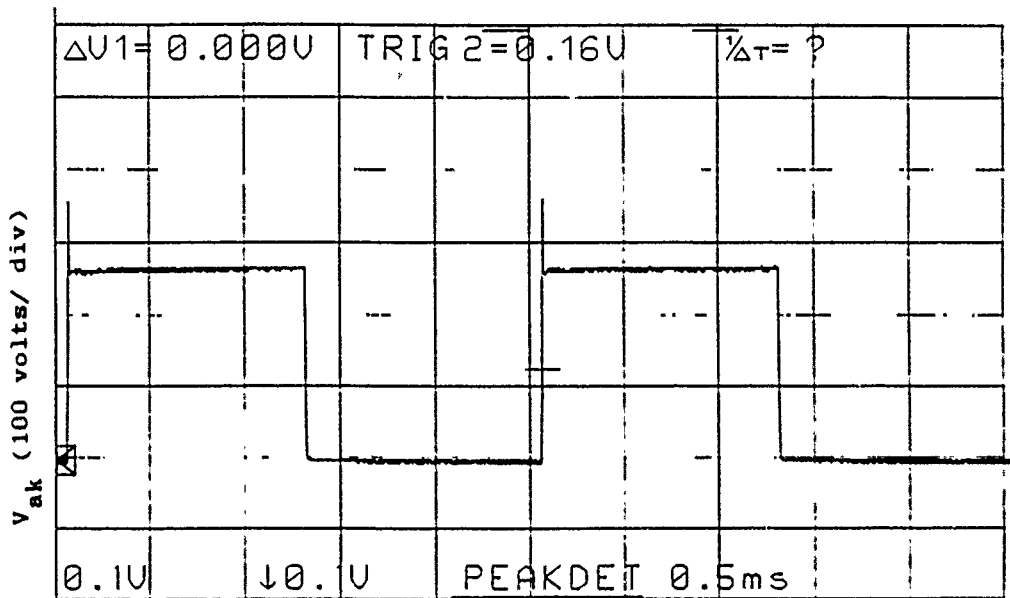
Figure 9.   Voltage across MCT with a snubber circuit for an
           inductive load at 270 v DC.

REFERENCE

1. V.A.K. Temple. MCTs - Thyristors for the Future. Power technics Magazine, November 1989, pp. 21-24.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/GRADUATE
STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OF SCIENTIFIC RESEARCH

Conducted by the

UNIVERSAL ENERGY SYSTEMS, INC.

FINAL REPORT

DEVELOPMENT OF A THREE-DIMENSIONAL FINITE-DIFFERENCE CODE FOR MODELING FLOW AND HEAT TRANSFER IN ROTATING DISK SYSTEMS

Prepared by: Mingking K. Chyu

Academic Rank: Associate Professor

                       Department of Mechanical Engineering
                       Carnegie Mellon University
                       Pittsburgh, PA 15213

Research Location: Wright Research and Development Center,
                       POOS-3, Dayton Ohio

USAF Researcher: Dr. Won S. Chang

Date: September 30, 1990

Contract Number: F49620-88-C-0053

DEVELOPMENT OF A THREE-DIMENSIONAL FINITE-DIFFERENCE CODE
FOR MODELING FLOW AND HEAT TRANSFER IN ROTATING DISK SYSTEMS

M. K. Chyu
Dept. of Mechanical Engineering
Carnegie Mellon University
Pittsburgh, PA 15213

W. S. Chang
AFWAL/POOS-3
Wright Research Development Center
WPAFB, OH 45433

ABSTRACT

A pressure-based, three-dimensional, finite-difference (finite-volume) code for modeling the fluid flow and heat transfer in a turbine disk system has been developed. The computer program developed, however, is capable of solving the governing equations for convection transport in elliptic nature. The computational procedure employs a colocated grid system extended for generalized non-orthogonal coordinates thus can handle irregular geometries. The velocity-pressure coupling uses a modified version of the pressure implicit split operator (PISO) originally proposed by Issa in 1985. The PISO algorithm not only gives time-accuracy results but also provides the computation with a much faster converging speed as compared to the SIMPLE based counterparts. Ongoing research using this computational program is to investigate the effects of radial and axial clearance on the heat transfer in a turbine rotor-stator system. Near-future study plan will focus on counter-rotating disk systems which appear to be the viable configurations used for future turbine engines.

ACKNOWLEDGEMENTS

NOMENCLATURE

g_{ij}	geometric coefficients
J	Jacobian
p	pressure
Pr	Prandtl number
S	source term
T	temperature
u_i	velocity at duct inlet or velocity components (u,v,w)]
	in cartesian coordinate (x,y,z)
U_i	contravariant variables (U,V,W)
x,y,z	cartesian coordinates
X,Y,Z	(x,y,z) normalized by D, see Fig. 2
μ	viscosity
ρ	density
ξ,η,ζ	general non-orthogonal coordinate system

Subscript

P,E,W,WW	variable values at node P, E...at the discretized domain

Superscript

n	at current time step
n+1	at next time step
*,**	at intermediate time step

I. INTRODUCTION:

The study of the heat transfer characteristics of rotating disk systems is a subject of importance in the design of various types of rotating equipment, particularly for the aeropropulsion and turbomachinery. In the continual quest for improved performance and efficiency, modern gas turbine designers strive to develop systems and/or materials that will permit increasing turbine operating temperatures. Consequently, much effort has been expended to enhance understanding of the hydrodynamic and heat transfer phenomena of various types of rotating disk systems.

Heat transfer from a disk rotating in a quiescent environment has been examined in numerous experimental and theoretical studies. Results of such investigations (Millsaps and Pohlhausen (1952), Cobb and Saunders (1956), Sparrow and Gregg (1960), and Ong (1988)) have provided extensive insight into the hydrodynamic and thermal characteristics of free disks. While these results may be of limited applicability in turbine design, they often serve as a baseline for gauging the effectiveness of various forced cooling techniques and for judging the heat transfer characteristics of rotor-stator systems, which are more appropriate idealized turbine disk models.

Unlike the free disk in which heat transfer is only affected by rotational speed, heat transfer in rotor-stator systems is influenced by the operating conditions as well as system geometry. In a rotor-stator system with no radial shroud, commonly referred to as an open rotor-stator, Kreith et. al. (1959, 1963) and Metzger (1970) demonstrated that the presence of the stator influences heat transfer from the disk at axial gap spacings less than 10 percent of the disk diameter. These studies also examined the effects that forced coolant flow introduced at the hub of the system had on the average heat transfer rate. As an extension of these studies, more recent evaluations by Metzger et al. (1979, 1989) and Popiel and Boguslawski (1986a& 1986b) have examined the effects of varying the location and rate of forced coolant flow. In these studies the effects coolant jet impingement have on rotor local heat transfer coefficients were investigated.

Another important geometrical effect in actual turbines is the presence of a radial clearance seal. While the primary intent of forced disk cooling systems is to enhance heat transfer from the rotating turbine disk, this flow is also intended to prevent the ingress of hot combustion gases. Phadke and Owen (1983) investigated the hydrodynamic effects of different seal designs to assess their ability to prevent ingress, but few studies have

explicitly examined the effect that a radial clearance shroud has on heat transfer rates. Metzger et. al. (1989) performed local heat transfer measurements across a disk face using a radially shrouded rotor-stator system, but the purpose of these studies was to investigate the influence that the radial location and rate of coolant flow have on the local heat transfer rates.

II. OBJECTIVES OF THE RESEARCH EFFORT:

The primary objective for this summer research is to establish a numerical code for modeling convection transport phenomena in turbine disk systems. The computational procedure developed should, at least, have the capability of addressing the important issues and research concerns involved in the rotating disk problems. Moreover, it is targeted to become one of the most powerful and modern computer programs for general elliptic problems in the area of computational fluid dynamics (CFD) and heat transfer. Many special features and algorithms of recent advent in the CFD community are also included in the procedure. With an extensive review and numerical testing, the final ramification is a pressure-based, finite-difference code capable of solving the incompressible, three-dimensional, full Navier-Stokes equations in non-orthogonal coordinates. One special feature of this finite-difference (finite-volume) procedure is the capability of dealing with irregular geometries, which is used to be monopolized by the finite-element methods. In addition, it adopts a non-staggered grid arrangement, thus the velocity components and other dependent variables are colocated at the same grid. The velocity-pressure coupling is resolved by the pressure implicit split operator (PISO) (Issa, 1985) with an extension for generalized non-orthogonal coordinates. Discretization of the governing equations uses the central-differencing scheme for the diffusion terms and the second-order upwind scheme for the convection terms. The solution of the discretized equations is obtained from a conjugate gradient method (Kershaw, 1978), instead of the conventional tri-diagonal matrix algorithm (TDMA) incorporated with a line-by-line iteration. The remaining part of this report describes the framework of the code.

III. MATHEMATICAL FORMUALTION:

For laminar flow of Newtonian fluid with constant properties, the governing equations for the unsteady flow and heat transfer in cartesian tensor form are

Continuity

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{1}$$

Momentum

$$\frac{\partial \rho u_i}{\partial t} + \frac{\partial}{\partial x_j}(\rho u_i u_j) = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j}\left[\mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\right] \tag{2}$$

Energy

$$\frac{\partial \rho T}{\partial t} + \frac{\partial}{\partial x_j}(\rho u_j T) = \frac{\partial}{\partial x_j}\left(\frac{\mu}{pr}\frac{\partial T}{\partial x_j}\right) \tag{3}$$

To treat arbitrary boundaries, these governing equations are transformed to a general form based on a non-orthogonal coordinate system (ξ, η, ζ). There are three different ways to represent velocity field in a non-orthogonal grid system; i.e., (1) cartesian velocity components (Rhie and Chou, 1983; Han, 1988), (2) contravariant velocity components (Maliska and Raithby, 1984) and (3) resolute or velocity projections along the coordinate directions (Karki and Patankar, 1989). The present transformation uses the cartesian velocity components, which has the advantage of excluding the so-called "curvature" terms (Rhie and Chou, 1983). Using the standard transformation formulae (Anderson et al., 1984), the continuity equation becomes

$$\frac{\partial U}{\partial \xi} + \frac{\partial V}{\partial \eta} + \frac{\partial W}{\partial \zeta} = 0 \tag{4}$$

and the remaining governing equations for the dependent variable ϕ in the general coordinate system can be expressed in a compact form as follows:

$$\frac{\partial}{\partial t}(\rho\phi) + \frac{\partial}{\partial \xi}(\rho U\phi) + \frac{\partial}{\partial \eta}(\rho V\phi) + \frac{\partial}{\partial \zeta}(\rho W\phi)$$

$$= \frac{\partial}{\partial \xi}\left[\frac{\Gamma_\phi}{J}(g_{11}^2 + g_{12}^2 + g_{13}^2)\frac{\partial\phi}{\partial\xi}\right]$$

$$+ \frac{\partial}{\partial \eta}\left[\frac{\Gamma_\phi}{J}(g_{21}^2 + g_{22}^2 + g_{23}^2)\frac{\partial\phi}{\partial\eta}\right] \tag{5}$$

$$+ \frac{\partial}{\partial \zeta}\left[\frac{\Gamma_\phi}{J}\left(g_{31}^2 + g_{32}^2 + g_{33}^2\right)\frac{\partial \phi}{\partial \zeta}\right] + JS^\phi$$

where U, V, and W are the contravariant variables that represent convective flux ;

$$\left.\begin{array}{l} U = g_{11}u + g_{12}v + g_{13}w \\ V = g_{21}u + g_{22}v + g_{23}w \\ W = g_{31}u + g_{32}v + g_{33}w \end{array}\right\} \text{ or } U_i = g_{ij}u_j \qquad (6)$$

and Γ_ϕ and S^ϕ are the associated diffusivity and source terms for the variable ϕ (=u,v,w,T). The detailed expressions of the source terms and the cross derivative terms due to grid non-orthogonality have been given by Han (1988). The geometric coefficient g_{ij} and the Jacobian J are defined as

$$g_{ij} = \begin{bmatrix} y_\eta z_\zeta - y_\zeta z_\eta & x_\zeta z_\eta - x_\eta z_\zeta & x_\eta y_\zeta - x_\zeta y_\eta \\ y_\zeta z_\xi - y_\xi z_\zeta & x_\xi z_\zeta - x_\zeta z_\xi & x_\zeta y_\xi - x_\xi y_\zeta \\ y_\xi z_\eta - y_\eta z_\xi & x_\eta z_\xi - x_\xi z_\eta & x_\xi y_\eta - x_\eta y_\xi \end{bmatrix} \qquad \text{and}$$

$$J = x_\xi y_\eta z_\zeta + x_\zeta y_\xi z_\eta + x_\eta y_\zeta z_\xi - x_\xi y_\zeta z_\eta - x_\eta y_\xi z_\zeta - x_\zeta y_\eta z_\xi \qquad (7)$$

These geometric coefficients are evaluated using central differencing schemes in the transformed domain.

IV. DISCRETIZATION:

Discretization of the governing, differential equations uses the finite-volume approach. Differencing in the temporal domain employs the Implicit Euler scheme. All the dependent and independent variables are stored at the same grid location and variables at the finite control-volume boundaries are interpolated between adjacent grid points. In the spatial domain, the diffusion terms and convective terms are approximated by the central differencing and a second-order upwind scheme, respectively. The present upwind scheme determines the prevailing value of a dependent variable at the finite-volume interfaces by referring to the sign of the contravariant fluxes. For example, if the contravariant flux U at point p is positive, the finite-volume face value ϕ_W (west to P) is calculated by linear

extrapolation from the corresponding variable at the two closest upstream neighboring points (ww means further _west_ to the _west_). This leads the discretization of the term

$$\frac{\partial(U\phi)}{\partial\xi} = A_p\phi_p - A_w\phi_w - A_{ww}\phi_{ww}$$

where

$$A_w = \frac{\Delta x_w (2\Delta x_{ww} - \Delta x_w) U_w + \Delta x_E (\Delta x_{ww} - \Delta x_w) U_e}{2\Delta x_w (\Delta x_{ww} - \Delta x_w)}$$

$$A_{ww} = \frac{-\Delta x_w \cdot U_w}{2(\Delta x_{ww} - \Delta x_w)}$$

and

$$A_p = A_w + A_{ww}$$

V. VELOCITY-PRESSURE COUPLING:

A typical computational procedure for compressible flow employs density as a primary variable and extracts pressure from an equation of state. This density-based method is difficult for incompressible or low-Mach number flows, since weak pressure-density coupling prevails under these situations. Computation for incompressible flow is generally performed using pressure-based method which is characterized by the use of pressure as one of the primary dependent variables. However, the absence of both unsteady term in the continuity equation and an explicit equation specifically for pressure imposes that the pressure can only influence the velocity field through the momentum conservation with the continuity as a compatibility condition. Hence provision of an algorithm for the velocity-pressure coupling is essential. One well-known algorithm employs semi-implicit iterative coupling procedure; e.g., SIMPLE (Patankar, 1981) and its variants (Van Doormaal and Raithby, 1985). In the present study, a more efficient non-iterative method based on the operator-splitting technique (PISO) (Issa, 1985) is extended to non-orthogonal grid coordinates. The discretized governing equations are solved in a time-marching fashion until a steady state solution is reached. In each time step, a two-stage correction procedure follows an initial predictor stage, and all three stages combined yields a time-accuracy result. The sequence of the scheme is as follows:

a) Predictor Step:

Using the pressure field at time step t^n, the discretized equation for u_i is

$$\frac{1}{\Delta t}\left\{(\rho u_i)^* - (\rho u_i)^n\right\} = H(u_i^*) - g_{ji}\nabla_j p^n + S_i \tag{8}$$

where $H(\ )$ represents the finite difference discretization operator. With $H(u_i^*) = H'(u_i^*) + A_p u_i^*$, the first step becomes

$$\left(\frac{\rho}{\Delta t} - A_p\right)u_i^* = H'(u_i^*) - g_{ji}\Delta_j p^n + S_i + \frac{\rho}{\Delta t}u_i^n \tag{9}$$

The coefficients A_p's involve contributions due to the convective fluxes as well as diffusive fluxes in the ξ, η, ζ directions.

b) First Corrector Step

The mass fluxes calculated from equation (9) unnecessarily satisfy continuity, the first corrector step is established by splitting the operator $H(u_i^{**}) = A_p u_i^{**} + H'(u_i^*)$, thus

$$\left(\frac{\rho}{\Delta t} - A_p\right)u_i^{**} = H'(u_i^*) - g_{ji}\nabla_j p^* + S_i + \frac{\rho}{\Delta t}u_i^n \tag{10}$$

Subtracting equation (9) from (10) and imposing the continuity equation ($\nabla_j U_j^{**} = 0$), the pressure correction equation is resulted by taking divergence of the incremental momentum correction equation:

$$\nabla_j\left[\left(\frac{\rho}{\Delta t} - A_p\right)^{-1} g_{ji}^2 \nabla_j (p^* - p^n)\right] = \nabla_j g_{ji} u_i^* = \nabla_j U_j^* \tag{11}$$

Note that the first corrector step is similar to the SIMPLE algorithm. In fact, if the loops are iteratively terminated here, it is identical to the SIMPLE method with the time-step-involving term as an under-relaxation factor. To account for the effects of convection and diffusion and to satisfy the momentum conservation, a second corrector step follows.

c) Second Corrector Step

Further splitting the operator; i.e.,

$$H\left(u_i^{***}\right) = A_p u_i^{***} + H'\left(u_i^{**}\right),$$

the momentum equation becomes

$$\left(\frac{\rho}{\Delta t} - A_p\right) u_i^{***} = H'\left(u_i^{**}\right) - g_{ji} \nabla_j p^{**} - S_i + \frac{\rho u_i^n}{\Delta t} \qquad (12)$$

Subtracting equation (10) from (12) gives the incremental momentum equation. Then, by invoking the continuity again, the pressure increment equation becomes:

$$\nabla_j \left[\left(\frac{\rho}{\Delta t} - A_p\right)^{-1} g_{ji}^2 \nabla_j \left(p^{**} - p^*\right) \right]$$

$$= \nabla_j \left[\left(\frac{\rho}{\Delta t} - A_p\right)^{-1} g_{ji} H'\left(u_i^{**} - u_i^*\right) \right] - \nabla_j \left(g_{ji} u_i^{**}\right) \qquad (13)$$

The splitting errors after two corrector stages are less than the truncation error associated with the Euler implicit approximation in the temporal domain. The prevailing values obtained here are taken to be the values at the next time-step, t^{n+1}.

It is known that the non-staggered grid system can produce the undesirable checkerboard phenomena in the pressure field, due to the decoupling effect between velocities and pressures. As suggested by Rhie and Chou (1983), an explicit fourth order pressure damping terms may be added to the right hand side of equation (11) to ensure an attainment of oscillatory-free pressure field. This approach can be further modified to include the transient term for time-marching formulations. However, the steady-state solution obtained in this way is found to be strongly influenced by the size of time step chosen, which is undesirable (Parameswaran and Sun, 1989). In this study, the inclusion of the transient term in the formulation is done by the operator interpolation technique first proposed by Peric (1988). For example, the U_j's evaluated at the cell interface $i+1/2$ becomes:

$$U_j^*\Big|_{i+\frac{1}{2}} = \frac{g_{ji}}{2} \left[\left(\frac{\rho}{\Delta t} - A_p\right)^{-1} \left(H'\left(u_{i+1}\right) + S_{i+1} + \frac{\rho}{\Delta t} u_{i+1}^n\right) \right.$$

$$+ \left(\frac{\rho}{\Delta t} - A_p\right)^{-1} \left[H'(u_i) + S_i + \frac{\rho}{\Delta t} u_i^n\right] + g_{ji}(p_{i+1}^n - p_i^n) \qquad (14)$$

The resulting mass fluxes are then substituted into the right hand side of pressure correction equation (11) and (13).

The system of finite difference approximation equations formed by equations (8), (11) and (13) produces a seven-banded matrix. For both pressure correction equations, the matrix is symmetric and positive-definite. To enhance the matrix solver efficiency, the conjugate gradient (CG) method of Kershaw (1978) is employed for reducing the residuals of these two equations to 10^{-3}.

In actual computation, it generally solves the non-dimensionalized governing equations. Dependent variables are scaled to the characteristic lengths, velocities and corresponding references. A preliminary result for a study concerning the radial and axial clearance effects on the heat transfer in a rotor-stator system shows that the present code performs far superior than the SIMPLE related counterparts, by a factor of 10 - 100 in converging speed. Most of the pressure-based codes in the public domain or commercially available are based on SIMPLE algorithm or its derivatives.

VI. RECOMMENDATION:

Besides the ongoing research for the rotor-stator system, this code is capable of performing similar modeling for other disk configurations. Because its feature of formulating governing equation in generalized non-orthogonal coordinates, the effects of geometric details; e.g. uneven thickness and surface curvature, can be studied with a grid generation. These issues have never been explored in the past. Another important configuration is the system involving counter-rotating disks, that is possibly to be used in the future high-performance turbine engines. Many existing theories and findings for the conventional rotor-stator system are inapplicable for the counter-rotating disks; both numerical modeling and experimental work for code validation are strongly desirable.

REFERENCES

Anderson, D.A., Tannehill, J.C. and Pletcher, R.H., *Computational Fluid Mechanics and Heat Transfer*, Hemisphere Publication, Washington D.C., 1984.

Cobb, E.C. and Saunders, O.A., "Heat Transfer from a Rotating Disk," *Proceedings, Royal Society*, Volume 236, 1956, pp. 343-351.

Dibelius, G.H. and Heinen, M., "Heat Transfer from a Rotating Disc," Proceedings of the Gas Turbine and Aeroengine Congress and Exposition, Brussels, 1990.

Issa, R.I., "Solution of the Implicit Discretized Fluid Flow Equations by Operator Splitting," *J. Computational Physics*, Vol. 62, p. 40, 1985.

Han, T., "A Navier-Stokes Analysis of Three-Dimensional Turbulent Flows Around a Bluff Body in Ground Proximity," AIAA Paper 88-3766, 1988.

Karki, K.C. and Patankar, S.V., "Pressure Based Calculation Procedure for Viscous Flows at All Speeds in Arbitrary Configurations," *AIAA J.*, Vol. 27, No. 9, 1989, pp. 1167-1174.

Kershaw, D.S., "The Incomplete Cholesky - Conjugate Gradient Method for the Iterative Solution of Systems of Linear Equations," *J. Computational Physics*, Vol. 26, p. 43, 1978.

Kreith, F., Taylor, J.H. and Chong, J.P., "Heat and Mass Transfer from a Rotating Disk," *Journal of Heat Transfer*, Volume 81, 1959, pp. 95-105.

Kreith, F., Doughman, E. and Kozlowski, H., "Mass and Heat Transfer from an Enclosed Rotating Disk with and without Source Flow," *Journal of Heat Transfer*, Volume 85, 1963, pp. 163-163.

Maliska, C.R. and Raithby, G.D., "A Method for Computing Three-Dimensional Flows Using Non-Orthogonal Boundary-Fitted Coordinates," *Int. J. of Num. Meth. in Fluids*, Vol. 4, p. 518, 1984.

Metzger, D.E., "Heat Transfer and Pumping on a Rotating Disk with Freely Induced and Forced Cooling," *Journal of Engineering for Power*, Volume 92, 1970, pp. 342-348.

Metzger, D.E., Mathis, W.J. and Grochowsky, L.D., "Jet Cooling at the Rim of a Rotating Disk," *Journal of Engineering for Power*, Volume 101, 1979, pp. 68-72.

Metzger, D.E., Bunker, R.S. and Bosch, G., "Transient Liquid Crystal Measurement of Local Heat Transfer on a Rotating Disk with Jet Impingement," Proceedings of Gas Turbine and Aeroengine Congress and Exposition, Toronto, 1989.

Millsaps, K. and Pohlhausen, K., "Heat Transfer by Laminar Flow from a Rotating Plate," *Journal of Aeronautical Science*, Volume 19, 1952, pp. 120-126.

Ong, C.L. *Computation of Fluid Flow and Heat Transfer in Rotating Disc Systems*, Ph.D. Thesis, University of Sussex, 1988.

Owen, J.M. and Rogers, R.H., *Flow and Heat Transfer in Rotating-Disc Systems*, Research Studies Press Ltd., Somerset, England, 1989, pp. 65-68, 93-124 and 147-149.

Parameswaran, S. and Sun, R., "Numerical Aerodynamics Simulation of Turbulent Flows Around a Car-like Body Using the Non-Staggered Grid System," AIAA 89-1884.

Patankar, S.V., "A Calculation Procedure for Two-Dimensional Elliptic Situations," *Numerical Heat Transfer*, Vol. 4, 1981, pp. 409-425.

Peric, M., Kessler, R. and Scheuerer, G., "Comparison of Finite Volume Numerical Methods with Staggered and Colocated Grids," *Computers and Fluids*, Vol. 16, p. 389, 1988.

Popiel, C.O. and Boguslawski, L., "Local Heat Transfer from a Rotating Disk in an Impinging Round Jet," *Journal of Heat Transfer*, Volume 108, 1986a, pp. 357-364.

Popiel, C.O. and Boguslawski, L., "Badanie Procesow Konwekcji Masy i Ciepla Metoda Sublimujacego Naftalenu," 1986b.

Phadke, U.P. and Owen, J.M., "An Investigation of Ingress for a 'Air-Cooled' Shrouded Rotating Disk System with Radial-Clearance Seals," *Journal of Engineering for Power*, Volume 105, 1983, pp. 178-183.

Rhie, C.M. and Chou, W.L., "Numerical Study of the Turbulent Flow Past an Airfoil with Trailing Edge Separation," *AIAA J.*, Vol. 21, p. 1525, 1983.

Sparrow, E.M. and Gregg, J.L., "Mass Transfer, Flow and Heat Transfer about a Rotating Disk," *Journal of Heat Transfer*, Volume 82, 1960, pp. 294-302.

Van Doormaal, J.P. and Raithby, G.D., "An Evaluation of the Segregated Approach for Predicting Incompressible Fluid Flows," ASME 85-HT-9.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Thin Film Behavior of Powder Lubricants
Mixed with Ethylene Glycol

Prepared by:	Don W. Dareing, P.E., Ph.D.
Academic Rank:	Professor
Department and	Mechanical Engineering
University:	University of Florida
Research Location:	WRDC/POSL Wright-Patterson AFB Dayton OH 45433-6563
USAF Researcher:	Ronald D. Dayton
Date:	Aug 90
Contract No:	F49620-88-C-0053

THIN FILM BEHAVIOR OF POWDER LUBRICANTS

MIXED WITH ETHYLENE GLYCOL

by

Don W. Dareing

ABSTRACT

The rheological properties of two powder lubricants mixed separately with ethylene glycol were determined experimentally. The two powder lubricants were graphite and molybdenum disulfide. Bingham plastic, power law and hyperbolic rheological models were compared with the experimentally determined shear stress-shear rate rheological data. All three models compare favorably with experimental data at high shear rates. The power law and hyperbolic rheological models are realistic candidates at low shear rates. The three rheological models were incorporated into fluid flow equations and solved for the case of pressure induced flow between parallel surfaces. Predicted flow rates of these non Newtonian slurries were checked against laboratory data which was obtained through a test rig designed and built during the study.

ACKNOWLEDGEMENTS

I wish to thank the Air Force Systems Command for giving me the opportunity to conduct research at Wright-Patterson. I appreciate the kindness and courtesy at all levels of activity. It was an honor to associate with the military professionals and a pleasure to watch our fine young adults in action on the Base.

I wish to thank the Air Force Office of Scientific Research for sponsoring the Summer Faculty Research Program. The program was especially beneficial to me as I have recently transferred to the university from industry. The program helped me focus on specific research that is of interest to the Air Force and that has the potential for long term funding.

I would like to thank Mr Don Campbell, Laboratory Director, for allowing me to be a part of the propulsion research program during the summer. Mr Benito Botteri, POS (Chief, Fuels and Lubrication Division) and Mr Howard Jones, POSL (Chief, Lubrication Branch) made me feel a part of their research team.

I am especially grateful for the direct involvement of Mr Ronald D. Dayton, POSL (TAM, Lubrication Branch), in the research. Mr Dayton suggested several research problems to consider during my pre-summer visit, provided laboratory space and equipment, and supported the project throughout the summer. In addition, he

made significant technical input to the research through periodic discussion of the work.

Mr Nelson Forster and Mr John Schrand took time from their projects to roundup equipment and instrumentation so that I might achieve my project goals during the ten-week summer period. For this support, I am most grateful.

INTRODUCTION

Tribology and lubrication of machine elements are an important consideration in the design and operation of machinery. Equipment literally "wears out" from lack of proper lubrication. I have been interested in this area of technology since conducting research on the lubrication of gear teeth for my Ph.D. (1962) at the University of Illinois. I have written several papers on lubrication films and have conducted research at Cambridge University on the subject of squeeze film damping as a visiting researcher. Two summer assignments with Pratt & Whitney dealt with structural stiffness of rolling contact bearings and traction coefficients of transfer films in contact with soft metal coatings. Also, I teach a graduate level course in lubrication of bearings at the University of Florida.

The Summer Faculty Research Program has given me the opportunity to become directly involved in tribological research that is of interest to the Air Force. Through the suggestion of Mr Ron Dayton, I chose to work on powder lubricant slurries. This area of research is unique in that the slurries are non Newtonian. The summer effort focused on measuring the rheological behavior of powder slurries, evaluating different rheological math models, and developing fluid flow equations for pressure induced flow between parallel surfaces. The long term application of the research is the Integrated high Performance Turbi ⁺ Engine Technology (IHPTET) program.

OBJECTIVES OF THE RESEARCH EFFORT

The objective of the research was to determine the thin film velocity and flow characteristics of two powder lubricants mixed separately with ethylene glycol. This goal is an initial step toward formulating equations of lubrication which include the effects of these non Newtonian slurries. Steps taken during the summer research program were:

A. Establish the rheological properties of two lubricant slurries through testing.

B. Evaluate mathematical rheological models against measured data.

C. Develop fluid flow equations for predicting velocity profiles and flow rates for non Newtonian slurries.

D. Check predicted flow rates against measured flow rates using test equipment designed and built during the summer program.

SUMMARY OF RESEARCH CONDUCTED

UNDER SUMMER FACULTY RESEARCH PROGRAM (1990)

A. Rheology Data

Shear stress versus shear rate relationships for the powder lubricant slurries was determined experimentally by using a Haake Rotovisco RV100 viscometer. This instrument has the capability of measuring the shear stress-shear rate relationship continuously from zero rate to a maximum shear rate of about 1000 sec-1 depending upon the particular sensor system used.

The viscosity of ethylene glycol was determined early in the study and the measured value checked with the book value of 23.0 cp at 60°F (Reference 1). Rheology data for ethylene glycol was determined for four temperature levels. These data provided a baseline reference for judging the effects of powder lubricant additives on rheology. The effects of adding molybdenum disulfide (MoS_2) to ethylene glycol are shown in the test data Figure 1. The data corresponds to a slurry mixture of one part (weight) of MoS_2 to one part (weight) of ethylene glycol. This mixture ratio (1:1) was arbitrary. These data show how shear stress changes with temperature. Note that the relationship between shear stress and shear strain is no longer linear.

RHEOLOGY DATA

MoS2 – Ethylene Glycol (1:1)
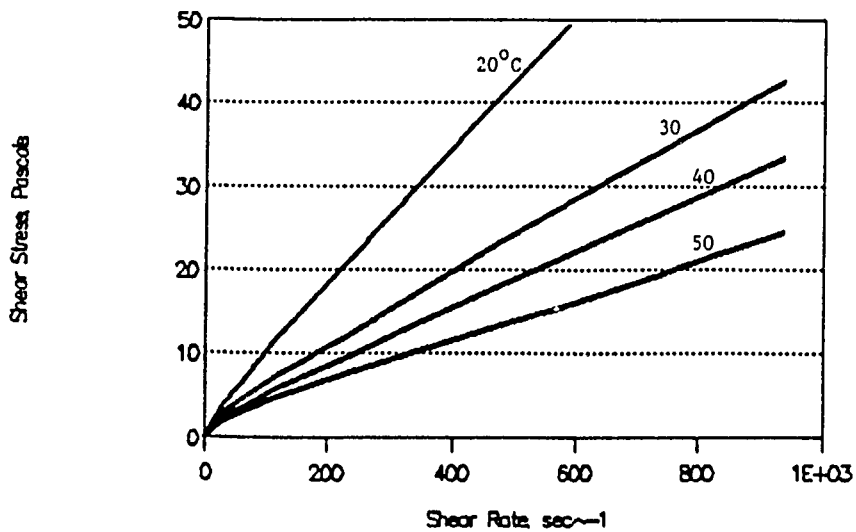
FIGURE 1

RHEOLOGY DATA

Graphite–Ethylene Glycol (1:8)
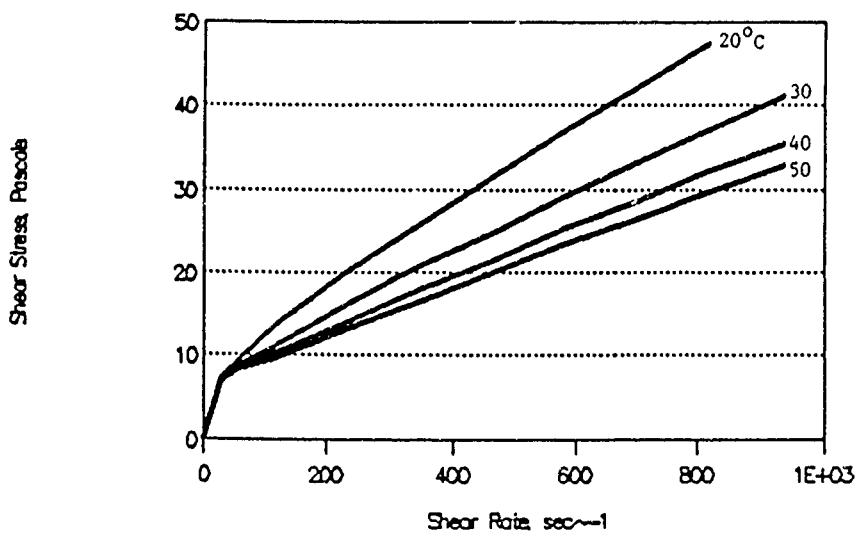
FIGURE 2

The effects on rheology of adding graphite to ethylene glycol are shown in the test data given in Figure 2. The data corresponds to a slurry mixture of one part (weight) of graphite to eight parts (weight) of ethylene glycol. This mixture ratio (1:8) was also arbitrary, however, higher ratios were very thick and difficult to test. The main characteristic of this slurry is the high degree of nonlinearity of the curves at low shear rates. These slurries are thixotropic as they shear thin at high shear rates.

B. Rheology Mathematical Models

Three different mathematical models were considered for approximating the experimental rheology data. Each of these models are discussed below.

1. Bingham Model: This model (References 2,3) is mathematically the simplest of the three because it is a straight line intersecting the shear stress axis at τ_0 called the yield point.

$$\tau = \tau_0 + K_0 \dot{\gamma} \tag{1}$$

This model describes a substance that has no shear rate until a certain shear stress level has been reached. This model deviates considerably from the measured data at low shear rates as the measured data shows no yield point. However, the model is accurate at high shear rates where many lubricating films operate.

2. **Power Law Model:** This model (Reference 2) is so named because the shear rate variable is raised to an "n th" power as defined in Equation 2. The model is useful for describing shear stress-shear rate relationships for non Newtonian fluids that have no yield point. This is the case here. The model is amenable to fluid flow formulations.

$$\tau = K \dot{\gamma}^n \tag{2}$$

The power law model may be most useful in predicting laminar flow through capillary tubes. This model also covers the full range of shear rates.

3. **Hyperbolic Model:** This model can be formulated to match the experimental data from zero shear rate to very high shear rates. However, its mathematical expression is more complex and thus velocity profile and flow rate predictions are more difficult to obtain.

$$\tau = [K_0 \dot{\gamma}^2 + 2K_0 \tau_0)]^{1/2} \tag{3}$$

This function approaches the Bingham line at high shear rates and matches reasonably well with the experimental data at low and near zero shear rates. The constants on the mathematical expression are the same as the ones used in the Bingham model expression.

C. Fluid Flow Equations

The fluid flow equations were developed for pressure induced flow between parallel surfaces. The purpose here is to gain insight into the behavior of these non Newtonian lubricant slurries in thin films. Velocity induced flow formulation and Reynolds equation follow directly from these equations.

The velocity profiles for the molybdenum disulfide and graphite slurries (Figures 3 and 4) show the effects of the nonlinearity of the shear stress-shear rate curves (Figures 1 and 2). The conditions for these curves are also

$$h = 0.0005 \text{ inches}$$

$$dp/dx = 20 \text{ psi/inch}$$

$$\text{Temp} = 20°C$$

The velocity profile curves in this case are more blunt near the center of the film. This effect would be amplified by larger values of τ_0. The zone of plug flow for the Bingham model is

$$y_0 = \frac{\tau_0}{\dfrac{dp}{dx}} \qquad (4)$$

The extent of the plug flow zone is identified by the point "a" in the figure. The hyperbolic model profile more nearly approximates the true velocity profile because

VELOCITY PROFILES

MoS2 – Ethylene Glycol (1:1)
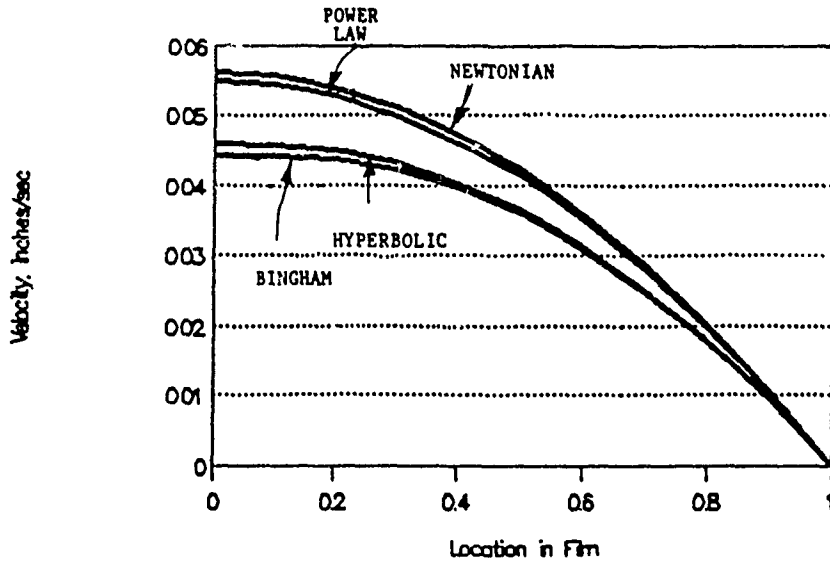
FIGURE 3

VELOCITY PROFILES

Graphite–Ethylene Glycol (1:8)
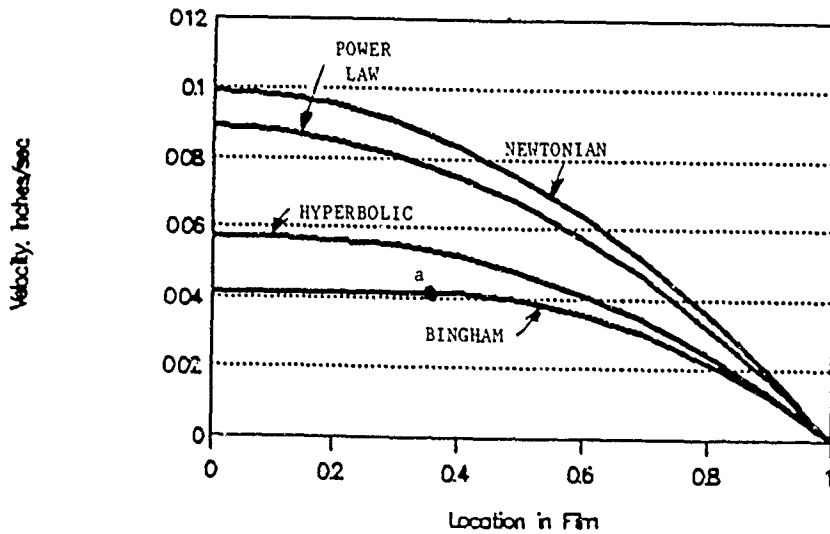
FIGURE 4

this model more nearly fits the shear stress-shear rate test data at low shear rates (0 to 1000 sec-1).

Flow rate calculations were made using the above rheological models. An example set of calculations are shown in Figure 5.

The ethylene glycol case was determined from classical Newtonian equations of flow (Reference 4). For the two powder slurries, all three rheological models predict essentially the same flow rates. This is due to the fact that shear rates within the films are far in excess of the 1000 sec^{-1} and in the operating region where all three models converge.

Flow Rate Predictions

Film Thickness is 0.002 inches
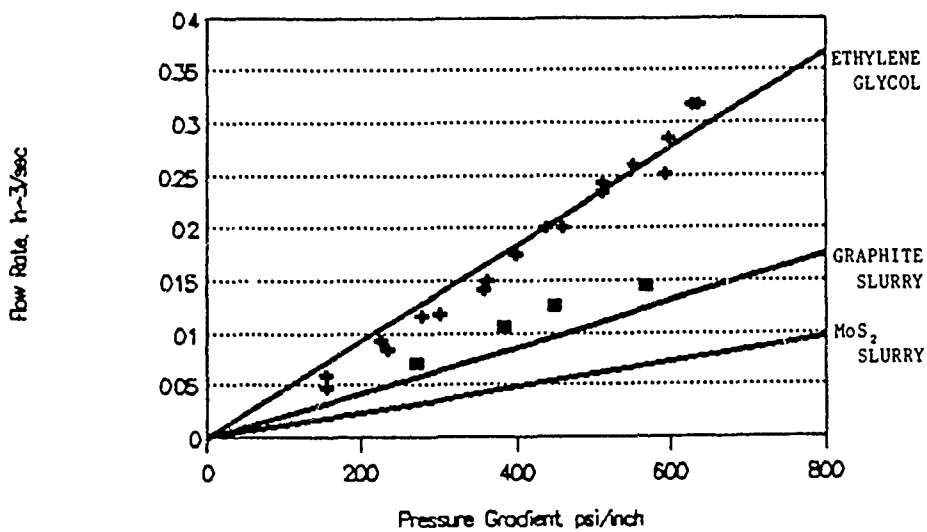
FIGURE 5

D. LABORATORY TESTING

A laboratory test rig was designed and built to check fluid flow predictions based on the rheological data and equations of flow given earlier. The rig was designed to generate pressure induced flow through parallel surfaces.

Parallel surfaces were approximated by an annular space having a circumference of two (2) inches. The length of the inner cylinder was one inch so that pressure gradient could be ascertained directly from pressure measurements which were determined from a load cell.

Flow across the parallel surfaces were established by advancing a piston by means of a power screw driven by a 1/2 horsepower motor thru a four to one timing belt pulley pair. This arrangement provided a positive displacement pump having the capability for low volume delivery at high pressures. After a test fluid is forced through the parallel surfaces, the piston is reset and the cylinder recharged for another test run. A tachometer gave volume flow while a load cell mounted in tandem with the piston rod gave fluid pressure.

The test rig allowed flow rate versus pressure data to be collected for small samples of the powder lubricant slurries.

Data from these experiments are also given in Figure 5. In general there was reasonably good agreement between theory predictions of flow and the test data.

RECOMMENDATIONS

The summer research established velocity and flow rates for pressure induced flow through parallel surfaces for two powder lubricant slurries. This was a useful first step, because pressure induced flow is an important component of flow in lubrication films. The extension of this research should combine the results of pressure induced flow with velocity induced flow to formulate equations of lubrication. Recommended future work is outlined below:

A.   Develop Reynolds equation of lubrication for rheological models which closely match the experimental data obtained under the SFRP (1990).

B.   Evaluate techniques for solving Reynolds equation for Non Newtonian fluids including methods

   1. found during extensive literature survey.
   2. developed during the research

C. Solve Reynolds equation for:

   1. Hydrostatic type bearing.
   2. Squeeze film type bearing.
   3. Slider bearing (no side leakage).

D.   Design and build test equipment to measure performance of the above three types of bearings.

E. Compare theory predictions of film performance against laboratory test data. Laboratory data should be obtained under different operating conditions for the three different types of lubrication film.

Longer term work should focus on the behavior of powder lubricant slurries in the elastohydrodynamic lubrication of rolling contact bearings.

REFERENCES

1. Windholz, M., <u>The Merck Index</u>, 10th Edition, Merck and Co., Inc., 1983.

2. Skelland, A. H. P., <u>Non Newtonian Flow and Heat Transfer</u>, John Wiley and Sons, 1967.

3. Cameron, A., <u>Principles of Lubrication</u>, John Wiley and Sons, 1966.

4. Fuller, D. D.: <u>Theory and Practice of Lubrication for Engineers</u>, 2nd Edition, John Wiley and Sons, 1984.

<u>NOMENCLATURE</u>

dp/dx	-	Pressure gradient
h	-	Film thickness
K	-	Power Law constant
K_o	-	Bingham model constant
n	-	Power Law exponent
y_o	-	Limit of plug flow
τ	-	Shear stress Pascals
τ_o	-	Yield stress in Bingham model
$\dot{\gamma}$	-	Shear rate, sec^{-1}

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Laser Velocimetry Measurements in Shock Tubes

Prepared by:	Paul A. Dellenback
Academic Rank:	Assistant Professor
Department and	Civil and Mechanical Engineering
University:	Southern Methodist University
Research Location:	WRDC/POTC Wright-Patterson AFB Ohio 45433
USAF Researcher:	Major Geoffrey Jumper, USAF
Date:	July 27, 1990
Contract No:	F49620-88-C-0053

Laser Velocimetry Measurements in Shock Tubes

by

Paul A. Dellenback

ABSTRACT

The overall objective of this research was to measure turbulence intensities in shock tubes with non-intrusive optical instrumentation. Laser velocimetry measurements applied to shock tube driven flows present several new problem areas that require attention. Among these are the implementation of appropriate circuitry for triggering and controlling high speed data acquisition, and the development of a suitable seeding material for these flows. Control circuitry was readily developed and demonstrated, but equipment limitations severely impeded success in resolution of particle seeding issues.

Acknowledgements

I. Introduction

The Air Force and those companies that build its propulsion engines are continuously involved in concentrated efforts to improve gas turbine engine performance. For the last decade, much of their attention has focused on higher turbine operating temperatures. Enhanced engine performance requires, among other things, a detailed understanding of both the convective heat transfer and the flow fields associated with the high temperature turbine components. The relationship between free stream turbulence intensity and local convective heat transfer rates has especially been of interest in recent years.

The overall research effort ongoing in this country to understand the interaction between convective heat transfer and free stream turbulence levels has many different facets. One component of that program is an in-house effort at WPAFB to directly measure both quantities in shock tubes designed to accommodate full scale rotating turbines. The motivation for using shock tubes is to reduce testing expenses to manageable levels, but at the same time, several new complications are introduced beyond those normally encountered in steady flow wind tunnel test facilities. Identification of, examination of, and resolution of these complications has been ongoing at WPAFB, and the purpose of this SFRP effort was to both contribute to this development effort and to initiate actual measurements of turbulence intensity in a full scale shock tube facility at Calspan Corp. in Buffalo, NY.

The author's possesses a solid background in the field of optical measurements as applied to turbulent flows, while at the same time, the author's principle research interests are in experimental convective heat transfer. Thus, there was an ideal match between the author's qualifications and the problem needs associated with the problem discussed above. Major Geoffrey Jumper, USAF, recognized this match and hosted the author's studies at WPAFB to assist in the shock tube velocimetry investigations.

II. Objectives of the Research Effort

The overall research objectives as outlined by Major Geoffrey Jumper included;

a) becoming familiar with use of POTC lab's laser two-focus (L2F) velocimetry system,

b) developing, in conjunction with technical support from Mr. John Schmoll, suitable triggering circuitry for use in a shock tube,

c) forwarding the work of Rivir, Elrod, and Dunn[1] by examining flow seeding alternatives suitable for the high temperature (1000 deg F) environments that were anticipated in Calspan's full scale shock tube,

d) travelling to Buffalo, NY for a two-three week period for the purpose of performing a series of measurements in Calspan Corp.'s full scale turbine test rig (a large scale shock tube),

e) returning to WPAFB for final data manipulation and writing up of findings and results.

The above series of events were viewed, even from the outset, as an ambitious program for a ten week time frame by both Major Jumper and the author. Unfortunately, equipment limitations severely compromised the overall success of achieving the objectives stated above.

III. Experiences with the L2F System

The only aspect of the laser two-focus (L2F) velocimeter familiar to the author at the outset of the project was the basic operating principle whereby two parallel laser beams separated by a known distance are used to monitor the time-of-flight associated with micron-sized particles passing through one beam and then the other. The particle velocity is then computed by dividing the beam spacing by the time-of-flight. The inherent advantages associated with L2F techniques include the ability to measure very near walls and generally superior signal-to-noise ratios (SNRs) compared with more conventional laser Doppler anemometry. However, the L2F measurement is very directionally sensitive so that only particles traveling through a very narrow band of angles will intercept both beams. The L2F technique thus has historically been limited

primarily to flows that have a single dominant flow direction and to those having fairly low levels of turbulence. The L2F is limited to these applications in order to assure that a single particle crosses both beams, as opposed to the ambiguous case where one particle passes through the first beam and a second particle passes through the second beam. This latter series of events becomes more probable for multi-dimensional or highly turbulent flows.

Familiarization with the basic Polytec L2F velocimeter occurred during the first week of the program. The shock tube facility to be used in support of the above-stated seeding and firing circuit development was located in the Air Force Institute of Technology (AFIT), and due to its primary role of teaching support, was unavailable during the initial three weeks of the project. Hence, initial investigations into L2F operations were conducted in a small (1 cm diameter) free jet seeded with droplets of propolene glycol.

Two problems were encountered during early L2F usage. The particular instrument used here contained two photomultiplier tubes (PMTs), with each monitoring the waist region (i.e. smallest diameter region) of a single beam. Thus, outputs from the two beams can be, and in fact must be, simultaneously monitored for alignment purposes as well as for correctly adjusting the controls of the dedicated L2F signal processor. Dr. R. Rivir, the last user of the L2F instrument, reported before the project began that one of the two PMTs was misaligned so that it was not obtaining *any* signal from the beam it was intended to monitor. However, with some good fortune, both PMTs were initially in alignment so that scattering signals were in fact being received. Unfortunately though, one of the two signals was oddly distorted so that rather than appearing gaussian as expected, it appeared somewhat like a cresting wave. This shape clearly would negatively impact the overall accuracy of the L2F measurements (albeit to a small degree) since the pulse shaper and discriminator in the signal processing electronics would be "confused" by the peculiar slope of the signal, thus producing a systematic

timing error. By swapping the two PMT's and their associated support circuitry in various ways, it was determined that the source of the misshapen signal was the receiving optics alignment. By disassembling the receiving optics and realignment, the problem was eventually resolved.

A second obstacle to immediate success was more readily managed. The rather complicated software supplied by Polytec is intended to be used in steady flows in which the plane containing the two laser beams can be leisurely rotated through 90 degrees to allow determination of the predominant flow direction. Unfortunately, the software will only work in this mode. Clearly then, the software is not satisfactory for transferring data from the dedicated L2F electronics and processing for the shock tube measurements because no such rotation of beam orientation is possible in the 10-40 msec time window for data acquisition. A scheme to "fool" the software by supplying arbitrary and artificially generated angular data was conceived and tested to assure that it produced no adverse effects on accurate data reduction. This scheme was determined as sound and subsequently used to reduce one-dimensional shock tube data. It perhaps should be noted that a data reduction program had previously been written by a former student for use with the L2F system, but critical inspection of this routine by Mr. Schmoll and the author established that there were at least eight conceptual and algorithmic errors in the code so that its use was abandoned as unfeasible without a substantial effort towards revision.

It should also be noted that neither data reduction program discussed above appears to incorporate a correction scheme for the velocity bias problem associated with finite-sized probe volumes, even though by the mid-1980s this effect has been established as a non-trivial contributor to errors in calculation of mean velocity and turbulence intensity[2].

The beam diameters at the waists and the beam spacing were measured as per the Polytec instruction manual by projecting the measurement volume through a graduated microscopic slide and a microscope objective lens. In this way, the beam spacing was

determined to be 350 microns and the beam diameter was determined to be about 12 microns.

The L2F does not lend itself to an easy determination of data rate since pairs of signals received by the PMTs must be correlated with many other pairs to determine the likelihood that a single particle has produced both scattering events. Thus, a data acquisition rate can only be determined at the end of a timed period by counting the number of properly correlated events. However, it is a straightforward matter to count the particle crossings seen by each PMT, and the Polytec L2F instrument provides such an on-line readout of "event rates". The event rates will always be larger than the data rate. The data rates are of critical interest in the present study since it is desired to perform a statistical analysis on the data (i.e. to determine the RMS fluctuation of velocity, that when normalized is referred to as "turbulence intensity"). Furthermore, the error associated with the sample size is inversely proportional to the square root of the number of samples, so it is desirable to obtain several thousand such data points per location. It is somewhat tricky to exactly reproduce the conditions in a shock tube from shot to shot, so that it is further desirable to obtain several thousand data points in any one 20 to 40 msec shot. Assuming a minimum number of samples is 4000, then in a 20 msec shot (which is actually a bit long for the AFIT shock tube), one must have a data rate of at least 200,000 Hz, and event rates that are larger still. Thus, mention of event rates will occur in the discussion which follows, and in fact the search for very high data rates was the motivation for examining the effects of system gain and seed density that are discussed in the following paragraph.

The impact of seed density was investigated in the free jet. In a series of tests it was determined that although signal noise is a strong function of PMT voltage (which is like an amplifier gain in the system), the accuracy of computed mean velocity and turbulence intensities (TIs) does not appear to be negatively impacted by even very high gains that result in noisy signals. A similar systematic investigation of various seed

densities (employing propolene glycol) in the free jet likewise showed little impact on the computed velocity statistics.

The event rates in the PMT gain and seed density studies discussed above were between about 1000 and 25,000 Hz on each channel of the L2F optics. While low by the criteria stated above, velocities in the free jet were about one order of magnitude lower than behind the M=1.3 shocks that were typically generated in the AFIT shock tube facility. Because theoretical data rate is proportional to jet velocity, or how fast seed droplets are swept through the measurement space, the observed event rate of 25,000 Hz in the free jet study suggests that event rates of 250,000 Hz might be possible in the AFIT shock tube if the flow can be seeded to a similar density. It should be noted that in the free jet studies, an event rate of 25,000 Hz was produced at the expense of very noisy signals and it was surprising to the author that the ensuing mean and RMS velocities agreed so well with cleaner, less noisy signals obtained at lower seed densities and lower system gains.

Eventually, the AFIT shock tube facility became available and was reconfigured to provide a test region one inch square. The tube was essentially set up as it was when Rivir, et Al.[1] made L2F measurements during the 1983/84 time frame.

The tube was first operated in a continuous flow mode to allow for determination of beam angle. With a mean velocity of 120 m/s (the largest available from the air supply), the impact of L2F angular bias on accuracy of mean and RMS velocities was investigated. The plane containing the beams was tilted about the mean flow direction a total of plus and minus 6 degrees, in increments of two degrees (see Figure 1). When the plane of the beams was not aligned with the flow direction, the indicated mean velocity was too low by about 4% per degree of beam-plane rotation. The sense and magnitude of this bias error in mean velocity are dependant on the seed density, but the magnitude of the 4% figure is interesting as it pertains to what seemed to be fairly typical measurement conditions. The TI tended to increase with beam angle error at about 6% per degree of

beam-plane rotation over the range examined here (for TIs computed based on the actual mean velocity of 120 m/s). The TI results are qualitatively as expected, and demonstrate the potential for biasing of data due to the directionally sensitive L2F measurement space.

With this operational experience, our attention turned to making measurements during 10 msec windows behind shock waves. The window duration of 10 msec was chosen after inspecting transient, representative shock pressure histories produced by a pressure transducer located just upstream of the optical window access and in the one inch square cross section. This period is only half that used previously by Rivir, et. Al., but it was felt that our choice would provide the more accurate measure of true turbulence intensity as the pressure (and presumably the flow velocity) is clearly decreasing for all time after passage of the shock. It has been assumed that the data rates obtained in our measurements could be roughly doubled if the window for accumulating data were increased to 20 msec. The strength of the shocks for all runs was about 1.3, as determined from calculations for an ideal shock tube based on the material in Gaydon[3].

Towards the end of a series of roughly 60 experience-generating tube shots, a careful set of ten shots were contrasted. For these cases, computed mean velocities were consistently in the 245 m/s regime (plus or minus about 5 m/s). However, TI were less consistent as they were observed to vary between 3.3% and 8.4%, depending on the many operator-settable parameters employed. The inconsistency of the TI measurements and our inability to determine which subtle influences were altering these values from run to run were very disturbing. Furthermore, the number of satisfactorily correlated data that were obtained in any 10 msec window were typically 100 to 200 velocity measurements. This, based on the preceding discussion, was also very disappointing since the sample size was generally too small for a satisfactory statistical analysis.

It was felt that development efforts should proceed with a seed media that was more like that we hoped to implement in the full scale Calspan tests that would require a

seed material with higher temperature capability than the propolene glycol. Solid particle seeds were investigated because they tend to be convenient to use (e.g. as opposed to titanium tetrachloride), but there may ultimately be some concern as to their abrasiveness in the Calspan tunnel. A small swirling-flow seeder was installed in place of the six-jet atomizer. Micro-balloons particles were used initially due their well-behaved nature and ready availability (that is, micro-balloons shows a minimal tendency to clump relative to many other common solid particle seed materials and they were given to us by a UD employee). During set-up of the solid-particle seeder with continuous flow through the shock tube, it was clear that a problem had developed.

Use of the solid seed generally required higher PMT voltages due to the smaller size of the new seed. As we attempted operation at the required PMT voltages, fairly regular 0.5 volt "events" spaced at approximately 5 micro-second intervals were in obvious profusion on one channel of the L2F. This noise was not related to seeding, as ascertained by turning off the seeder. Eventually, the source of the noise was determined to be the PMT itself. This caused the L2F work to come to a halt since the nature of the noise was such that it looked like scattered light from particles passing with a frequency of 250 kHz was being accumulated by the PMT. The L2F electronics was all too happy to process these signals. A replacement PMT was ordered, but the quoted delivery period (4 to 6 weeks) was so extensive that alternative approaches to the problem at hand were clearly in need of consideration, given the short 10 week period available to complete the work.

IV. Experiences With the LDA

A laser Doppler anemometer (LDA) was available for use in these experiments. The author's background in LDA measurements and the technique's general superiority over the L2F method for making measurements of TI led to optimism that it could be

usefully employed on the present problem. It was also anticipated that the replacement PMT for the L2F might arrive towards the end of the investigator's 10 week tenure so that simultaneous, side-by-side L2F and LDA measurements might be conducted for purposes of comparison and confidence building in the L2F's TI measurement capability. It was recognized that the LDA's inherently lower SNRs might negatively impact the instrument's performance in this environment where backscatter light collection is required.

A plan was formulated whereby it would be attempted to install the LDA into the AFIT lab and make velocity measurements. Two logistic problems associated with actually using the LDA presented themselves. First, there was not a proven data acquisition system and interface available to work with the IBM XT computer that was available to support these experiments. Secondly, simultaneous measurements with both the L2F and LDA seemed to require a physically complex arrangement for support and mounting of the two systems. The second problem was to be managed by incorporating a fiber optic probe with the LDA, thus allowing remote placement of the laser and optics. The penalty to be paid was that the fiber optic probe was of a rather old design such that only very low light transmission could be obtained.

The problem associated with hardware interfaces, software interface drivers, and data acquisition and processing software was capably tackled by Mr. Schmoll, who had previously generated a substantial amount of code for exactly these purposes. Modifications to Mr. Schmoll's routine required about one week, but this step was important to determining the viability of LDA use in this investigation since there was concern about the maximum data through-put capability of the LDA processor and computer interface. Through-put rate is not a concern in the L2F instrument where data points are accumulated in real time by a multi-channel analyzer and then later dumped to the computer for processing at the user's leisure. Eventually, Mr. Schmoll's program development was sufficiently advanced that we could measure the through-put rate.

However, the software was never completed due to Mr Schmoll's efforts being required on other projects, and thus no capability of storing data and statistically processing data was available. The data through-put rate of the TSI interface was determined to be marginally adequate, but no single quantitative value was determined since the rate proved to be a function of the validated data rate determined by the counter. Generally though, the through-put rate was approximately 80% of the validated data rate for low data rates, and as little as 30% for the highest data rates examined. We guessed that in optimal conditions, a thousand or so data points could be accumulated in our 10 msec window, and if this were in fact the case, it would represent a substantial improvement over our prior experience with the L2F system.

Based on the through-put measurements, the decision to move and try the LDA was made. A delay of almost two weeks ensued as power for the LDA's laser was installed in the AFIT lab. Note that the L2F's laser and the LDA's laser had substantially different power requirements (single phase vs. three phase, respectively), thus requiring the new electrical service.

The LDA was installed and aligned. Alignment proved very tedious with that portion of the system which "piped" the laser beams into the optical fibers. With a beam collimator, two mirrors, a beamsplitter, Bragg cell, and fiber-optic link installed in the optical train, the best optical transmission efficiencies for the entire system were about 5%, thus providing a maximum of 45 mW of optical energy at the probe volume (employing the 1 watt, 488 nm line from the argon laser). When contrasted with the 200 mW available at the probe volume from the L2F system and the smaller laser beam foci associated with the latter, it is clear that the L2F system had a substantial advantage in potential SNR that could be achieved (a rough computation suggests that the SNR under conditions of these tests might be as much as 1900 times higher with the L2F). The LDA probe volume could not be successfully imaged back through the fiber-optic probe, even under ideal fixed-scattering source conditions. In tracking down this apparent alignment

problem, the fiber at the PMT head was either discovered to be severed, or was severed as a consequence of our investigations. The fiber was not repairable at that time due to a lack of tools and experience in such matters. Thus the probe volume was imaged via a standard optical housing and PMT mounting assembly. Light collection at the front end of the receiving optics was improvised with a lens having focal length of about 100 mm and aperture of 40 mm.

The make-shift LDA proved successful in operation with propolene-glycol seeded, low velocity flows. However, success with the LDA configuration used here in shock tube measurements was minimal.

One persistent problem precluded success with high velocity steady flow measurements. Shot noise from the PMT could be observed on an oscilloscope as a single, classic, spike. The PMT signals were processed in a dedicated LDA counter that was typical in that its front end contained an amplifier and bandpass filters. The amplifier apparently introduces substantial overshoot into the incoming signal so that the PMT shot noise was modified. After passing through the filtering and amplification stages of the counter, observation of the shot noise signal on an oscilloscope showed that it contained a number of distinct cycles. This signal contained an even greater number of cycles if the PMT output were first passed to the frequency downmixing circuitry of the Bragg cell power supply and then to the counter, as is the usual practice. Thus, the shot noise was being modified by the processing electronics to appear exactly like a Doppler burst! Not surprisingly, the timing circuits in the counter were incapable of detecting a difference between the artificially generated shot noise "bursts" and true velocity samples (note that the shot noise problem occurred even when the PMT aperture was masked off from all light sources). Apparent data rates due just to the shot noise contribution were measured at 330,000 Hz when measuring the time for 8 cycles per burst. Increasing the number of cycles per burst to 16 only halved this data rate - that is, the modification to the original shot noise signal contained at least 16 cycles. Counting 32 cycles per burst

was similarly a problem, but judged not feasible due to the calculated number of fringes (50) in the probe volume. Filtering of these signals was also of no use since the apparent frequencies were about 20 MHz, which fell in the range expected for steady flow measurements. For measurements behind shock waves, the combination of expected velocities (determined as about 250 m/s from L2F measurements), fringe spacing (4.89 microns), and possible frequency shifts suggested that the shot noise generated signals could be filtered out by passing those signals with frequency content greater than 50 MHz, and this could indeed be accomplished with the filter circuitry of the counter.

The final conclusions pertaining to the shot noise problem should be noted by those who intend to do LDA work at higher velocities and therefore higher frequencies: The TSI counter processor takes a noise signal that a user has every expectation will be ignored, and modifies it to appear as a plausible Doppler burst, which the counter is then happy to process. Obviously, there will be a severe biasing of data if this problem is unaccounted for. It would be interesting to see how the counter reacts to a clean pulse generated by a pulse generator. As an editorial note, it is remarkable what junk one may be sold for a $20,000 outlay (i.e the price of a TSI counter).

The shot noise problem fooled this author for some time. When discovered and filtered out for the shot tube tests, the LDA system used here proved to be inadequate for the job. An event counter was set up to record the number of samples during a 10 msec window behind the shock wave, but results from seemingly identical runs were so inconsistent that no conclusions could be reached. For several shots, several hundred data points were apparently generated, while for most runs, no bursts with frequencies above 50 MHz were recorded. Captured PMT signals on an oscilloscope suggested that the electronics was working correctly, as there were no discernable bursts in the background noise. Because they could not be repeated, those runs for which several hundred data samples were accumulated are very suspicious.

In conclusion, it appears that SNRs were simply too low for detection and processing by the counter. Note that even a ten-fold increase in probe volume intensity that could be expected from eliminating the fiber-optic link (or by employing a state-of-the-art fiber optic system) might not be sufficient to give satisfactory results. This observation can be put in perspective by noting that the SNRs estimated from oscilloscope observations of the L2F signals were often less than 10, and even after LDA system improvements, the LDA probe volume intensities (and by inference, the SNRs) would still be 190 times lower than for the L2F.

Low SNRs in these experiments were further aggravated by the necessity to use off-axis backscatter light collection. An attempt to use forward-scatter light collection was made, but the requirement to remove and clean the shock tube windows after about every 4 shots, coupled with the short focal lengths associated with the optics available to the author, rendered this effectively impossible. Specifically, the forward-scatter optics had to be realigned after every window cleaning. Realignment required a window-dirtying flow with scattering source, so that the number of shots until the next cleaning and realignment was reduced to about one or two. Any future effort to use an LDA for verifying L2F accuracy should employ forward scatter optics with sufficiently long focal lengths so that realignment is not required after each window removal and cleaning. The forward scatter optics in combination with removal of the fiber-optic link should provide SNRs on the same order as those available from the L2F velocimeter (i.e., SNRs could be expected to be at least 1000 times higher than available with the optical arrangement employed here). Alignment of receiving optics also proved to be tedious due to the make-shift receiving lens arrangement, scattered light from the rear test section wall (or window in the case of forward scatter), and to an unusually bright background glow of laser light in the test region.

V. Conclusions and Recommendations

It is clear that performing measurements in shock tube driven flows adds complication over and above that encountered in steady flows. With respect to the original objectives of the investigation, it can be noted that Mr. Schmoll built a timing and gating circuit that worked satisfactorily throughout the latter eight weeks of the study. However, the effort this summer became so bogged down in various other areas that the issue of a suitable seed material is yet unresolved.

A problem not addressed during the brief periods of testing with a solid particle seed material, but one by no means trivial, was the abrasiveness associated with the particle's high speed impact with surfaces. Although the particles are small (on the order of 0.5 to 1 micron), they are on the same order as the size of polishing grit employed to fabricate thin-film heat flux gauges on the turbines tested at Calspan. Hence, it is not clear that solid particle seeding, even if optimal for the optical measurements, would be acceptable to use in the Calspan tunnel or other tunnels due to possible abrasive action on other instrumentation.

This author's concerns associated with the accuracy of L2F techniques for measurement of TI are also unresolved. It is this author's opinion that any effort made specifically for the purpose of measuring TI, and in which the principle instrument used is an L2F, will be viewed as suspicious by the experimental turbulence and heat transfer communities. Thus, I would urge a renewed effort to collaborate the L2F's ability to accurately measure TI in this difficult flow measurement problem through side-by-side comparison with either LDA or hot-wire measurements. As shock tubes become more common for use in what were traditionally steady flow investigations (e.g. WPAFB's new ATARR facility), it will be important to establish early on the reliability of one's measurement technique.

Any future efforts to verify L2F operation with an LDA should employ forward scatter measurements to insure SNRs on the same order as those available with the L2F instrument. (It is recognized that forward-scatter optics can not be used in many actual measurements, as for example at Calspan, but it could be useful for accuracy verification in the AFIT shock tube.)

One difficulty that may arise in attempting to contrast simultaneous hot-wire and optical measurement methods is that hot-wires may not be sturdy enough to withstand bombardment by the seed particles. There is also the possibility of using the hot-wire simultaneously in a liquid droplet-seeded flow, but there will be a problem with liquid accumulation on the probe and possible "bombardment breakage" of the wire (?). There is some local experience[4] with using hot-wire techniques to measure turbulence intensities in shock tube driven flows that could be helpful to future efforts.

References

1) Rivir, R.B., M.G. Dunn, and W.C. Elrod, "Feasibility Study of the Application of a Two-Spot Laser Velocimeter to Measure Velocity and Turbulence in a Shock Tube and in a Turbine Stage", AFWAL-TR-84-2054, 1984.

2) Stevenson, W. H., H.D. Thompson, and T.C. Roesler, "Direct Measurement of Laser Velocimetry Bias Errors in Turbulent Flow", AIAA J., 20, pp. 1720-1723, 1982.

3) Gaydon, A.G. and I.R. Hurle, The Shock Tube in High Temperature Chemical Physics, Reinhold Pub. Co., 1963.

4) Rockwell, R., P. King, and W. Elrod, "An Electrical Analog Circuit for Heat Transfer Measurements on a Flat Plate Simulating Turbine Vane Heat Transfer in Turbulent Flow", AIAA Paper no. 90-2412, 1990.
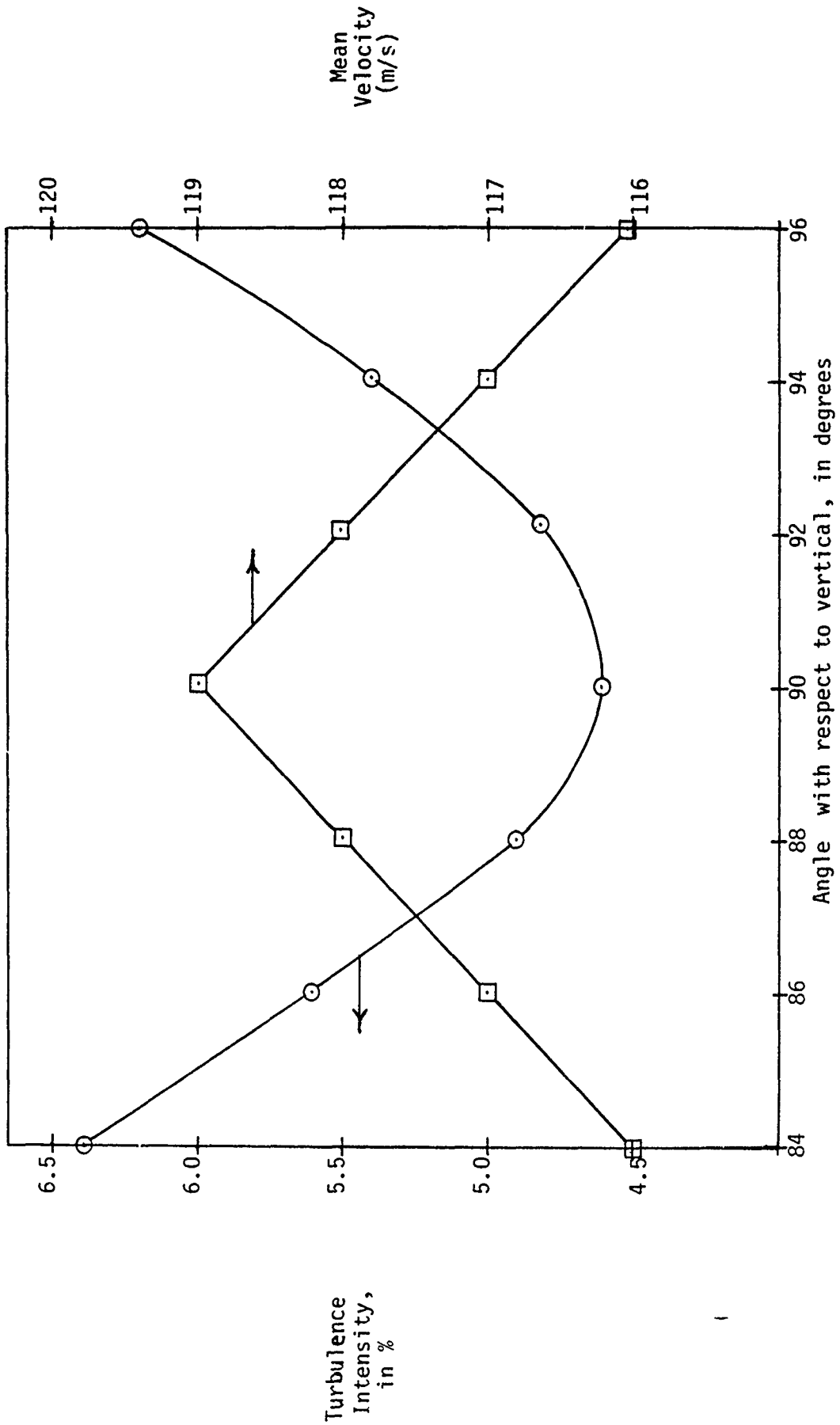
Figure 1 – Impact of L2F beam angle orientation errors on mean velocity and turbulence intensity in a steady, 1-D flow

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Thermal Analysis of Potential Solid Lubricant Candidates

Prepared by:	Dennis R. Flentge, Ph. D.
Academic Rank:	Associate Professor
Department and	Science and Mathematics
University:	Cedarville College
Research Location:	WRDC/POSL Wright Patterson AFB Dayton, OH
USAF Researcher:	Phillip Centers, Ph. D.
Date:	September 14, 1990
Contract No:	F49620-88-C-0053

Thermal Analysis of Potential Solid Lubricant Candidates

by

Dennis R. Flentge, Ph. D.

ABSTRACT

A series of tungsten, molybdenum, phosphorus, and zirconium compounds were studied using thermogravimetric analysis, differential thermal analysis, and mass spectrometry. Interaction of these materials with silicon carbide and silicon nitride were also examined. Some evidence was found for the catalytic effect of lead cations on the conversion of the carbide and nitride to silicon dioxide. Oxythiomolybdates and oxythiotungstates released sulfur dioxide when heated and promoted the conversion of silicon carbide and silicon nitride to silicon dioxide.

ACKNOWLEDGEMENTS

I. INTRODUCTION

The lubrication laboratory in the Propulsion Laboratory of the Wright Research and Development Center at Wright Patterson AFB has been investigating a series of compounds which could be used as high temperature lubricants. During the past six years I have studied the degradation of polyphenyl ethers (a class of liquid lubricants) and performed some preliminary studies on antimony sulfides (a class of solid lubricants). My graduate studies were in the area of surface catalysis. This background prepared me well for the research conducted this summer.

II. OBJECTIVES OF THE RESEARCH

The United States Air Force has a research goal of substantially increasing the output of its turbine engines during the next 15 years. This increase in output will be accompanied by an increase in operating temperature which will place an increased demand on the performance of the lubricant being used. Several solids are under consideration as potential lubricant candidates or as precursors of acceptable candidates. Since molybdenum compounds have shown good lubricating capabilities at lower temperatures several molybdenum compounds and their analogues are being considered for the higher temperature use. In this research a series of molybdates, oxythiomolybdates, tungstates, oxythiotungstates, phosphates, and zirconates were studies using thermogravimetric analysis, differential thermal analysis, and mass spectrometry.

Our primary goals were to (1) determine the thermal and oxidative behavior of the molybdates, oxythiomolybdates, and analogous compounds, and (2) determine the nature of the chemical and physical reactions that occur between the surfaces of the these materials and silicon carbide or silicon nitride.

III. EXPERIMENTAL

The thermal analysis data were collected using the DuPont Thermal Analyst 2100 System with the 951 Thermogravimetric Analyzer and the 910 Differential Scanning Calorimeter. Gas phase products generated during the thermogravimetric analysis were analyzed using the VG Quadrupoles Micromass PC.

Lead tungstate, lead molybdate, lead phosphate, cadmium tungstate, cadmium molybdate, zinc molybdate, and zinc zirconate were synthesized by Comprehensive Research Chemical Corporation. Cesium oxytrithiomolybdate was synthesized by Pennwalt Chemical Corporation. The remaining materials were synthesized for the lubrication laboratory under an Air Force contract.

Mixtures of silicon nitride (Si_3N_4) or silicon carbide (SiC) with the substances studied were prepared in a 1:1 mass ratio of substance to nitride or carbide.

IV. RESULTS

The compounds studied can be grouped into two categories--compounds containing no sulfur and oxythiomolybdates or oxythiotungstates.

A. SULFUR-FREE COMPOUNDS

Table I contains the thermal and mass spectral data for the sulfur free compounds. Cadmium tungstate, magnesium tungstate and lead tungstate show little mass loss when heated. Only the lead tungstate shows a mass gain when SiC or Si_3N_4 were added to the pure material. These mass gains are accompanied by the liberation of heat and generation of carbon dioxide (CO_2) or nitrogen monoxide (NO).

Table I Tungstates and Analogous Compounds—Thermal and Mass Spectral Data

	TGA mass % change*	MS Gases Formed	Percent Conversion to SiO_2
1) $CdWO_4$	<1	None	--
w/SiC	+3	CO_2	12
w/Si_3N_4	1.5	NO	11
2) $MgWO_4$	<1	None	--
w/SiC	+3	CO_2	12
w/Si_3N_4	+1.5	NO	11
3) $PbWO_4$	<1	None	--
w/SiC	+15	CO_2	60
w/Si_3N_4	+12.5	NO	90
4) $Pb_3(PO_4)_2$	-7.5	None	--
w/SiC	+25	CO_2	100
w/Si_3N_4	+12.5	NO	90
5) $ZnZO_4$	<1%	None	--
w/SiC	+2	CO_2	8
w/Si_3N_4	+0.5	NO	4
6) $LiO.MoO_3$	+3	None	--
w/SiC	+22	CO_2	90
w/Si_3N_4	+12	NO	90

*The mass changes that occur with SiC and Si_3N_4 represent the actual increase in mass percent during the oxidation of the carbide or nitride.

The mass gain in each of these situations is the result of the oxidation of silicon carbide or silicon nitride to silicon dioxide. Since significant conversion of the carbide and nitride occurs only with the lead tungstate, lead cations may catalyze the conversion.

Mixtures of lead phosphate and the carbide or the nitride also produce large conversions to the oxide. Lead cations may play an important catalytic role in these reactions as well.

B.   OXYTHIOMOLYBDATES AND OXYTHIOTUNGSTATES

Table II summarizes the thermal and mass spectral data for the oxythiomolybdates and oxythiotungstates studied. The mass loss for each of the compounds is consistent with the amount needed to remove all of the sulfur from the anions in all of the substances except antimony oxytrithiotungstate. In that substance more than 75% of the sulfur is removed so at least some of the product could be antimony tungstate. Each of the compounds showed significant interaction with the silicon carbide and the silicon nitride as evidenced by the oxidation of the ceramics to silicon dioxide.

Table II Oxythiomolybdates and Oxythiotungstates--
Thermal and Mass Spectral Data

	TGA	MS	
	Mass Percent Change*	Gases Formed	Percent Conversion to SiO_2
Cs_2MoOS_3	-13	SO_2	--
w/SiC	+14	CO_2	56
w/Si_3N_4	+5	NO	36
Cs_2WOS_3	-8	SO_2	--
w/SiC	17	CO_2	70
w/Si_3N_4	8	NO	60
$Sb_2(WOS_3)_3$	-11	SO_2	--
w/SiC	6	CO_2	24
w/Si_3N_4	2.5	NO	17
Zn_2MoOS_3	-40	SO_2	--
w/SiC	14	CO_2	56
w/Si_3N_4	11.5	NO	85
$Zn_2MoO_2S_2$	-22	SO_2	--
w/SiC	8	CO_2	32
w/Si_3N_4	5	NO	36

*The changes that occur with the SiC and Si_3N_4 represent increases associated with the oxidation of the carbide and nitride and occur after the oxythiomolybdates and oxythiotungstates have lost mass.

V. RECOMMENDATIONS

Analytical studies should be conducted to determine the composition of the reaction products generated when the oxythiomolybdates and oxythiotungstates are heated.

Scanning electron microscopy should be used to determine the crystalline nature of the products.

Infrared spectroscopy should be used assist in the identification of the reaction products.

Salts containing lead, cesium, zinc, cadmium, antimony, or magnesium cations could be used in conjunction with the substances examined in this study to determine whether any of them functions as a catalyst for the conversion of silicon carbide and silicon nitride to silicon dioxide.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.,

FINAL REPORT

EFFECT OF EVAPORATION ON THE DRIVING CAPILLARY
PRESSURE IN CAPILLARY PUMPED
AEROSPACE THERMAL MANAGEMENT SYSTEMS

Prepared by: Kevin P. Hallinan, Ph.D. / David Welter
Academic Rank: Assistant Professor / Graduate Student
Department and   Mechanical and Aerospace Engineering
University: University of Dayton
Research Location: Aero Propulsion Laboratory
                    Wright-Patterson AFB
                    Wright-Patterson AFB, OH 45433
USAF Researcher: Michael Morgan
Contract No: F49620-85-C-0013

EFFECT OF EVAPORATION ON THE DRIVING CAPILLARY PRESSURE IN CAPILLARY PUMPED AEROSPACE THERMAL MANAGEMENT SYSTEMS

by

Kevin P. Hallinan
David Welter

ABSTRACT

Research has been conducted to determine the effect of evaporation on the driving capillary potential in capillary pumped heat transport devices used in aerospace thermal management. These devices primarily include heat pipes and capillary pumped loops. Current design criterion for such devices rely upon what has been termed a maximum capillary potential to evaluate the maximum heat transport limitations. Preliminary analytical results, based upon an idealized model of pores within a heat pipe evaporator wick indicate that in high powered heat pipes and capillary pumped loops where evaporator heat fluxes are approaching than 100 W/cm^2 that dynamic forces owing to evaporation from the liquid-vapor interfaces within the heat pipe evaporator wick can noticeably increase the driving capillary potential relative to static conditions (if boiling is not occurring). For apparent contact angles of less than 10° and assumed isothermal interfacial conditions, normal viscous forces were shown to affect the capillary pressure for capillary numbers greater than 10^{-5}. This conclusion is particularly true if the working liquid in these devices nearly perfectly wets the wick structure in the vicinity of the liquid-vapor interfaces. As a verification of the analytical efforts, an experimental facility has been constructed to actually measure the influence of evaporation on the capillary pressure existing at a curved, liquid-vapor interface.

Acknowledgements

I wish to thank the Aero-Propulsion Laboratory and the Air Force Office of Scientific Research for sponsorship of this research. Universal Energy Systems must be mentioned for their concern and help to me in all administrative and directional aspects of this program

My experience this summer, and that of my graduate student, David Welter, was educational. Working in laboratories as well equiped as the Thermal Systems group within the Aero-Propulsion laboratory was exciting. In particular I would like to thank Michael Morgan, Won Chang, Brian Hagar, John Leonard, Mike Ryan, and Don Reinmuller for their help in acquiring and then constructing the experimental facility to measure the dynamic capillary pressure.
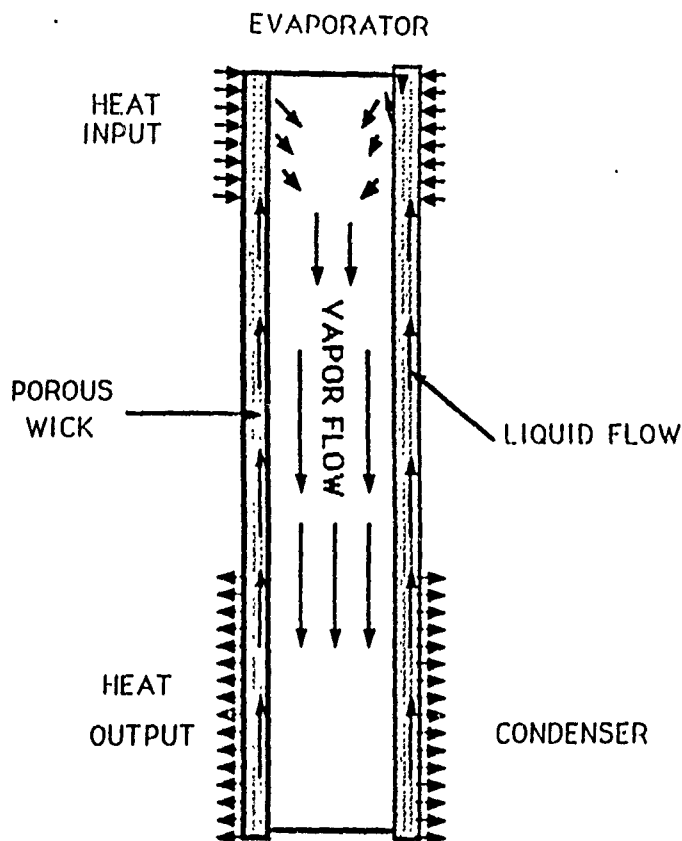
Figure 1: Schematic of a simple heat pipe.

1 Introduction

Heat pipes and capillary pumped loops are perhaps the most significant heat transport device invented to date. Their applications in all facets of society are potentially vast. In the aerospace field, their most attractive use has been related to the cooling of satellites where their extremely high heat transport capability is sufficient to handle evaporator heat fluxes of up to 100 kW/cm² offers space and weight savings relative to other types of heat exchange devices. An excellent review of the application of heat pipes in the aerospace field is offered by Chang and Hager[1]. Another important advantage of heat pipes is that they do not require any external power sources. They operate passively relying upon capillarity to transport the liquid from the condenser to the evaporator through the wick. The vapor returns to the condenser by a pressure gradient arising from a typically small temperature difference between the evaporator and condenser sections of heat pipes. Figure 1 helps to illustrate the simplicity inherent in the operation of heat pipes in their most elemental configuration.

Despite the apparent simplicity of the heat pipe operation a number of very funda-

mental issues associated with the thermal/fluid behavior within these devices is relatively unknown. Further, the phenomenon of capillarity, which is the principle mechanism inherent in the heat pipe operation is not well understood. Other problematic issues requiring further research include: porous media flows; solid-liquid-vapor interaction at the interfaces; and nucleate boiling.

The lack of a proper fundamental understanding of the operation of heat pipes and capillary-pumped devices is demonstrated by the apparent inability of current models and design criterion to predict the maximum heat transport capabilities of such devices, even in steady-state conditions. Gottschlich has given evidence that predicted performance of heat pipe operation in steady-state is often under- and over-predicted by orders of magnitude [2]. Obviously, the models predicting transient performance are even worse.

A review of the design literature for heat pipes (Dunn and Reay [3], Chi [4], and the B and K Handbook [5]) all utilized a troubling concept: the maximum capillary pressure. The maximum capillary pressure is used to evaluate when a heat pipe will fail (i.e., the evaporator wick dries out and the temperature of the pipe increases significantly). For a given heat input to the evaporator in steady conditions a certain liquid flow rate within the wick and vapor flow rate within the vapor core will be expected. The heat pipe is said to fail when the liquid and vapor pressure drops associated with the flow becomes greater than the maximum capillary pressure. Mathematically this is represented as:

$$(P_{cap})_{max} \geq \nabla p_l + \nabla p_v + \nabla p_g \tag{1}$$

Above, ∇p_g represents the gravitational head that must be overcome. In microgravity conditions this term is usually negligible. The problem with this convention is that three major assumptions are associated with the typical determination of the maximum capillary pressure. First, it assumes that static conditions exist at the liquid-vapor interfaces within the evaporator and condenser wicks. Second, it assumes that a flooded condition is present within the condenser. Third, it assumes that the liquid perfectly wets the wick within the evaporator. The first and third assumptions are not generally valid. The second assumption has yet to be verified. Thus, the basic premise of the maximum capillary pressure is questionable. Nevertheless, the collection of these assumptions leads to the following expression for the capillary pressure:

$$(P_{cap})_{max} = \frac{2\sigma}{r} \tag{2}$$

where σ represents the surface tension of the interface and r is the mean pore size within the evaporator.

Given the uncertainty in the maximum capillary pressure, the present research aims to at least answer the appropriateness of the first assumption utilized in its development. Namely, is it valid to neglect the viscous and inertial forces acting upon a curved liquid-vapor interface in the determination of the pressure drop across the interface (i.e., the capillary pressure)?

2 Objectives of the Research Effort

The goal of this study is to determine how the capillary pressure, $(p_v - p_l)$, at a curved interface is effected by viscous forces resulting from the motion of the fluid due to evaporation. In particular the investigation will concentrate on determining improved measures to predict the driving capillary forces in such devices both when film evaporation or boiling are occuring within the porous wick. Specifically, the research objectives were:

- Show that when film evaporation is occuring within the evaporator in capillary pumped thermal transport devices that inertial, viscous, and surface tension forces acting on the liquid-vapor interfaces can affect the pressure difference between the liquid and the vapor at the interface and, thus, the driving capillary pressure. The conditions characterizing the importance of these forces in actual heat pipe or capillary pumped loop operation were to be determined.

- Show theoretically that heat pipe operation can be dramatically improved by insuring that a nucleate boiling regime is stably maintained within the evaporator. A feasible mechanism is proposed which can explain why the maximum opearating conditions of heat pipe devices are affected by the start-up heating protocol. Generally, higher power transmission is obtained when the heating is increased gradually. This gradual start-up should allow for stable non-boiling operation and a higher degree of superheat within the evaporator. Eventually, when boiling does occur, smaller bubbles will be stably generated, delivering a higher capillary pressure (pressure difference between the liquid surrounding the bubbles and the vapor in the bubbles). In gravitational fields buoyant forces are responsible for removal of the bubbles from the wick. In microgravity applications surface tension gradients acting along the

bubbles are postulated to be responsible for the bubble migration through the wick to the vapor core. An investigation of this mechanism will be addressed.

The objectives of the summer research were too ambitious. The work performed during the summer primarily addressed the first goal, related to determing the dependence of the capillary pressure on fluid dynamics. The second task described above will be the subject of an additional AFOSR proposal.

3   Analytical Solution of the Dynamic Capillary Pressure

This phase of the research concentrated upon the development of a strictly analytical solution to demonstrate analytically how the capillary pressure changes at curved liquid-vapor interfaces when film evaporation is occurring. The probability that dynamic forces can be important at an interface stems from the generally accepted equation representing the balance of forces at a curved interface normal to the interface. In non-dimensional terms, this interfacial normal stress balance is given by (See Prosperetti [6])

$$P_v - P_l + Ca * \vec{n} \cdot (\tau_l - \lambda\tau_v) \cdot \vec{n} = K, \tag{3}$$

where Ca is the capillary number (defined below), τ is the dimensionless shear stress tensor, λ is the ratio of the vapor to liquid kinematic viscosity, $\vec{n}$ is the vector normal to the interface pointing to the vapor side, P is the dimensionless pressure, and K is the dimensionless interfacial curvature. In the typical definition of the maximum capillary pressure the viscous terms above are neglected. This can only be true if the capillary number is small.

Regular perturbation theory is employed to determine dynamic contributions. With this technique velocities, stream functions, and pressures (within the liquid and the vapor) are all expanded in terms of the capillary number, Ca, defined as

$$Ca = \frac{\mu_l U_l}{\sigma}. \tag{4}$$

Above μ refers to the kinematic viscosity and U refers to the average velocity normal to the interface. The subscript l refers to the liquid. The significance of the capillary number is that it has been shown to adequately represent the ratio of viscous to surface tension

forces acting upon an interface thus it represents a reasonable parameter with which the velocities and pressures can be expanded. For example:

$$P_{v,l} = (P_{v,l})_0 + Ca(P_{v,l})_1 + \cdots$$
$$\vec{u}_{v,l} = (\vec{u}_{v,l})_0 + Ca(\vec{u}_{v,l})_1 + \cdots$$

Actual values of the capillary numbers present near evaporating interfaces in real heat pipes range from 10^{-8} in water heat pipes to as high as 10^{-5} in lithium heat pipes. Apparently in all cases the capillary number is small. However, such values do not necessarily imply that viscous forces are unimportant relative to surface tension forces. For example in free convection, buoyant forces are only important relative to viscous forces when the Grashof number, defined as a ratio of buoyant to viscous forces, is on the order of 10^7. Thus, the capillary number does not have to be on the order of 1 to represent a balance of viscous to surface tension forces.

The physical model of the evaporating interface focuses on just one pore within the evaporator wick, and particularly the contact line region. A wedge type model of this region, as shown in Figure 2 is utilized similar to that employed by Huh and Mason [7] and Cox [8] in describing the unrelated problem of an advancing (non-evaporating) liquid-vapor interface within a capillary tube. The liquid-vapor interface is assumed to be isothermal (a realistic assumption considering the typically small pore sizes on the order of 10^{-5} m, making it likely that the liquid and pore structure temperature will locally be uniform). As a consequence, the evaporation rates are uniform over a majority of the interface. Only near the wall where Derjaguin[9] and Wayner[10] have shown that solid-liquid intermolecular forces for wetting fluids locally reduce the liquid pressure near the contact line and thus choke off the evaporation rate. Thus, at the wall the evaporation rate is zero. However, the lengths, s, at which these long-range intermolecular forces are important is only on the order of 50 nm (Potash and Wayner [11]) and thus this region is small relative to the entirety of the pore. It is argued by Kafka and Dussan [12], that an imprecise treatment of this near contact line region affects only slightly the flow and pressure fields outside of this region. Thus, the model employed utilized a uniform evaporation rate over most of the interface. At a distance s from the apparent contact line the evaporation rate is assumed to experience a linear decline from the uniform rate to zero at the apparent contact line. (Note that there is a wealth of literature relevant to the exact definition of the 'apparent' contact line.)

80-8
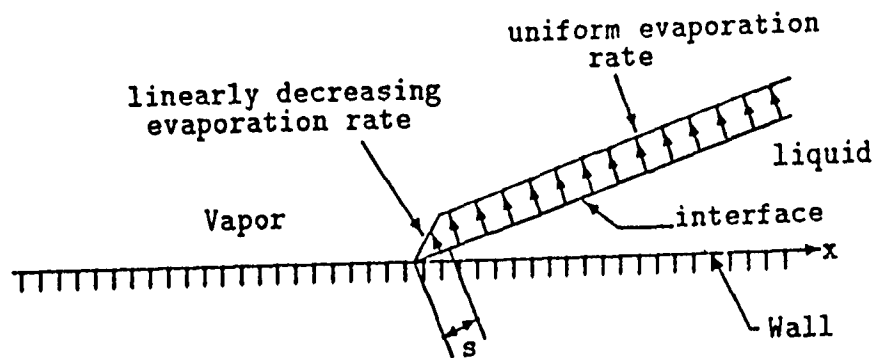
Figure 2: Wedge model used for the analysis.

Steady-state conditions are considered. Plane-polar coordinates are used, implying that the circumferential curvature of the pores is small. Low Reynolds and capillary number flows typical of heat pipes are considered. A review of a significant number of papers describing heat pipe performance has revealed that capillary numbers up to about 10^{-5} are observed in liquid metal heat pipes and up to 10^{-7} in low temperature heat pipes. Pore Reynolds numbers of below 10^{-5} are universally observed. Because of these extremely low pore Reynolds numbers associated with the evaporator liquid flow the lubrication approximation is used to describe the flow near the contact line within a representative pore in the evaporator wick.

With these assumptions the momentum equation in each phase is written as :

$$-\nabla p_{l,v} + \mu_{l,v}\nabla^2 \vec{u}_{l,v} = 0 \qquad (5)$$

The momentum equation is made dimensionless by scaling the lengths with R, the pore radius, and the velocities with U, the average liquid velocity at the interface, and dividing by the surface tension, σ. In dimensionless form the liquid and vapor momentum equations

are: (neglecting body forces which could easily be accounted for).

$$Ca\nabla^{*2}\vec{u}_l^* - \nabla^* P_l^* = 0 \tag{6}$$

$$\lambda Ca\nabla^{*2}\vec{u}_v^* - \nabla^* P_v^* = 0 \tag{7}$$

$$\tag{8}$$

The asterisk implies dimensionless quantities, λ represents the ratio μ_v/μ_l, and Ca is the capillary number. The velocity and pressure fields are expanded in terms of the capillary number for small values of the capillary number. The asterisk is dropped but all quantities are to be considered dimensionless.

$$\vec{u} = \vec{u}_0 + Ca\vec{u}_1 + \cdots \tag{9}$$

$$P = P_0 + CaP_1 + \cdots \tag{10}$$

Notice that in the dimensionless form of the momentum equation the viscous terms are multiplied by the capillary number. Thus, at $O(Ca^{+0})$ the only term present is the pressure term (and the body force term). As expected if there is no flow the pressure field is governed by hydrostatic conditions. Since our interest is in determining the first-order correction to the capillary pressure, we need only compute the zeroth order velocity fields.

Near the contact line the velocity fields within the liquid and vapor phases are subject to the following boundary conditions:

- At wall:

$$\quad - \quad \text{no-slip and impermeability} \qquad . \tag{11}$$

- At interface:

$$\quad - \quad \text{continuity of evaporative flux} \tag{12}$$

$$\quad - \quad \text{continuity of liquid and vapor tangential velocities} \tag{13}$$

$$\quad - \quad \text{continuity of tangential shear stress} \tag{14}$$

$$\tag{15}$$

Additionally the conservation of normal momentum at the interface yields an expression relating the liquid and vapor pressures at the interface.

The two-dimensionality of the problem allows for the representation of the velocity fields in terms of stream functions.

$$(\vec{u}_l)_r = \frac{1}{r}\frac{\partial \Psi_l}{\partial \phi} \quad \text{and} \quad (\vec{u}_l)_\phi = -\frac{\partial \Psi_l}{\partial r} \tag{16}$$

The stream-functions can be likewise expanded in terms of the capillary number so that the zeroth order velocities are related to the zeroth order stream functions as shown:

$$(\vec{u}_{l,0})_r = \frac{1}{r}\frac{\partial \Psi_{l,0}}{\partial \phi} \quad \text{and} \quad (\vec{u}_{l,0})_\phi = -\frac{\partial \Psi_{l,0}}{\partial r} \tag{17}$$

Taking the curl of the liquid and vapor momentum equation and representing the velocities in terms of the stream functions for both phases, yields at zeroth order the biharmonic equation:

$$\nabla^4 \Psi_{l,v_0} = 0 \tag{18}$$

The no-slip boundary conditions for $\vec{u}_{l_0}$ and $\vec{u}_{v_0}$ at the wall gives:

$$\Psi_{l_0} = 0 \quad , \quad \frac{\partial \Psi_{l_0}}{\partial \phi} = 0 \quad \text{on} \quad \phi = 0, \tag{19}$$

$$\Psi_{v_0} = 0 \quad , \quad \frac{\partial \Psi_{v_0}}{\partial \phi} = 0 \quad \text{on} \quad \phi = \pi. \tag{20}$$

$$\tag{21}$$

The evaporative flux boundary condition, represented according to Figure 2, is given by:

$$\frac{\partial \Psi_{l_0}}{\partial r} = \frac{\partial \Psi_{v_0}}{\partial r} = -\left[\frac{r}{\epsilon} + H(r - \epsilon)(1 - \frac{r}{\epsilon})\right] \tag{22}$$

where H is the Heaviside function. Continuity of tangential velocity requires that:

$$\frac{\partial \Psi_{l_0}}{\partial \phi} = \frac{\partial \Psi_{v_0}}{\partial \phi} \quad \text{at} \quad \phi = \theta_0 \tag{23}$$

The tangential shear stress continuity at the interface is represented by:

$$\lambda\left(\frac{1}{r^2}\frac{\partial^2 \Psi_{v_0}}{\partial \phi^2} - \frac{\partial^2 \Psi_{v_0}}{\partial r^2} + \frac{1}{r}\frac{\partial \Psi_{v_0}}{\partial r}\right) = \left(\frac{1}{r^2}\frac{\partial^2 \Psi_{l_0}}{\partial \phi^2} - \frac{\partial^2 \Psi_{l_0}}{\partial r^2} + \frac{1}{r}\frac{\partial \Psi_{l_0}}{\partial r}\right) \tag{24}$$

The balance of the normal stresses on the interface is given dimensionally by:

$$p_v - p_l + \vec{n}\cdot(\tau_l - \tau_v)\cdot\vec{n} = \sigma K \tag{25}$$

The curvature is defined by $K = d\theta/ds$, where θ is the angle of tangency to the interface and s is the arc length. In the wedge approximation for the meniscus, $s \to r$, and thus,

$$K \approx \frac{\partial \theta}{\partial r} \tag{26}$$

The normal stress equation then can be represented non-dimensionally by:

$$P_v - P_l + 2Ca\left[\lambda(-\frac{1}{r}\frac{\partial^2\Psi_v}{\partial\phi\partial r} + \frac{1}{r^2}\frac{\partial\Psi_v}{\partial\phi}) - (-\frac{1}{r}\frac{\partial^2\Psi_l}{\partial\phi\partial r} + \frac{1}{r^2}\frac{\partial\Psi_l}{\partial\phi})\right] = \frac{\partial\phi}{\partial r} \quad (27)$$

Since the angle of tangency, $\theta = \theta_0 + Ca\theta_1(r) + \cdots$, is assumed planar at $O(Ca^{+0})$, only at $O(Ca^{+1})$ is any curvature considered and is a result of the viscous forces. Notice again that at zeroth order only the pressure and surface tension are of consequence. Thus, neglecting body forces, at zeroth order, away from the thin film region the pressure drop across the interface is controlled by curvature. At first order the normal stress equation is given by:

$$P_{v_1} - P_{l_1} + 2\left[\lambda(-\frac{1}{r}\frac{\partial^2\Psi_{v_0}}{\partial\phi\partial r} + \frac{1}{r^2}\frac{\partial\Psi_{v_0}}{\partial\phi}) - (-\frac{1}{r}\frac{\partial^2\Psi_{l_0}}{\partial\phi\partial r} + \frac{1}{r^2}\frac{\partial\Psi_{l_0}}{\partial\phi})\right] = \frac{\partial\theta_1}{\partial r} \quad (28)$$

A solution to the biharmonic equation for both phases satisfying the boundary equations at zeroth order is given by:

$$\Psi_{l,v_0} = r\left[(A_{l,v_0}\phi + B_{l,v_0})\cos\phi + (C_{l,v_0}\phi + D_{l,v_0})\sin\phi\right] \quad (29)$$

$$+r^2\left[(E_{l,v_0}\phi + F_{l,v_0})\cos\phi + (G_{l,v_0}\phi + H_0)\sin\phi\right] \quad (30)$$

$$(31)$$

The constants A_{l,v_0} through H_{l,v_0} are obtained from the imposition of the boundary conditions in both the liquid and the vapor.

With the stream functions determined the velocity fields were computed. The pressure fields were then determined from the momentum equations up to a constant. A solution for the liquid and vapor pressure fields of the form shown was used.

$$P_{l,v_1} = \frac{h_{l,v_1}(\phi)}{r} + ln(r)i_{l,v_1}(\phi) + K_{l,v} \quad (32)$$

The functions h_{l,v_1} and i_{l,v_1} were found to be:

$$h_{l,v_1} = 2A_{l,v_0}\cos(\phi) + 2C_{l,v_0}\sin\phi \quad (33)$$

$$i_{l,v_1} = E_{l,v_0}(-\cos\phi - \phi\sin\phi) - F_{l,v_0}\sin\varphi$$

$$+ \quad G_{l,v_0}(\phi\cos\phi - \sin\phi) + H_{l,v_0}\cos\phi \quad (34)$$

The prime interest in the study lies in the calculation of the pressure jump across the interface $(P_v - P_l)$. Thus, it is only necessary to solve for the difference in the constants, $K_v - K_l$, which can be defined to be another constant, K_{lv}. The constant K_{lv} can be
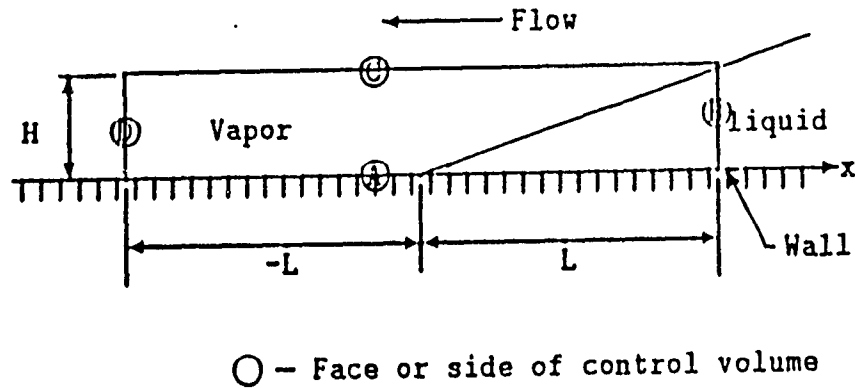
Figure 3: Control volume straddling contact line used for force balance.

determined by performing a simple force balance on a control volume straddling the contact line. The balance of forces in the flow direction yields in dimensional terms the following expression.

$$
\int_0^H \rho_l u_l^2|_{x=L} dy + \int_0^H P_l|_{x=L} dy - \int_L^0 \mu_l \frac{\partial u_l}{\partial y}|_{y=0} dx - \int_0^{-L} \mu_v \frac{\partial u_v}{\partial y}|_{v=0} dx
$$

$$
+ \int_L^{xi} \mu_l \frac{\partial u_l}{\partial y}|_{y=H} dx + \int_{xi}^{-L} \mu_v \frac{\partial u_v}{\partial y}\Big|_{y=H} dx - \int_L^{xi} \rho_l u_l v_l|_{y=H} dx - \int_{xi}^{-L} \rho_v u_v v_v|_{y=H} dx
$$

$$
- \int_0^H P_v|_{y=H} dy - \int_0^H \rho_v u_v^2|_{x=L} dy = 0 \tag{35}
$$

The inertial terms on faces A, C, and D for small Reynolds number flows are all negligible relative to the pressure terms and the viscous terms acting at the wall and on face C. The following simplified equation results.

$$
+ \ Ca\left[-\int_L^0 \frac{\partial \bar{u}_l}{\partial \bar{y}}|_{y=0} d\bar{x} - \int_0^{-L} \lambda \frac{\partial \bar{u}_v}{\partial \bar{y}}|_{y=0} d\bar{x} + \int_L^{xi} \frac{\partial \bar{u}_l}{\partial \bar{y}}|_{y=H} d\bar{x} + \int_{xi}^{-L} \lambda \frac{\partial \bar{u}_v}{\partial \bar{y}}|_{y=H} d\bar{x} \right]
$$

80-13

$$+ \int_0^H P_l|_{x=L}d\bar{y} - \int_0^H P_v|_{x=-L}d\bar{y} = 0 \tag{36}$$

Expanding the velocities and pressures in terms of the capillary number as before ($u_{l,v} = Cau_{l,v_1} + \cdots$, etc.) and substituting into the force balance equation yields expressions of the force balance at different orders in the capillary number. At order zero, static conditions (assumed known) exist. Dropping the bar depicting dimensionless parameters the zeroth order term is:

$$\int_0^H P_{l,0}|_{x=L}dy - \int_0^H P_{v,0}|_{x=-L}dy = 0 \tag{37}$$

At first order the following equation results:

$$- \int_L^0 \frac{\partial u_{l,0}}{\partial y}|_{y=0}dx - \int_0^{-L} \lambda \frac{\partial u_{v,0}}{\partial y}(y=0)dx + \int_L^{xi} \frac{\partial u_{l,0}}{\partial y}|_{y=H}dx + \int_{xi}^{-L} \lambda \frac{\partial u_{v,0}}{\partial y}|_{y=H}dx$$

$$+ \int_0^H P_{l,1}(x=L)dy - \int_0^H P_{v,1}|_{x=-L}dy = 0 \tag{38}$$

Thus, the first order correction to the capillary pressure depends again on the zeroth order correction for the velocity field which is known. Additionally the zeroth order corrections for the vapor and liquid pressures are known up to the constants K_l and K_v. The constant K_{lv} is chosen so that the equality is satisfied. This comes from the fact that the integral given by:

$$\int_0^H P_v|_{x=-L} - P_l|_{x=L}dy \tag{39}$$

can be written in terms of a known part plus an unknown contribution due to the constants. Simply, this is as shown.

$$\int_0^H [\text{Known part of } (P_v - P_l)] \, dy + \int_0^H K_{lv}dy \tag{40}$$

Thus K_{lv} is determined by requiring Equation 38 to be satisfied.

In summary the velocity, and pressure fields are computed at zeroth and first order in capillary number from the Navier- Stokes equations, subject to standard boundary conditions. The lubrication or low Reynolds number approximation of the Navier-Stokes equations is employed. Exact solutions for the velocity and pressure fields were obtained.

Results

A sampling of the results from this study are included in this report. Figure 4 shows the streamlines (dimensionless) for a case where the dynamic contact angle is 10 degrees and the kinematic viscosity ratio is 0.01. The liquid region is to the right of the interface
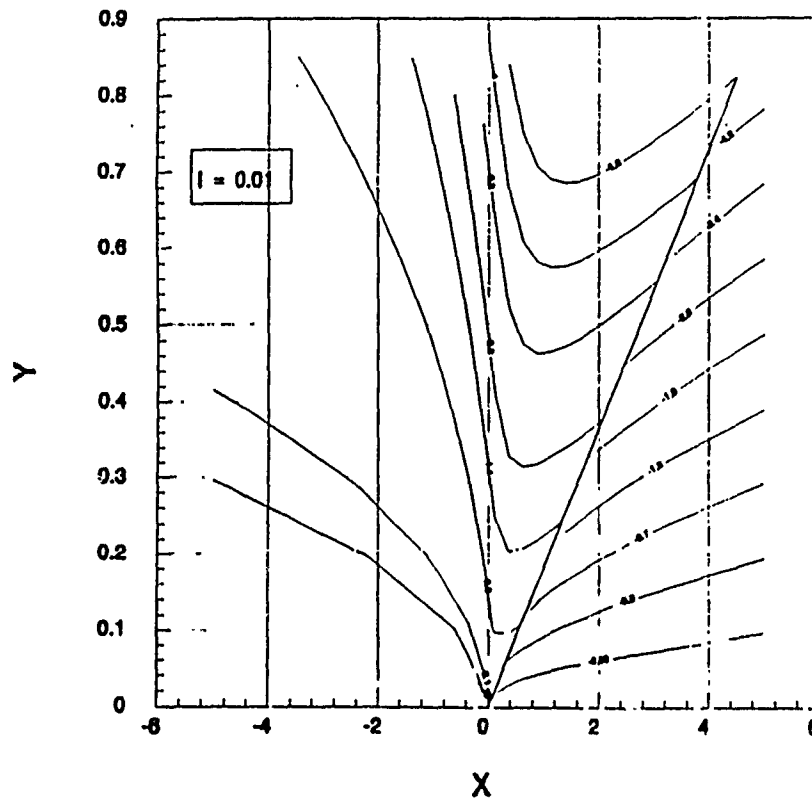
Figure 4: Representative flow streamlines near contact line at a dynamic contact angle of 10 degrees.

line. The vapor region is to the left. It is interesting to observe the curvature of the liquid streamlines toward the solid wall as the interface is approached, owing to the requirement of a uniform evaporation rate. This trend is expected. Upstream of the interface a parabolic velocity profile might be expected due to the representation of the pore as a wall, and because of the shear stresses at the wall. Due to the uniformity of the evaporation rate, the liquid velocity profile must undergo a transition from parabolic to near uniform, and thus, liquid must flow toward the wall. At the interface the liquid transitions to vapor. Continuity of tangential velocity between the liquid and vapor and the requirement of a continuous normal evaporative flux at the interface necessitates that the vapor upon leaving the interface must initially flow toward the wall. Downstream, however, the presence of the wall decelerates the flow causing the vapor streamlines to diverge from the wall.

What is more interesting is how these streamline trends affect the capillary pressure. Consider that in static conditions there is no flow. The pressure drop (from vapor to liquid) is determined from a static pressure balance yielding the expression for the maximum capillary pressure described earlier. Now, in order to have a uniform evaporation rate the

liquid has to move toward the interface. Thus the liquid pressure at the interface has to be reduced. Additionally, in order for the vapor to move away from the interface, the vapor pressure at the interface must increase above its static value. Thus, the difference between the vapor and liquid pressure across the interface must increase. This trend is shown in Figure 5 where the first order correction in the dimensionless capillary number, $(P_{cap})_1$, is presented as a function of the distance from the apparent contact line. A viscosity ratio of 0.01 is common to all four curves, representing data for 2, 5, 10, and 20 degree apparent contact angles. Two trends are observed. First, in all cases the capillary pressure increase at first order declines with increasing distance from the contact line. More importantly, the capillary pressure correction increases rapidly as the contact angle decreases. For a 2 degree contact angle, values for $(P_{cap})_1$ ranged from 2×10^7 at $R = 0.01$ to a near constant value of over 10^4 at large R. These numbers are significant relative to the zeroth order or static dimensionless capillary pressure which is on the order of 1. Capillary numbers in heat pipes of up to 10^{-5} have been observed. Since the total capillary pressure existing at an interface is the contribution of the static plus the first order: $P_{cap} = (P_{cap})_0 + Ca(P_{cap})_1$, then if $(P_{cap})_1$ is in the range reported for the smaller contact angles, it can be concluded that the fluid dynamics of an evaporating interface can in fact increase the capillary pressure existing at the interface. Thus, in high-powered heat pipes, it is argued that dynamic effects might be incorporated into the definition of the maximum capillary pressure. However, it is recognized that the exact correlation of the model used in this study to real heat pipes is sketchy. The results merely show that for apparent contact angles less than $10°$ that dynamic effects can be significant when the capillary number is on the order of 10^{-5}. The experiments described below will help draw a better correlation.

4   Facility to measure the capillary pressure

A facility has been constructed and assembled to afford the opportunity to actually measure the capillary pressure exsiting across a curved liquid-vapor interface. A schematic of the facility is shown in Figure 6. A vertical capillary tube is mounted within an evacuated pressure vessel. A $CO2$ laser beam, directed upon the interface, is to be used to affect evaporation. The power of the CO2 beam is variable. A large reservoir is to supply liquid to the capillary tube. Another capillary tube will be used to connect the reservoir and the
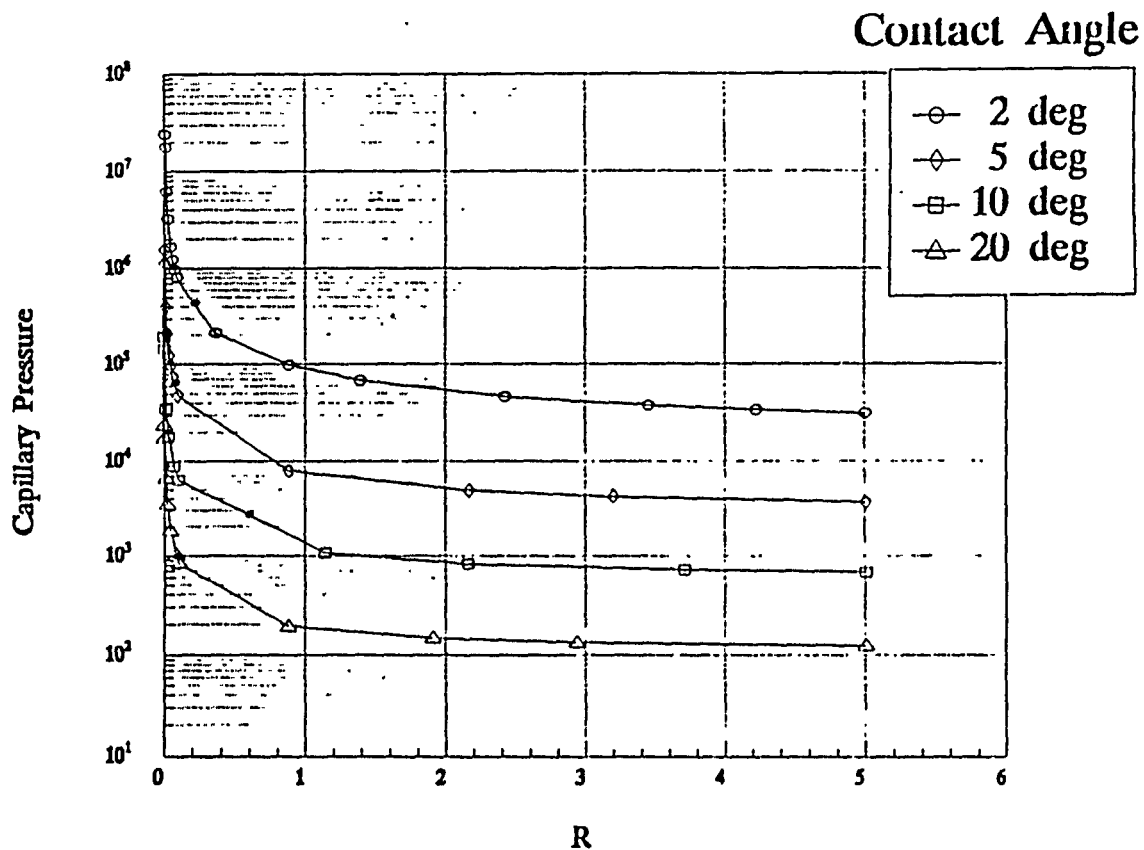
Figure 5: First order correction to the dimensionless capillary pressure dependence upon the distance from the interface.

evaporator, across which, the change in pressure will be measured. A cathetometer is to be used to measure the difference in levels between the reservoir and the meniscus within the capillary tube. The vapor pressure above the reservoir and within the evaporator are to be measured. Thus, in stable evaporation, the pressure jump across the meniscus can be compared to the pressure losses between the reservoir and the evaporator. The capillary pressure can then be inferred from the measurements. The pressure drop associated with the flow will be determined from the knowledge of the flow rate and the assumption of fully-developed conditions within the interface. Mathematically, this is written as:

$$P_{cap} = P_{\text{c}} - P_{\text{r}} + \Delta P_{\text{horizontal tube}} + \Delta P_{\text{vertical tube}} + \rho_l g \Delta h. \qquad (41)$$

In this equation P_{cap} is the capillary pressure existing across the evaporating interface, P_c is the vapor pressure within the chamber housing the vertical capillary tube, P_r is the vapor pressure above the reservoir supplying liquid to the capillary tube, $\Delta P_{\text{horizontal tube}}$ and $\Delta P_{\text{vertical tube}}$ are the flow pressure drops, and $\rho_l g \Delta h$ is the head difference between the reservoir and the meniscus.

These experiments will be performed for a variety of evaporation rates, including,
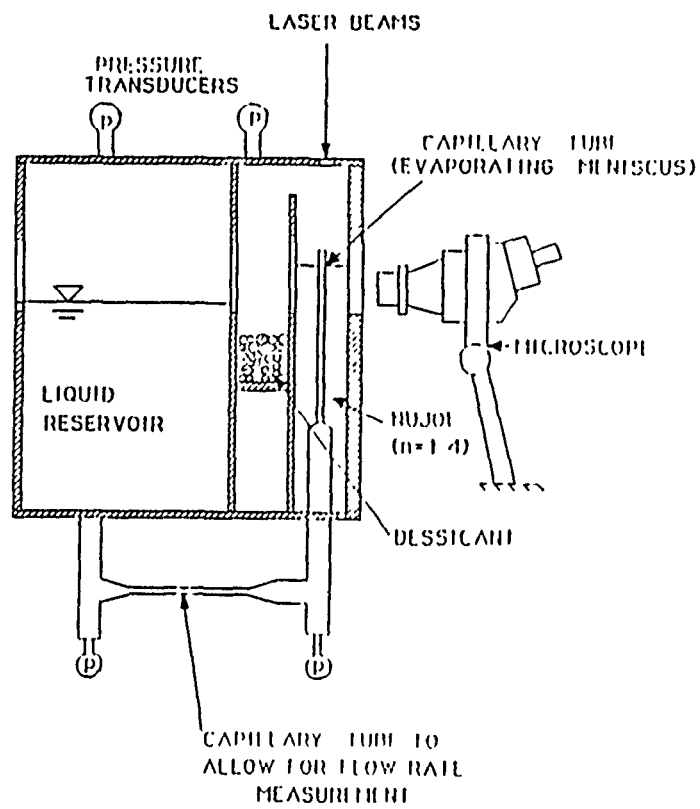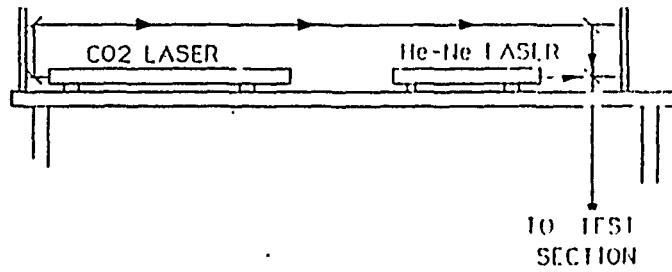
Figure 6: Schematic of experimental facility to study effects of fluid dynamics on capillary pressure.

situations where boiling is occuring. A microscope will be used to observe the evaporating meniscus to additionally observe how the shape of the interface changes with increasing evaporation rates. It is noted that these experiments are similar to those performed by Wayner [10]. However, a significant difference exists. Wayner did not measure the dynamic capillary pressure indirectly as is proposed here. Instead, he photographed the meniscus at varying evaporation rates. From these photographs he inferred the dynamic curvature, K_{evap}. The dynamic capillary pressure was then related to this dynamic curvature using Laplace's equation (which is valid only for static conditions). This reduction of the data was incorrect, as shown by Equation (3). The normal viscous stresses ultimately are responsible for the change in the capillary pressure.

The facility is now complete and ready for testing. The mini-grant is a necessity to allow for the opportunity to complete the experiments.

5 Recommendations

The work described in this report represents only the beginning of a total research effort deemed necessary to gain a clear understanding of the driving capillary potential in heat pipes. A methodical approach is suggested beginning with simple idealized problems and gradually approaching physical geometries more realistic of actual heat pipe wicks. First, support is essential for the completion of the experiments with the existing facility, looking at evaporation from a single capillary tube. Next, the vertical capillary tube should be replaced by a larger diameter tube. Real wick materials will be inserted into the tube. The CO_2 beam will be expanded to fall nearly uniformly over the top of the porous material. The dynamic capillary pressure will once again be inferred from the measurements. These results would definitely constitute a significant contribution to heat pipe designers, in that for real heat pipes, for evaporation rates typical of heat pipes, the dynamic effects on the capillary pressure will be known.

Of course, additional work is also needed to investigate the condensing problem at a curved interface to assess whether the flooded assumption typically used in heat pipe condensers is valid. A similar project can be performed to determine the influence of condensation of the capillary pressure existing across the liquid-vapor interfaces within the heat pipe condenser wick.

A number of additional studies are necessary to fully understand the phenomenon of

capillarity within heat pipe wicks. These and other questions must be answered. How is the driving capillary pressure effected by the occurrence of nucleate boiling within the evaporator wick? If bubbles are generated within the evaporator wick, is it possible for these bubbles to migrate from the wick? What is the transient behavior of the driving capillary pressure during start-up conditions? These questions must eventually be addressed to overcome the problems currently inherent in the use of heat pipes and capillary pumped loops.

References

[1] Chang, W. S. and Hagar, B. G., ASD90 1490 Technical Report, 1990.

[2] Gottschlich, J. M., 1989 SAE Aerospace Tech. Conf., Anaheim, CA.

[3] Dunn, P. and Reay, D. A., 1982, **Heat Pipes**, 3 ed., Pergamon Press, New York, New York.

[4] Chi, S. W., **Heat Pipe Theory and Practice**, MacGraw-Hill, 1967.

[5] Brennan, P. J., **B and K Heat Pipe Design Handbook**, B and K Engineering, Inc., Towson, MD, June 1979.

[6] Prosperetti, A., 1979, **Meccanica**, p. 34.

[7] Huh, C. and Mason, S. G., 1977, **J. Fluid Mech.**, 81, p. 401.

[8] Cox, R., Γ., 1986, **J. Fluid Mech.**,168, p. 169.

[9] Derjaguin, B.V., 1955, **Colloid J. USSR**, 17, p. 207.

[10] Wayner, P. C., Jr., 1978, **J. Heat Transfer**, 100, p. 100.

[11] Potash, M., Jr. and Wayner, P. C., Jr., **Int. J. Heat Mass Transfer**, 15, p. 1851.

[12] Kafka, F. Y. and Dussan, V., E. B., 1979, **J. Fluid Mech.**, 95, p. 539.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

<u>FINAL REPORT</u>

<u>Investigation of the Combustion Characteristics of a Confined Coannular Jet with a Sudden Expansion</u>

Prepared by: Paul O. Hedman, Ph.D., P.E.

Academic Rank: Professor

Department and   Chemical Engineering

University: Brigham Young University
                      Provo, Utah 64602

Research Location: Aero Propulsion Laboratory
                      Wright-Patterson AFB, OH 45433

USAF Researcher: W. M. Roquemore, Ph.D.

Date: 20 July 1990

Contract No: F49620-88-C-0053

Investigation of the Combustion Characteristics of a Confined Coannular Jet with a Sudden Expansion

by

Paul O. Hedman, Ph.D., P.E.
Chemical Engineering Department
Brigham Young University
Provo, Utah 84602

ABSTRACT

This report contains a brief summary of the work done to investigate the operation al characteristics of a burner that was designed to "specifically reproduce recirculation patterns and LBO processes that occur in a real gas turbine combustor." Measurements of lean blowout limit were conducted. The lean blowout limit was found to closely correspond to the lean flamability limit found in a well stirred reactor. LDA measurements of the gas velocity were made to determine the pattern of the complex flows in the combustor, and to identify regions of recirculation flame flow patterns. Measurements were made using spontaneous OH uv emission to characterize the fuel equivalence ratios where the flame was attached to the bluff step just outside the air jet and where a fully lifted flame occurred. At very fuel rich fuel equivalence ratios, the flame was attached to a greater or lesser degree. As the fuel equivalence ratio approached about 1.06, there was a point of demarcation where the flame became full lifted. Two-dimensional images of the flame and flow structure were taken with laser sheet lighting and a CCD camera. These images have shown the complex nature of the shear layers between the outer recirculation zone, the air jet, and the fuel jet. The eddies in these shear layers are of about the same scale as the annular space between the edge of the air jet and the fuel tube.

ACKNOWLEDGEMENTS

I.   INTRODUCTION:

This report presents a brief summary of results of an investigation to determine the flow and flame characteristics of a burner with a confined, coannular jet with a sudden expansion (Pratt and Whitney Task 100 Combustor).which has been developed to study the phenomenon of lean blowout (LBO) in modern annular aircraft gas turbine combustors. The combustor has been carefully designed (Sturgess, et al. 1990) to "specifically reproduce recirculation patterns and LBO processes that occur in a real gas turbine combustor."

The combustor consists of coaxial jets with a 29 mm diameter central fuel jet surrounded by a 40 mm diameter annular air jet. The jets are located in the center of a 150 mm diameter duct. A sudden expansion, rearward facing bluff body, with a step height of 55 mm, is located at the exit plane of the coaxial jets. The combustor test section incorporates flat quartz windows to accommodate laser and other optical access, but uses a metal shell with metal corner fillets to reduce the vorticity concentration and eliminate its effect of the bulk flowfield in the combustor. This box-section combustor with corner fillets allows reasonable optical access, while providing a cross section that approximates a two-dimensional axisymmetric cross section. The bluff body provides a recirculation region that can stabilize the flame.

When operated in a very fuel rich mode, the flame is very stable and is attached to the bluff body near the outer edge of the air tube. As the fuel equivalence ratio is reduced, the flame becomes less stable, and eventually reaches a point where it lifts from the base region, and becomes stabilized on the outer recirculation zone. Thus, there are two distinct operating modes for the burner, a fully attached flame, and a lifted flame. As the fuel equivalence ratio is further reduced, the flame becomes very unstable, and eventually completely blows out of the burner. This lean blowout limit has been found to correspond to a fuel equivalence ratio of about 0.49, very close to the lean flammability limit of the propane fuel and air (ca ϕ = 0.50, Lewis and von Elbe, 1987). Other studies (Sturgess, et al. 1990; Longwell, et al., 1953) suggest that the combustor is behaving like a well stirred reactor when it is near its lean blowout limit.

The objective of the project was to determine the combustion and flow characteristics of the burner over a range of operating conditions. Specifically, the study was to examine the mechanisms that contribute to the transition between a flame attached to the bluff body of the sudden expansion, and a lifted flame. These characteristics relate to a flame blowout modeling study being conducted by other investigators from Pratt and Whitney (Sturgess, et al. 1990). A secondary purpose of this study has been to collect data that could be used in validating a computer code that predicts the flame phenomena and blowout limits (Sturgess, et al. 1990).

The study has developed a better understanding of the overall behavior of the combustor, has investigated those phenomena that contribute to the transition from a flame that is well attached to a lifted

flame that is being stabilized on the outside recirculation zone, and has examined how the variation in fuel equivalence ratio affects the operation of the burner and the lean blowout.

II.   APPROACH:

The approach used in this study was to investigate the flow and flame structure within the burner over a wide range of fuel equivalence ratios. The characterization involved measurement of the fuel equivalence ratio at lean blowout, LDA measurements of the gas velocity throughout the combustor with and without combustion, laser sheet flow visualization of the combustion flows using $TiCl_4$ seed and uv laser sheet lighting to stimulate OH LIF emission, and spontaneous OH emission measurements in the uv to determine quantitative measurements of the fuel equivalence ratio where the flame transitions from an attached flame to a lifted flame.

III.   RESULTS:

Lean Blowout Limit Measurements. The lean blowout limit (i.e. the fuel equivalence ratio where the burner extinguishes) was determined by setting the air flow in the combustor, and then by slowly reducing the fuel flow rate (gaseous propane) until the reactor would extinguish. The fuel flow at blowout was somewhat dependent on the time temperature history of the combustor. The reactor would sustain a flame to somewhat lower fuel equivalence ratios after the reactor had reached thermal equilibrium. Consequently, the data reported were taken after sufficient time to allow thermal equilibrium to be reached. Undoubtedly, this thermal effect has lead to some of the data scatter and uncertainty in the data. The Pratt and Whitney Task 100 Combustor has been designed to accommodate several different exhaust extensions and exit restrictions. In the configuration tested in this study, the 10 inch exhaust extension was used with and without a 45% exit restriction.

Lean blowout experimental results were measured over a range of operating conditions and reactor configurations. Results for two combustors (SRL Building 450 combustor, and UDRI Building 490 combustor) with measurements taken at the two different facilities were compared Both combustors were tested with the 10 inch exhaust extension, but other hardware variations included tests with or without the 45% exhaust restriction, and tests with different arrangements of metal wall plates or quartz windows installed. Example results are presented in Figure 1. These results show the effect of the 45% exhaust restriction on the fuel equivalence at blowout. Quadratic least squares fits of the data are shown. There is a slight decrease (3.4% at 1500 slpm air to 5.5% at 4000 slpm air, an average effect over all air flow rates of about 3 9%) in the fuel equivalence ratio at blowout when the combustor is operated with the 45% exhaust restriction While the trend seems to be quite consistent over the range of air flowrates tested, the differences are generally within the experimental band of uncertainty, and the conclusion that there is a significant effect of the 45% exhaust restriction must be used with caution.
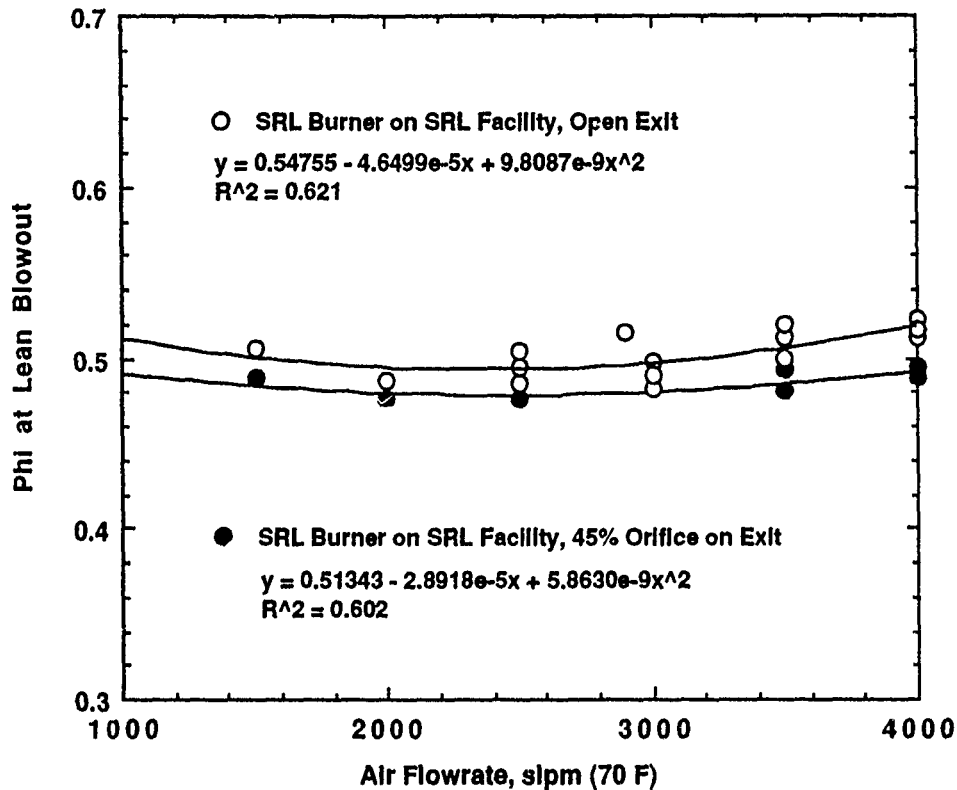
**Figure 1 - Effect of Exit Restriction (45% Orifice)
on Lean Blowout Limit**

Demarcation of Lifted Flame Region by Spontaneous OH uv Emission Measurements.

Photographs of the flame clearly illustrate the different attached and detached flame operational regimes of the burner. However, visual observation and conventional photographic techniques were unable to quantify the precise fuel equivalence ratio where the flame transitioned from an attached flame to a detached flame. Measurements of OH emission from the flame in the region of the attachment point were made to determine change in flame structure as a function of fuel equivalence ratio. These measurements have clearly identified the two operational regions of the flame, and have quantified the point of demarcation between the regimes of attached and detached flames. FFT analysis of the OH emission data have also provided an indication of the frequencies of the fluctuations in the flame associated with the attachment phenomena. LDA measurements of gas velocity in the flame at fuel equivalence ratios corresponding to a fully detached flame, to the point of demarcation between the detached flame and the attached flame, and to a well attached flame have been made, and used to

quantify the characteristics of the major recirculation zones, and the mixing shear layers. The influence of fuel equivalence ratio on the flame structure has been determined.

The burner operates in either a mode where the flame is attached to the bluff body just outside of the air annular jet, or in a lifted flame mode. Visual observations of the flame as fuel equivalence ratio was changed indicated that there was a point of demarcation in ϕ between these two regions where the flame was always lifted. At fuel equivalence ratios in excess of this value, the flame seemed to be lifted part of the time, and attached part of the time. Example results from these emission experiments are shown in Figure 2. In an effort to quantify the exact fuel equivalence ratio where this occurred, an optical system was set up that would look at a region of the flame near the attachment point,
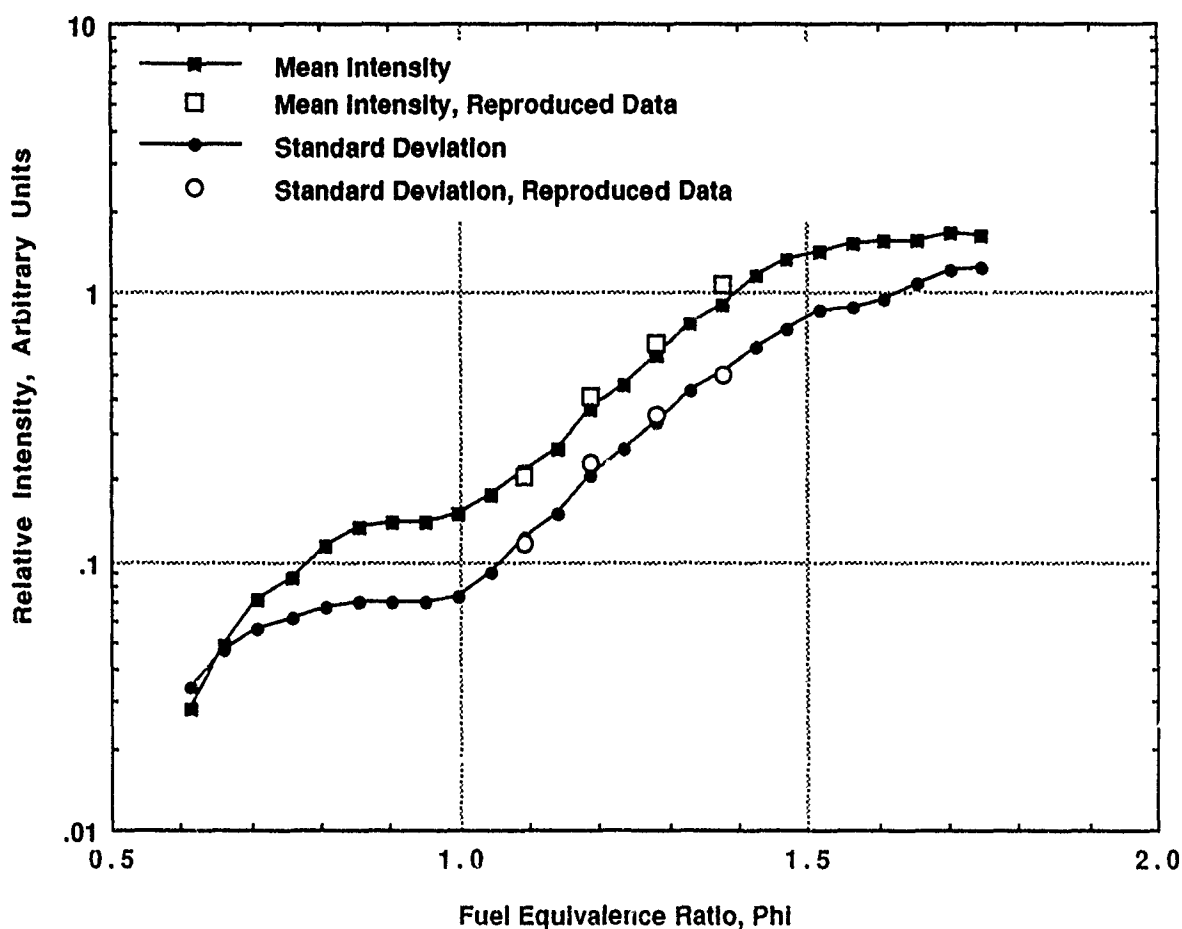
Figure 2 - Flame Attachment Data, 1000 slpm (70 F) Air flow, 45% Orifice on Exit (Calibrated Phi)

and measure the spontaneous OH emission in the near uv at about 308 nm. The optical system was a line of sight system that had a very narrow view angle, that was focussed through the flame attachment region,

but nevertheless, integrated the emission signal completely across the duct. Figure 2 illustrates clearly, that there are two distinct regions of operation, and that there is a definite point of demarcation between these regions. Similar measurements at other air flow rates have shown that this point of demarcation occurs at a fuel equivalence ratio of about 1.06 independent of the air flow rate.

LDA Velocity Measurements. A laser Doppler anemometer (LDA) was used to make extensive measurements of gas velocity in the burner. Measurements were made in cold flow simulations of the flame, and in flames at two fuel equivalence ratios ($\phi = 1.05$ and $\phi = 1.56$).

The air flowrate used was 1000 slpm (70 F), and the fuel (propane) flowrates used were 41 slpm (0 C) and 63 slpm (0 C). The cold flow simulations of the flame also used 1000 slpm (70 F), but used either air, nitrogen, or CO_2 in the fuel tube. The cold flow measurements were primarily completed to resolve operational difficulties in the burner, and in the LDA system.

Higher flowrates were desired in the burner, but heatload limitations of the exhaust hood prevented the continuous operation needed for LDA measurements to about 70 slpm (0 C) of propane. Consequently, a set air flowrate of 1000 slpm (70 F) was established for the experiments, and fuel flow rates were selected at the point of demarcation between a lifted flame [fuel flowrate of 41 slpm (0 C) propane, $\phi = 1.05$] and a well attached flame [fuel flowrate of 63 slpm (0 C) propane, $\phi = 1.56$]. Other fuel flow rates were used for some of the imaging measurements, notably fuel flowrate = 30 slpm (0 C) ($\phi = 0.78$) and 52 slpm (0 C) ($\phi = 1.31$) which correspond to a well lifted flame, and a moderately attached flame respectively. It would be desirous to have complete sets of data at all four fuel equivalence ratios considered, but the amour of data that could be taken during the short time of the study was limited to the measurements at $\phi = 1.05$ and $\phi = 1.56$.

LDA measurements were taken in two different zones in the reactor. Due to the limitations in optical access, and the orientation of the laser beams for the two component LDA, axial and radial velocity components could be obtained from about -31 mm to +31 mm in the radial X direction. In the other coordinate direction, axial and tangential velocity data could be obtained from about -70 mm to +70 mm in the Y coordinate radial direction. Normally, radial velocity profiles were only taken on one side of the duct, from the centerline out. Complete profiles were usually taken for the data near the let exit to determine the boundary condition to the reactor, and to determine the symmetry of the inlet flow conditions. Radial velocity profiles in either quadrant were take from about 6 mm axial location (closest position possible to the bluff body sudden step) to 300 mm axial location. Most of the interesting flame structure existed between these two axial locations. The axial and radial velocity components are the much preferred components to evaluate the flows in the axisymmetric flowfield.

The velocity measurements permit the flow structure in the burner to be quantified. One example set of data is shown in Figure 3. This figure presents data from the near field of the burner (Z = 6 mm) out to a 200 mm axial location. In the near field of the burner, the high velocity associated with the flow through the air annulus is clearly evident. A more complete profile taken at 1 mm steps in the other coordinate direction shows that the velocity across the air duct is relatively flat at about 34 m/s. from about 14.5 mm to 20 mm radius. The course grid of this data set shows only the general details of this region. The velocity profiles in the fuel tube clearly show that there is a recirculation bubble in the fuel tube that extends from about 10 to about 32 mm. The exact size of this recirculation zone depends on the fuel flow
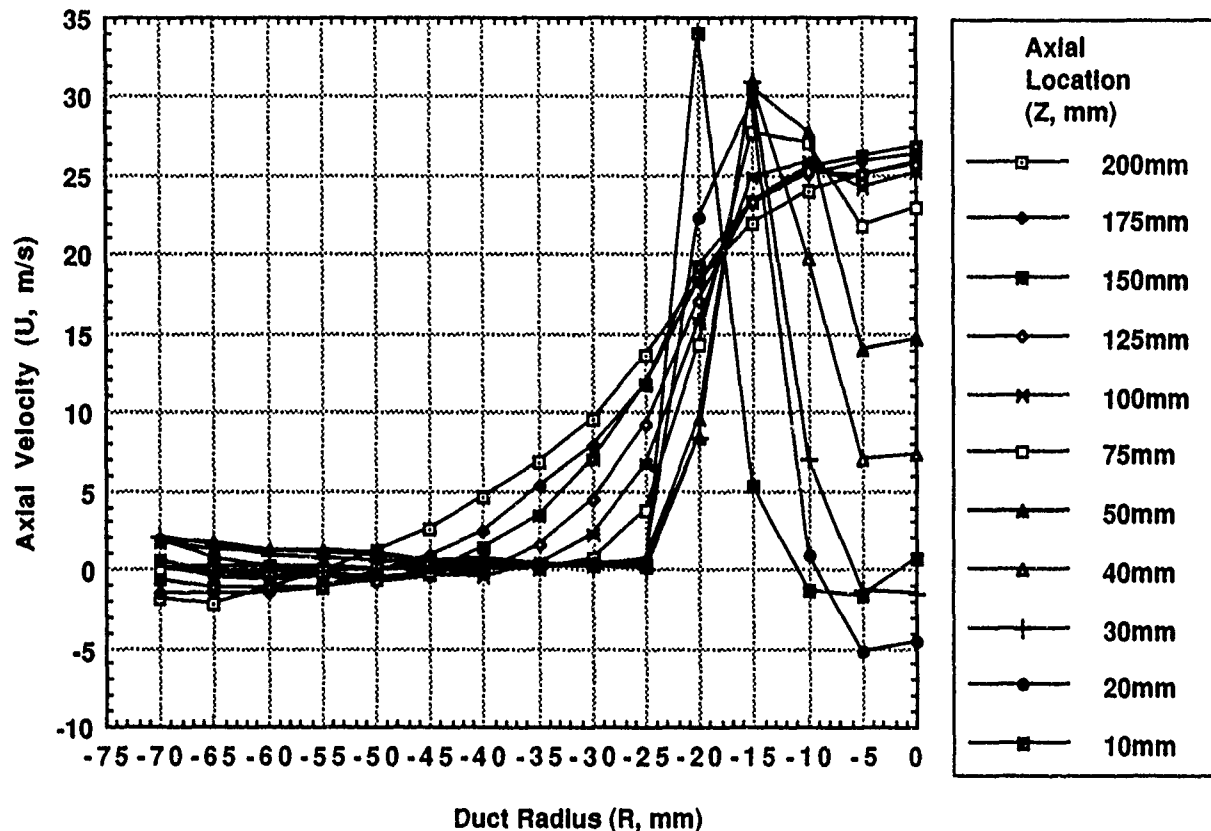
Figure 3 - LDA Axial Velocity Measurements in the
Pratt and Whitney Task 100 Combustor
1000 slpm Air, 41 slpm Propane, Phi = 1.05

rate The change in velocity profiles with increased axial distance is clearly evident in Figure 3. It can be seen that the flow in the channel is flattening, and is expected to approach a typical turbulent flow profile by the reactor exit Preliminary analysis of the velocity data obtained shows that there is a very complex recirculation zone associated with the bluff body sudden step.

<u>Laser Sheet Lighting Imaging Measurements.</u> Two different schemes were used to examine the flow and combustion details of the flame, $TiCl_4$ seeding and LIF emission of OH. In each scheme, a sheet of planar laser light was used to capture an instantaneous (10 ns) two dimensional image of the flame and flow structure. In each case, the plane of the laser light bisected the test section of the reactor, passing through the centerline of the reactor. The laser light came through the window on the X coordinate. The image of the illuminated field was taken by a CCD camera located on the Y coordinate. Two different CCD cameras were used, a 1024x1024 array unintensified camera for the more intense $TiCl_4$ seeded experiments, and a 384x576 intensified CCD camera for the less intense images of the OH LIF. The CCD cameras were able to capture the full width of the optical window. The camera used with the $TiCl_4$ seeded experiments was able to capture a vertical image in the quartz window about 145 mm in height. The camera used in the OH LIF imaging was able to capture an image about 79 mm in height. In order to get a more complete record of the OH images, the burner was located a four different axial locations, 0 mm, 75 mm, 150 mm, and 225 mm.

In the first technique, a sheet of frequency doubled Nd/Yag laser light (532nm) was focused through the test chamber. Mie scattering of this laser light from particles of TiO_2 was used to obtain instantaneous images of the flame and flow patterns. An example of one of these images is shown in Figure 4. The images were recorded on a 1024x1024 array CCD camera. The TiO_2 seed was obtained

Figure 4 - Example of Laser Sheet Lighting of $TiCl_4$ Seeded Flame, 1000 slpm Air, 30 slpm Propane, Seed in Air
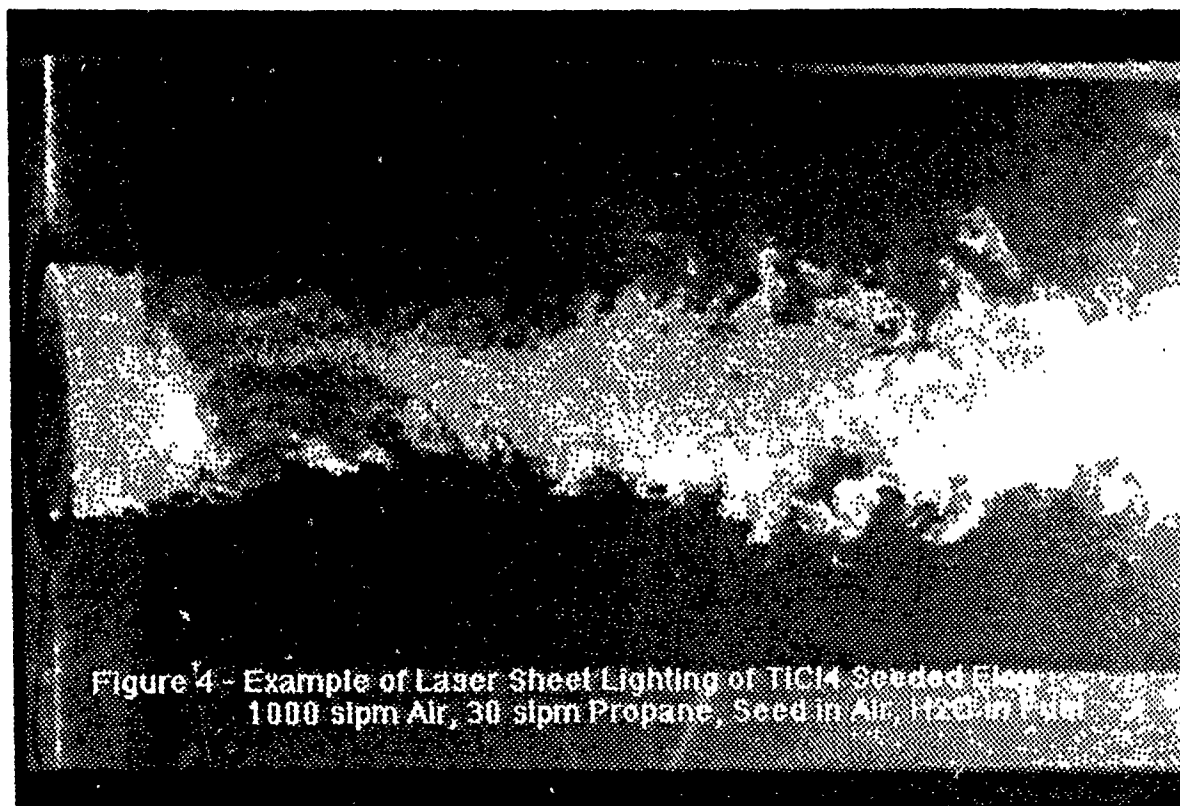
from the near instantaneous reaction of $TiCl_4$ vapor and moisture to form the $TiO2$ particles, and HCl vapor. The $TiCl_4$ vapor was seeded with either the dry air flow or the propane fuel. Moisture was added with the opposite flowing stream to provide ample $H2O$ for the chemical reaction. Moisture generated in the combustion process also participated in the chemical reaction. This figure shows the results of seeding the $TiCl_4$ vapor into the air stream, and adding moisture to the fuel tube. The flow structure associated with the flow through the fuel tube are evident. The eddies in the shear layer between the fuel jet and the annular air jet are quite evident. The turbulent structure downstream is also clearly evident.

In the second imaging technique, the laser light from a tunable dye laser was used to form a sheet of laser light. The frequency used was about 283 nm. This frequency was used to pump a fluorescence transition of OH which resulted in a LIF signal at about 308 nm. Two dimensional images of the flame zone shown by the OH LIF were obtained with a 384x576 intensified CCD camera. These images have been used to examine the location and shape of the flame zones in the burner. An example of one of these images in shown in Figure 5. This image is located in the near region of the fuel and air jets, and clearly shows the combustion zones associated with the flame. The convoluted shape of the flame front clearly shows the complex structure of the flame and the associated eddies.

Figure 5 - Example of Laser Sheet Image of OH LIF
1000 slpm Air, 63 slpm Propane
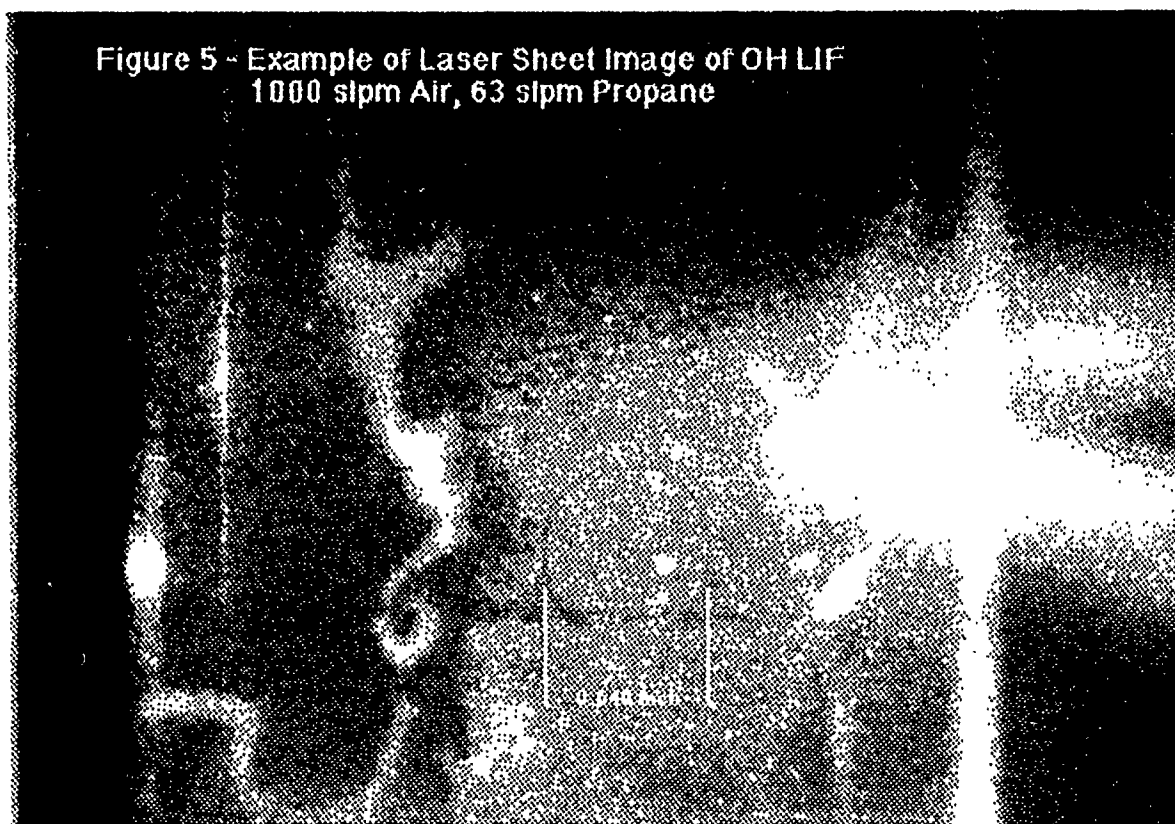
IV.   DISCUSSION OF RESULTS (CONCLUSIONS):

Lean Blowout Limit Measurements. The following preliminary conclusions have been drawn from the results of the lean blowout limit measurements.

1. In general, the fuel equivalence ratio at lean blowout is independent of air flowrate. Some of the data, particularly the UDRI data suggest a slight decline in ϕ with increasing air flowrate, while other sets of data, particularly the SRL data shows a slight increase in ϕ with increasing air flowrate. The noted slight increase or decrease are well within the data scatter and measurement uncertainty, and are probably not meaningful.

2. The fuel flow meters are only readable to 1 slpm of propane at the SRL facility, and 1 slpm of air (times 0.36 to convert to propane) at the UDRI facility. This translates to about 1 part in 20 to about 1 part in 80 over the range of propane flow rates being used, or about 5% at 1000 slpm air flowrate down to about 1.3% at 4000 slpm air flowrate. Trying to determine low percentage trends within these error bands is risky.

3. There is a slight effect observed in the performance of the burner with and without the 45% exhaust restriction installed. The 45% exhaust restriction lowers the fuel equivalence ratio at lean burnout by an average of about 3.9% over the range of conditions tested. The effect seems consistent, even though it is of the same order as the scatter and experimental uncertainty in the data.

4. There was no significant effect observed on the effect of quartz windows or metal plates installed in the reactor. Limited data taken on the SRL facility showed a small effect, but a more exhaustive set of data taken on the UDRI facility over a much wider range of test conditions showed the the effect was insignificant.

5. There was a slight difference in the performance of the SRL and UDRI burners, that could not be explained. In general, the fuel equivalence ratio at lean blowout for the SRL burner was always higher than that observed for the UDRI burner. The differences were slight, generally on the same order as the data scatter and measurement uncertainty, but consistently in the same direction on both the SRL and UDRI test facilities. There was no conclusive explanation for the observed differences.

Demarcation of Lifted Flame Region by Spontaneous OH uv Emission Measurements. The following preliminary conclusions have been drawn from the results of the measurements of spontaneous uv emission at about 308 nm.

1. The OH uv emission measurements showed a strong change in relative signal intensity as the fuel equivalence ratio was changed from a very fuel rich case where the flame was clearly attached, to a very fuel lean case where the flame was very clearly lifted.

2. There existed a marked inflection in the intensity of the OH emission curve that appeared to separate the regions of the flame where the flame was attached and lifted. This point of demarcation appeared to mark the location where the flame was fully lifted.

3. The point of demarcation was very insensitive to air flow rate. The value corresponded to a fuel equivalence ratio of about 1.06 to 1.08 over the entire range of air flows tested.

4. The point of demarcation was used to set the fuel equivalence ratios used for other experiments. Notably, fuel equivalence ratios of about 0.7, 1.06, 1.3, and 1.6 were established to denote a well lifted flame, the point of demarcation, an attached flame, and a well attached flame respectively.

LDA Velocity Measurements. The following preliminary conclusions have been drawn from the results of the LDA velocity measurements.

1. A small recirculation zone exists in the primary fuel tube. This recirculation zone preferentially attaches to one side of the fuel tube, which cause a regular asymmetry in the velocity profiles in the near field region. This asymmetry was also seen in cold flow experiments where CO_2 was used to simulate the fuel flowrate. When air or nitrogen was used in the fuel tube, or when the air in the annular air jet was not flowing, the velocity profiles in the fuel tube were very symmetric.

2. Limited analysis of the velocity in the step region of the combustor has shown the flow to be very complex. There is a suggestion from the data that there may be multiple vortices in the corner formed by the side wall and the bluff body.

3. At times, it was difficult to make LDA measurements in the facility because of the unreliability of the signal processors, and the difficulties in uniformly seeding the flow with the Al_2O_3 seed particles.

4. Velocity measurements at higher air and fuel flow rates, although highly desirable to fully understand the operation of the burner, were not possible because of the heat load limitations of the exhaust hood.

Laser Sheet Lighting Imaging Measurements: The following preliminary conclusions have been drawn from the results of the laser sheet images taken.

1 The two dimensional images show a complex structure of the shear between the air jet and the fuel jet and between the air jet and the outer recirculation zone.

2. The eddy scale in the shear flows appears to be about the same scale as the annular gap between the fuel tube, and the air tube.

3. The OH images in particular have shown that the flame is rather fragmented. There does not seem to be a continuous flame sheet. There do appear to be short regions of flame sheet, but in general they seem detached from other segments of the flame.

V.   RECOMMENDATIONS:

Lean Blowout Limit Measurements. The following recommendations might be considered, if additional, more precise lean blowout data are required.

1. The flow meters, particularly the fuel flow meter needs to be replaced with a meter of greater precision. The meter needs to be very carefully calibrated, and should be able to be read to a precision of 0.1 slpm.

2. The fuel flow capacity in the SRL test facility needs to be increased. At present the propane flow rate is limited to 108 slpm if a single flow meter/controller is used. If two meters/controllers are used in parallel, the flow rate can be increased up to about 124 slpm propane.

3. The hood in the SRL test facility needs to have a greater capacity.

LDA Velocity Measurements. The following recommendations might be considered, if additional, more precise lean blowout data are required.

1. Replace the exhaust hood with one of higher thermal capacity (A new hood is currently in the process of being fabricated, and will be installed shortly.)

2. Increase the flow rate potential of the propane fuel system so that measurements can be made at higher flowrate conditions.

3. Modify the LDA processors to improve their reliability.

4. Incorporate new particle seeders of improved reliability and control.

5. Conduct additional experiments at other fuel equivalence ratios and air flowrates.

Laser Sheet Lighting Imaging Measurements: The following recommendations might be considered when additional laser sheet lighted images are taken in the burner.

1. There is considerable evidence to suggest that the burner flow and flame characteristics have a considerable three-dimensional character. Consequently, it is suggested that some off axis images be taken. Illuminating the test section diagonally through the corner windows (i.e. at 45 degrees), with the CCD camera remaining on the orthogonal coordinate may allow images to be obtained in or at least near to the corner metal supports.

Suggestions for Follow-on Research. There is considerable work yet needed to fully understand the operation of this burner. There is a suggestion that the burner has many three-dimensional features the need to be quantified. Thus, additional off axis two-dimensional imaging measurements, and off axis LDA velocity measurements are needed. LDA velocity on the diagonal through the corner windows could also give insight to the flows the exist behind the metal corners. Measurements of gas temperature using a CARS system are also needed.

The measurements during this period were limited to 1000 slpm air, and about 70 slpm of propane because of the heat limitations of the exhaust hood. A hood of larger capacity needs to be installed, so that higher flowrates can be accommodated. A new exhaust hood has been fabricated, and will be installed by the end of summer.

Application for a Mini Research Initiation Grant. An application for a mini research grant will be prepared as a result of the summer fellow research program. The proposed project would include the installation at the Brigham Young University of a burner identical to the one tested in this summer project at the Aero Propulsion Laboratory, Wright Patterson AFB during this summer fellowship project. The burner would have been modified to incorporate a swirling burner supplied by Pratt and Whitney Aircraft Co, East Hartford, Connecticut. This proposed study would allow testing of swirl components in a single near axisymmetric burner. The Aero Propulsion Laboratory, Wright Patterson AFB has a companion programs, the Task 100 burner program, and the Task 200 burner program. The Task 100 burner was the burner investigated during this summer program. The Task 200 burner incorporates four swirling burners in a linear array, and is designed to burn liquid fuels. The proposed project would use the Task 100 combustor, and a single Task 200 swirling injector. The proposed BYU project would fill a need to investigate the swirling burner in a single burner configuration. In addition, to basic combustion measurements, the proposed program at BYU would begin to investigate the formation of NO_x pollutants at the higher fuel equivalence ratios associated with higher performance gas turbines. The proposed mini grant would involve the installation and checkout of the burner components that would be supplied by The Aero Propulsion Laboratory, Wright Patterson AFB

REFERENCES

Lewis, B., and G. von Elbe, (1987), Combustion, Flames and Explosions of Gases, 3rd Edition, Academic Press, Inc.

Longwell, J.P., E.E. Frost, and M.A. Weiss, (August 1953), "Flame Stability in Bluff Body Recirculation Zones", Industrial and Engineering Chemistry, 45(8), pp 1629-1633

Longwell, J.P., and M.A. Weiss, (August 1955), "High Temperature Reaction Rates in Hydrocarbon Combustion", Industrial and Engineering Chemistry, 47(8), pp 1634-1643

Sturgess, G.J., D.G. Sloan, A.L. Lesmerises, S.P. Henneghan and D.R. Ballal. (June 1990), "Design and Development of a Research Combustor for Lean Blowout Studies", 35th International Gas Turbine and Aeroengine Congress and Exposition. Brussels, Belgium.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the Universal Energy Systems, Inc.

<u>FINAL REPORT</u>

<u>AIRCRAFT HVDC POWER SYSTEM - STABILITY ANALYSIS</u>

Prepared by: K. Sankara Rao, Ph.D. and Allen Olheiser, B. S.

Academic Rank: Professor

Department and   Electrical and Electronics Engineering Department

University: North Dakota State University, Fargo, North Dakota

Research Location: WRDC/POOC1, Wright Patterson Air Force Base

USAF Researcher: Joe Weimer

Date: September 7, 1990

Contract Number: F49620-88-C-0053

AIRCRAFT POWER SYSTEMS - STABILITY

by

K. Sankara Rao
Alan Olheiser

ABSTRACT

Analysis and modeling of aircraft 270V dc electrical power systems are the main topic of the research project. HVDC at 270 volts has many advantages over the currently used three phase electrical power systems in an aircraft. There are some problems, particularly instability in the presence of a constant power load, which are addressed in this research. Computer models have been developed for the various components of the HVDC system and constant power load. The analysis using EMTP is included in this report.

Acknowledgment

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. Universal Energy Systems must be acknowledged for their help and concern in all administrative aspects of this program.

My experience was rewarding because of many influences. Joseph Weimer provided me with the necessary help, support, encouragement, and an excellent working environment for conducting the research.

1   INTRODUCTION:

Aircraft Electrical Power Systems, at present use three phase AC systems and research and development into the possible use of 270V dc as an alternative is underway.

Electrical Laboratory of Aero Power Propulsion Laboratory at Wright Patterson Air Force Base is very much concerned with computer modeling of HVDC systems. One of the main problems foreseen is instability when constant power loads are applied to the HVDC distribution system.

My research interests are in electrical power systems and modeling of electrical power systems. I have used the Electromagnetic Transients Program (EMTP) for modeling electrical power system components such as, alternators, transmission lines, rectifiers, inverters etc. and believe that EMTP is an excellent package for analyzing Aircraft Power Systems. My past experience in modeling three phase 400 Hz distribution systems in aircraft was a factor in my being assigned for this research.

2   OBJECTIVES OF THE RESEARCH EFFORT:

Electrical load on an aircraft power system usually consists of a combination of lighting and motor loads. In addition there is a considerable amount of dc load in an aircraft for flight critical computers etc. The main power supply in an aircraft at present is a three phase supply at 400 Hz. This frequency must remain constant

in the presence of variable speed turbines which supply mechanical power to the electrical generators. There are two methods generally used for obtaining constant frequency.

1. Constant Speed Constant Frequency Systems (CSCF): A Constant Speed Drive (CSD) is used to keep the generator running at a constant speed even while the input to the constant speed drive varies considerably. The advantage is that the electrical output from the alternator is at a constant frequency with very little harmonic content. The main disadvantage of this system is the very high maintenance required of the CSD.

2. Variable Speed Constant Frequency Systems (VSCF): In this system a variable speed alternator supplies a variable frequency three phase power. This output is rectified and inverted to produce constant frequency three phase power. The main advantage of this system is low maintenance cost. The major disadvantage is the presence of harmonics in the output due to the inverter operation.

As can be seen above, the three phase ac distribution system has some drawbacks and research and development is under way for using a 270V dc supply as the main electrical power supply. In this system, the inverter portion of the VSCF system can be eliminated and all of the dc loads can be directly applied to it. The low voltage dc loads at 28V can be supplied by using dc-dc converters. As far as motor loads are concerned, inverter fed brushless dc motors, inverter fed induction motors or inverter

fed switch reluctance motors can be used.

My assignment, as a participant in the 1990 Summer Faculty Research Program (SFRP), was to model a 270V dc power system on a computer and study the stability of the system. The Electromagnetic Transients Program (EMTP), which was developed in the early seventies for Bonneville Power Administration (BPA), is an excellent tool for analyzing transient behavior of a power system. This package, used in the present research, is constantly being revised and modified. It has been found that constant power loads cause instability in HVDC systems and use of a large capacitor and harmonic filters eliminate the instability and reduce the ripple current drawn from the power system.

A more detailed analysis of the HVDC system when motor loads are present should be undertaken. Since EMTP can be used to model various types of motors, it is possible to analyze the system under motor loads using EMTP. A proposal for an AFOSR Mini Grant for continuing this work at North Dakota State University is being submitted.

3   EMTP

The Electromagnetic Transients Program was developed in the early seventies by Dr. W. Scott Meyer of Bonneville Power Administration. This program is being constantly updated and has numerous users all over the world. The program is very

extensive and has all the features necessary for the analysis of Aircraft Power Systems. Some of the highlights of the program are as follows:

- The various kinds of elements that can be represented are:

  1. RLC branches

  2. Distributed lines

  3. Switches including diodes, thyristors, power transistors and time controlled switches.

  4. Sources

     (a) Voltage sources, both ac and dc

     (b) Current sources both ac and dc

     (c) Voltage and Current sources controlled by any other variable

     (d) Alternators with their full representation

     (e) DC generators and motors with their full representation

     (f) Induction motors with their full representation

     (g) Analytical sources

  5. Control System Blocks. These blocks can be linear or non linear. They can be represented by transfer function blocks for linear components. The inputs and outputs of the control system blocks can be interfaced with the voltages and currents of the electrical network.

- The output results are conveniently obtained in graphical form on a CRT or on any type of plotter or printer. In addition, there is a tabular output which provides highlights of all the necessary results.

4 CONSTANT POWER LOAD ON AN IDEAL HVDC SYSTEM

As a starting point for this research, an ideal 270V dc source is chosen as the power system. The series impedance with the ideal source is chosen as a small inductance in series with a parallel combination of an inductance and a resistance. The series inductance represents the subtransient reactance of the alternator in the actual HVDC generating system. The parallel combination of resistance and inductance represent the time constant and the open loop error. The rectifier filter is represented by a series RC circuit across the load. The constant power load is represented by a voltage controlled nonlinear current source such that the product of the current and the voltage equals the power. The constant power load is a cyclic load occurring 100 times a second with a duty cycle of 0.5. The circuit diagram is shown in Fig. 1.

The simulation results are shown in Fig. 2. Some of the main points of the results are the following.

1. There is considerable ringing in the voltage across the load.

2. The current drawn from the source has a very high ripple content with a fun-

damental frequency of 100 Hz.

An analysis of the circuit shows that the system is unstable for constant power loads. The system can be stabilized by including a large capacitor across the constant power load. The size of the capacitor depends on the magnitude of the load, the value of the inductance in series with the source and the resistance in the circuit. The simulation results, with a 3000 microfarad capacitor in parallel with the load, are given in Fig. 3. The ringing in the voltage is now absent but the ripple content of the source current is still high.

The ripple current can be reduced by inserting series LC circuits with resonant frequencies equal to 100 Hz and its odd harmonics, in parallel with the load. Fig. 4 shows the effect of inserting these harmonic filters.

5   SIMULATION OF A RECTIFIER WITH A CONSTANT POWER LOAD

The second phase of the research consisted of simulating an ideal three phase source rectified and feeding a constant power load cycling 100 times per second. The circuit diagram for this system is shown in Fig. 5 and the results are shown in Fig. 6. The following conclusions can be drawn from the results.

1. The constant power load causes instability and this instability can be eliminated by inserting a large capacitor in parallel with the load. The ripple content in

the load voltage is still high. The current drawn has a high ripple content of 100 Hz and its harmonics.

2. Introduction of harmonic filters as suggested in the earlier section reduces the ripple content of the voltage and the current drawn from the supply.

In the next phase of the simulation, the three phase voltage source magnitude is controlled so as to have the voltage across the load to be 270 volts. The overall control system is of type one.

The results are shown in Fig. 7 and show that the instability can be eliminated by inserting a large capacitor.

6 BRUSHLESS DC GENERATOR WITH A CONSTANT POWER LOAD

In the next phase of the research, the stability of a brushless DC generator in the presence of constant power load is studied. As in the earlier cases, the overall system is unstable under a constant power load. The instability is due to the fact that the constant power load appears as a negative resistance and the overall impedance as seen by the generator has a negative real part. Introduction of a large capacitor in parallel with the load makes the overall impedance to be stable and the overall system is stable. The value of the capacitor needed to make the system stable depends on the level of constant power that is applied as the load. The circuit diagram of the

system is shown in Fig. 11 and the results are shown in Fig. 8 and 9.

7 CONSTANT POWER LOAD

In all phases of the earlier research it was assumed that the load has a characteristic of demanding constant powe irrespective of the voltage across it. This load is simulated as a current source whose magnitude is such that the product of the source current and the voltage across it is the negative of the power demand.

In actual practice, the constant load, as it is assumed is a power conditioner whose output voltage is maintained constant irrespective of the input voltage. When this power conditioner is connected to a resistive load, which is constant the output power is constant. In this phase of the research a buck switching type of regulator is simulated and used to replace the current sources. The overall system stability is studied. The circuit diagram of the regulator is shown in Fig. 10 and the results are shown in Fig. 11.

An examination of the results shows that a. y constant power load, which is derived from a switching regulator does not pose any problems at all when the input voltage is derived from ideal sources. The output voltage of the regulator remains constant while the input voltage is varied within wide limits. The settling time of the transient response, when the load resistance is changed, is of the order of a tenth of a millisecond. This makes the circuit fully capable of operating a pulse load at 100 cycles per

second.

8 RECOMMENDATIONS

The Electromagnetic Transients Program (EMTP) is a very useful tool for analyzing aircraft power systems and will be used by this researcher in all simulation work. This simulation will be compared with the simulations using other software where possible.

The preliminary study shows that a constant power load, when viewed as a resistor with a hyperbolic v-i characteristic, poses a stability problem for dc systems. However, when the constant power load is considered as a constant resistive load on a power conditioner, the dc system seems to behave normally. Further simulation and analysis should be done to confirm these studies.

A proposal for mini-grant is being submitted to the AFOSR. The work proposed includes simulation and analysis of an HVDC system with various load configurations such as motor loads, actuator loads etc. The motor loads considered will be dc motors and induction motors connected through inverters using MCTs.
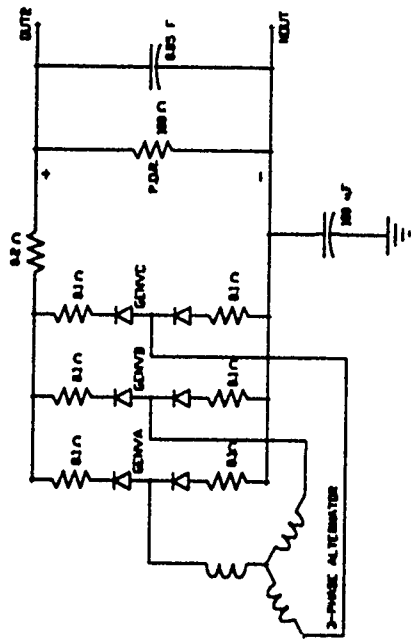
270V DC POWER SUPPLY
Fig. 5

BUCK SWITCHING REGULATOR
Fig. 10
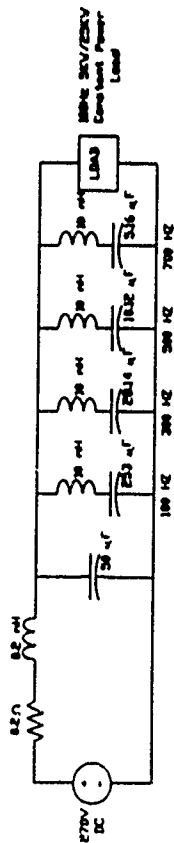
270V DC SOURCE
CONSTANT POWER LOAD AND HARMONIC FILTERS
Fig. 1
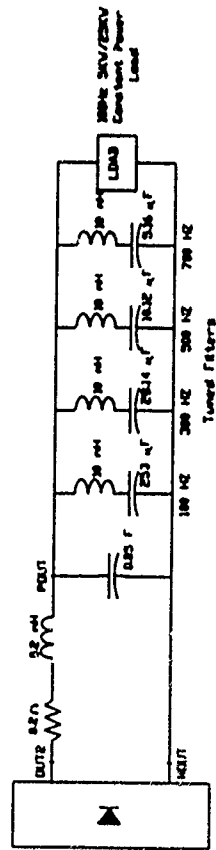
CONSTANT POWER LOAD WITH HARMONIC FILTERS
Fig. 8

Fig. 2b

Fig. 2d

Fig. 2a

Fig. 2c

82-14

270V DC SYSTEM
DCSOURCE WITH A LARGE CAPACITOR ACROSS THE LOAD

OC   56.04
RMS   67.89
RMS(RIPPLE) 38.32

CURRENT DRAWN FROM DC SUPPLY

DC2   08/22/90

Fig. 3b

270V DC SYSTEM
DCSOURCE WITH A LARGE CAPACITOR ACROSS THE LOAD

DC2   08/22/90

Fig. 3d

270V DC SYSTEM
DCSOURCE WITH A LARGE CAPACITOR ACROSS THE LOAD

OC   270.00
RMS   270.03
RMS(RIPPLE)   3.77

LOAD VOLTAGE

DC2   08/22/90

Fig. 3a

270V DC SYSTEM
DCSOURCE WITH A LARGE CAPACITOR ACROSS THE LOAD

DC2   08/22/90

Fig. 3c

82-15

Fig. 4b

Fig. 4d

Fig. 4a

Fig. 4c

Fig. 7a

Fig. 7c

Fig. 6

Fig. 7b

PHASE A GENERATOR VOLTAGE

Fig. 9c

PHASE A DIODE CURRENT

Fig. 9d

POWER SUPPLY CURRENT

Fig. 9a

FOURIER ANALYSIS OF THE POWER SUPPLY CURRENT

Fig. 9b

SYNCHRONOUS MACHINE, 100KW DC AND 25KW CONSTANT POWER LOAD   270V DC SYSTEM

FOURIER ANALYSIS OF THE LOAD VOLTAGE(100HZ,30WHZ....)

frequency

mag.

SYNCHMCH 08/30/90

Fig. 9g

SYNCHRONOUS MACHINE, 100KW DC AND 25KW CONSTANT POWER LOAD   270V DC SYSTEM

DC   268.91
RMS   269.05
RMS:RIPPLE)

seconds

LOAD VOLTAGE

VOLTS

SYNCHMCH 08/30/90

Fig. 9e

SYNCHRONOUS MACHINE, 100KW DC AND 25KW CONSTANT POWER LOAD   270V DC SYSTEM

DC   269.67
RMS   269.70
RMS:RIPPLE) 4.04

seconds

LOAD VOLTAGE

VOLTS

SYNCHMCH 08/30/90

Fig. 9f

82-19

Fig. 11c

Fig. 11d

Fig. 11a

Fig. 11b

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

DESIGN OF A DYNAMIC TEMPERATURE MEASUREMENT SYSTEM

FOR REACTING FLOWS

Prepared by: Larry A. Roe, Ph.D.

Academic Rank: Assistant Professor

Department and   Mechanical Engineering Department

  University: Virginia Polytechnic Institute and State University

Research Location: WRDC/POPT
                    Wright-Patterson AFB, OH   45433

USAF Researcher: Abdollah S. Nejad

Date: 31 July 1990

Contract No: F49620-88-C-0053

DESIGN OF A DYNAMIC TEMPERATURE MEASUREMENT SYSTEM

FOR REACTING FLOWS

by

Larry A. Roe

ABSTRACT

A system for the acquisition of spatially and temporally resolved temperature data in combustion systems was designed. This approach utilizes a dual-junction thermocouple probe for determination of instantaneous convective characteristics of the junction, necessary for compensation of the inherently low transient response. A data acquisition system for coincident recording of velocity with a laser Doppler velocimeter was configured; incorporating the probe, amplifiers, filters, analog-to-digital converters, and software. An extensive literature review was conducted, the system designed, requisite hardware specified and purchasing initiated. Recommendations for an experimental evaluation of the concept were presented.

ACKNOWLEDGMENTS

I express appreciation to all those in the Advanced Propulsion Division who contributed to making my stay both personally enjoyable and technically satisfying, and to the Air Force Systems Command, Air Force Office of Scientific Research, for providing sponsorship.

I. INTRODUCTION

The acquisition of spatially and temporally resolved temperature data in combustion systems is of vital importance for stability characterization and control, the informed interpretation of velocity data, and the evaluation of turbulence correlations and compressibility effects. An application of primary interest is the concurrent acquisition of temperature with velocity data acquired via laser Doppler velocimeter (LDV). The LDV provides a sampling of the fluid velocity by measuring the Doppler shift of laser light scattered from particles in the flow. Concurrent measurement of temperature at the same spatial location as the velocity sample is necessary to properly assess the influence of reaction and compressibility on turbulence, mean velocities, and the validity of the LDV measurements themselves.

Data rates must also be sufficient to track fluctuations in the flowfield, whether they be traditional turbulence, large-scale structures, or instabilities. This requires a method capable of providing thermal data at about 2 kHz for the anticipated application to research ramjet combustors. Non-intrusive temperature measuring techniques do not, at present, offer the combination of spatial and temporal resolution and high data rates necessary for correlation with LDV data. A probe measurement is therefore necessary, with thermocouples offering the best experience base from which to draw. While thermocouples have been in use for many years, their scientific application to reacting flows is still in the developmental stage.

Probe techniques pose survivability and system disturbance problems. Minimal disruption of the mechanical, thermal, and chemical character of the flow leads to the selection of small devices, which generally have reduced survivability relative to more robust designs. The required

compromise leads to the selection of the smallest probe which can sur-

vive the environment for a practical data-collection time. Even the

smallest physical probe can significantly perturb the flowfield, espe-

cially in recirculating and swirling flows, so the parameters measured

are not the same as would exist in the absence of the probe. If, how-

ever, the purpose of the investigation is to evaluate compressibility

effects, stress terms, LDV averaging techniques, or general temperature-

velocity correlations, properly configured thermocouples can provide

reliable information if analyzed properly.

Thermocouple voltages give very accurate information about the tem-

perature of the junction. The relationship between junction temperature

and instantaneous gas temperature requires some evaluation. The primary

considerations are conduction effects between the wire and the probe

support, radiation exchange, catalysis of chemical reactions on the

probe surface, and the transient response of the probe to changing con-

ditions in the flowfield. Moffat (1962) provides a good introduction to

some of these problems and offers some considerations for dealing with

them.

Briefly, radiation and conduction corrections can often be

neglected, catalysis is minimized by ceramic coatings, and response can

be extended analytically or electronically. Since the physical size and

thermal capacity of even the smallest survivable thermocouple probe pre-

vent it from responding directly to fluctuations with frequencies above

approximately 200 Hz, response compensation is necessary. This tran-

sient, real-time, response compensation is absolutely essential for

time-resolved results and requires the most consideration on our part.

For the case of minimal radiation and conduction effects, the response

of the junction will be such that:

$$T_c = T_j + a(dT_j/dt) \tag{1}$$

where T_c is the corrected temperature, T_j is the junction tempera-
ture, and a is usually referred to as the time constant. The time con-
stant is the ratio of junction thermal capacity to convective heat flux.
Since this is related to the local instantaneous thermal properties of
the reacting gas mixture, the information required to calculate the time
constant directly is not available. The evaluation of the time con-
stant, and the implementation of this first-order response compensation,
are topics considered by many researchers.

II. OBJECTIVES OF THE RESEARCH EFFORT

The goal for the summer program was identify a technique for the
acquisition of spatially and temporally resolved temperature data in a
reacting flow, to identify procedures for the application of this tech-
nique, and to specify hardware and software requirements. The ultimate
goal of the research is to integrate these techniques with an existing
laser Doppler velocimeter and to acquire time-resolved, correlated velo-
city and temperature data in a swirling, reacting flowfield in a dump
combustor. Hardware lead-time restraints were expected to prevent the
implementation of this program during the summer term, so the short-term
emphasis was on system design.

III. LITERATURE REVIEW

An extensive review of thermocouple techniques was conducted during
the term of the Summer Faculty Research Program; this review has been
significantly edited to fit the space constraints of this document. The
primary items of interest are: flowfield type, size of junction or wire,
coating, radiation and conduction corrections, temporal resolution, and

response compensation. Virtually all the referenced experimental studies used wires of Pt and/or Pt-Rh.

Fristrom and Westenberg (1965) provide detailed instructions on the fabrication of small thermocouples and generating silica coatings for the elimination of catalysis effects. The coating technique involves passing the junction through a flame containing silicon dioxide generated from the combustion of a compound such as dimethyl siloxane. Kent (1970) reported that silica is not appropriate for temperatures above about 1100C, as free silicon diffuses into the junction, altering the properties of the thermocouple. An yttria-beryllia coating was suggested as an alternate, tested, and found to give good results.

Yule et al. (1978) offer a good discussion of transient compensation concepts, and provide a correction for non-cylindrical junctions. (Heat transfer to a cylinder has often been assumed in predicting thermal response.) The time constants for the actual junctions were determined by observing the response to a step-change in temperature by measuring the decay in measured temperature after a heating current was switched off. It was stated that this technique will not work unless the thermocouple wires to the junction are much smaller than the probe lead wires. Wire size and junction size were 25 and 70 microns, respectively; a typical time constant was 30 msec. Some reasonable agreement was achieved between predicted average time constants based on Nusselt number relations and measured data. In a propane-air diffusion flame, average time constants were determined at different locations in the flowfield by pulsed overheating (again, a response to a known step change in temperature) and used to correct the observed temperatures. These thermocouples were apparently not coated. Independent confirmation of the observed temperatures was not provided. A two-wire technique, with one

wire constantly heated to provide heat transfer data for response correction, was suggested as a future development.

Lockwood and Odidi (1973) measured mean temperatures in a turbulent diffusion methane flame with a 40 micron thermocouple. No corrections for radiation or conduction were made and the thermocouple was not coated. An average time constant for each location in the flame was determined by an AC excitation method, a typical value was 10 msec. This time constant was then applied to an on-line compensation circuit to obtain a corrected voltage. The use of an average, rather than instantaneous time constant was estimated to generate errors of up to 10 percent in the mean temperature and 20 percent in the instantaneous temperature.

Lockwood and Moneib (1980) measured the temperature in a nonreacting, electrically heated free jet with a 12.7 micron thermocouple compensated for transient response. (They also concluded that the results of the previous work were not as good as originally claimed.) An on-line, digital compensation network was developed; the appropriate time constant was still an average value determined by a pulsed overheat method. The time constant was on the order of 15 msec; noise limited the max frequency response to 5 kHz. Prodigious quantities of results were plotted, including mean profiles, PDF's, flatness, and skewness, and were described as "physically realistic."

Heitor et al. (1985) measured temperature simultaneously with LDV data in premixed, disc-stabilized, natural gas flames. The maximum error due to radiation was estimated to be 100K; a specific correction was not made. The thermocouple output was digitized and stored, with the compensation being done off-line by determining the gradient dT_j/dt from the stored data. An average time constant was not used, rather, an

instantaneous time constant was determined based on the measured velocity, physical characteristics of the junction, and a heat transfer relation for fine wires. This was a good idea, but didn't work very well in the reacting flows, as the gas properties are not uniquely determined by the temperature because of large variations in chemistry. Plots of temperature PDF's near the edge of the flame showed extreme overcompensation by this method, so the "instantaneous" time constants were all multiplied by 0.65 (trial-and-error value) to normalize the PDF's to ambient temperature. The junctions were left uncoated to keep the time constants as low as possible. Significant build-up of the titanium dioxide seeding particles on the thermocouple surface caused the response to degrade rapidly with run time, with LDV data rates of 100/sec being high enough to render the compensation technique essentially useless. The wire diameters were again 15, 40, or 80 microns.

Yoshida and Tsuji (1978) evaluated the temperature and velocity distributions in a premixed propane flame, but not simultaneously. The thermocouple was 50 micron diameter and uncoated. Transient compensation was accomplished in a conventional manner (initially) by determining the time constant from an overheat-response method, then incorporating this value in a 5 kHz, RC compensation circuit. This gave maximum instantaneous temperatures higher than adiabatic flame temperature, and minimum temperatures lower than ambient. To correct this discrepancy, a series of different time constants was used in the compensation circuit until a value was found which normalized the output between the ambient and adiabatic flame temperatures. This value (40 msec) was then used throughout the flowfield. Some "mismatching" was expected by the authors from this procedure. Temperature fluctuations of 400C, at frequencies above 1 kHz, were found.

Katsuki et al. (1987) developed a linearization technique for time-response compensation, and evaluated the procedure by rapidly vibrating the thermocouple junction across the flame front of a laminar diffusion flame. Both coated and uncoated wires were used. The technique utilized for compensation essentially determines a film temperature for the wire and evaluates the instantaneous time constant based on that temperature. Variations in chemistry are apparently not accounted for. Differences in measured values between coated and uncoated junctions were ascribed to radiative effects in laminar diffusion flames and catalysis in turbulent premixed flames. Results from the new technique were compared to results using a standard, average time constant compensation technique. Neither procedure was found to give especially impressive measurements; the authors concluded that simultaneous velocity data were necessary to properly account for the variations in time constant. It was also concluded that coating did not adversely affect the determination of real-time data; although the time constant was increased due to the larger diameter, compensation produced the same temperature PDF's as for the compensated, uncoated, thermocouple.

Lenz and Gunther (1980) used an uncoated, 50 micron, frequency compensated thermocouple to determine time-resolved temperature in a free-jet diffusion flame. Compensation was accomplished by determining the average time constant at each position in the flow by an overheat-response method, then using the resulting value in an on-line electrical compensation network. Conduction, radiation, and catalytic effects were not corrected. Response to 8 kHz was obtained, limited by noise. Several apparent discrepancies in the temperature data were explained by physical arguments.

Brum et al. (1983) used a 25 micron, uncoated, compensated thermocouple to acquire temperature data simultaneously with LDV data in a swirling, reacting flow. The compensation was similar to that of Lockwood and Moneib, using an in-situ overheat method to determine an average time constant, and on-line electronic compensation of the signal. The authors indicated that this average time-constant approach was not optimum, and attributed some apparent discrepancies in the data to the instrumentation. It was determined that a variation of 10 percent in the time constant could lead to errors as high as 50 percent in local heat flux. An alternative technique, utilizing the instantaneous velocity to modify the time constant, was suggested. (This approach was later used by Heitor et al., with minimal success.) Probe perturbation effects were also studied, with large variations in velocity discovered when the probe was inserted, due to both local and large-scale effects. This is to be expected in elliptic flows.

Farrow et al. (1982) compared CARS data to results obtained from uncompensated thermocouples in a laminar methane diffusion flame. The thermocouples were 50 micron diameter and silica coated. Conduction and radiation losses were accounted for off-line by using the symmetry of the flame to generate a known environment. The CARS and thermocouple temperatures were obtained simultaneously, with a "slight" distance between the measurement regions. Temporally resolved records were not presented; the two techniques provided good agreement on average temperature in some regions.

Chandran et al. (1984) measured temperature with a coated, 25 micron thermocouple, coincident with LDV data, in a premixed turbulent flame. An average time constant was determined by a cross-spectral analysis technique, using the response from two closely spaced thermocouples to

determine the time constants for both. It was claimed that the use of an average time constant does not introduce significant error in the temperature results. Probe interference was minimized by adjusting the probe configuration until the locally velocity closely matched the velocity measured when no probe was in the flow.

In a theoretical investigation, Chandran et al. (1985) used a two-time-constant model for a premixed flame with an assumed bimodal temperature PDF. A linear response was assumed for each of the two parts of the flow and time constants calculated using assumed flow properties. For an assumed square wave input, the response of the system was modeled and the calculated mean temperature was found. Using only a single time constant generated errors on the order of 10 percent.

Experimental evaluation of these results in a premixed methane-air flame was conducted with yttria-beryllia coated thermocouples of 25 and 75 micron wire, operated simultaneously. The junction separation was 1 mm. A cross-spectral analysis technique was used to determine the average time constant for each junction, assuming the same environment for both. An RC circuit was tuned to this mean value to accomplish on-line compensation. Comparison to coincident Rayleigh scattering results (with about 1 kHz resolution) showed reasonably good agreement for temporally-resolved temperature, and errors on the order of 20 percent for the RMS fluctuation and 10 percent for the mean temperature. These errors were predicted by the response model.

Elmore et al. (1986) conducted a specific program for dynamic temperature measurement using a dual-junction method with off-line compensation. The dual junctions were used to determine the response characteristics of the smaller thermocouple, which was then corrected. Cor-

rection was not done on a point-by-point basis, but rather in the frequency domain.

Essentially, the technique determined the convective heat transfer coefficient as a function of instantaneous frequency, corrected the measured temperature for convective response in the frequency domain via extensive use of FFT's, then inverse-transformed (when appropriate) to get temporal waveforms. Frequency response was limited to about 250 Hz. The development program was not intended to provide temperatures in reacting regions, so the assumption of air as the working fluid was justified. In this case, the compensation may very well be a function of frequency only. Changes in local chemistry, and catalytic effects, were not addressed. For reacting flows, this technique would probably suffer from the same problems observed by Heitor et al. A finite-difference conduction correction was applied to all data. A Fortran program to provide for compensation and conduction corrections was developed. The design and fabrication details for the probe, and initial testing, were described in an earlier report (Elmore et al., 1983).

Ozem and Gouldin (1989) used an uncoated, 25 micron thermocouple to determine temperature statistics in a turbulent, premixed, methane-air flame. Radiation and conduction were neglected, and response compensation was conducted off-line. The thermocouple output was amplified, low-pass filtered at 20 kHz, and digitized at 40kHz for transfer to a MicroVax II. An average time constant was used, calculated from a Nusselt number relation rather than an experimentally determined response.

The bulk of the evidence indicates that time-response compensation is necessary, at least in turbulent flows, if accurate, temporally resolved temperature data are to be obtained. The use of an average time constant does not appear to be sufficient, and using velocity alone

as an instantaneous "corrector" to the time constant does not provide much improvement as the response is a function of temperature, velocity, and chemical composition. A two-wire technique is definitely indicated, but frequency domain compensation loses the temporal information required for coincidence with LDV data and still does not account for chemistry. An instantaneous, real-time correction, based on the differential response between two thermocouples of different sizes, is the most promising approach.

IV. CONCEPT DEVELOPMENT

It will be assumed that conduction along the thermocouple wires is insignificant relative to convection between the junction and the flowfield; this has been shown to be valid for wire l/d ratios greater than about 50 (Elmore et al., 1983). Transient radiation will also not be accounted for; the thermocouple response will be;

$$T_c = T_j + a(dT_j/dt). \tag{1}$$

The time constant is related to the thermal capacity of the junction and the convective heat flux;

$$a = (\rho_j c_j d_j{}^2)/(4\, k\, Nu), \tag{2}$$

where ρ_j is the density of the junction wire, c_j is the specific heat of the junction wire, d_j is the junction diameter, k is the gas conductivity, and Nu is the Nusselt number. For small wires, the Nusselt number will have the form

$$Nu = c_1 Re^{C_2} Pr^{C_3}, \tag{3}$$

where Re is the Reynolds number based on wire diameter, and Pr is the gas Prandtl number. The relations for Nu, Re, and Pr can be substituted into the relation for the time constant a to yield

$$a = ((\rho_j c_j d_j{}^{2-C_2})/4)\, (\mu^{C_2}/(V^{C_2} \rho^{C_2} Pr^{C_3} c_1 k)). \tag{4}$$

The first grouping of terms is a function of the junction construction only, while the second group is related to the instantaneous gas properties, such that

$$a = C_4 G \tag{5}$$

and

$$T_c = T_j + C_4 \, G \, dT_j/dt. \tag{6}$$

For two junctions with differing values of C_4 exposed to the same environment, the instantaneous temperatures T_{jI} and T_{jII} will be different, out the gas parameter G and actual temperature T_c will be the same. With known values for the junction properties ρ_j, C_j, and d_j, the values of C_4 are known, and the only remaining unknown is G, which can be determined;

$$G = (T_{jI} - T_{jII})/((C_4 dT_j/dt)_I - (C_4 dT_j/dt)_{II}) \tag{7}$$

With the gas parameter known, the time constant for either junction can then be quantified, allowing a solution for the instantaneous gas temperature.

It should be noted that this is a dynamic calibration, dependent on the different responses of the two junctions to the same transient input. The only required information on the heat transfer characteristics is the Reynolds number exponent, C_2. This value will be about 0.5 (Moffat, 1962), but there is a sensitivity to this assumption. An overheat technique in a nonreacting flow can provide an empirical value for the exponent.

V. RECOMMENDATIONS

Temperature Measurement System

It is recommended that an experimental verification of this concept be conducted. An initial program will not require coincident velocity

measurement, but the data acquisition system should provide for this capability. The recommended system design point is 1 kHz LDV rate, 3000F gas temperature.

The probe should consist of two junctions, of 25 and 50 micron Pt/Pt-13%Rh wire (Type R). The separation of the two junctions should be approximately 0.5 mm, providing spatial resolution comparable to that of the LDV system. Lead lengths of 2.5 mm will be sufficient to minimize conduction effects. The junction wires are to be welded to larger diameter wires of the same composition, contained in a protective sheath. An yttria-beryllia coating should be applied after junction welding, using the technique described by Kent (1970).

The thermocouple outputs should be amplified and low-pass filtered. Appropriate amplifier/filter systems are available and quotes have been obtained. An amplification factor of 100 is needed for 2.0 volts output at design temperature; the recommended filter cuts are 10 Hz, 500 Hz, 1 kHz, 5 kHz, and 10 kHz.

The amplified output should then be digitized for analysis purposes. A minimum of three input channels are necessary, two for the amplified and filtered thermocouple signals, and a third for recording of velocity from the LDV system. Coordination of the temperature records with data from the LDV computer is thus possible. A initial data acquisition rate of 20 kHz is recommended. Approximately 10 to 100 temperatures should be recorded for each velocity, so that the gradient dE/dt can be accurately calculated for response compensation.

Compensation can be done off-line, using the known physical characteristics of the wires and the assumption that the two junctions are exposed to the same environment. While no papers using such a technique have been reviewed, the abstract of a Japanese-language paper (Yamazaki

et al., 1983) appears to describe a similar concept. (A copy of this paper has been obtained, and translation by USAF agencies has been requested.) Since the assumption of identical junction environments is required, a spectral analysis of the thermocouple voltages, including coherence and phase information, should be conducted as an indicator of the maximum reliable frequency response of the system.

Combustors

To properly characterize the capabilities of this novel measurement technique, tests should be conducted with multiple combustor geometries. Initial testing should be in a burner providing a laminar premixed flame and a concentric flow of inert gas. A variable frequency, electro-mechanical actuator system may be used to rapidly alternate the probe between the hot reaction zone and the cold outer flow. This will generate a known-frequency transient in the thermocouple output. The chemical composition and temperature of the outer flow can be changed to verify the convective heat transfer calculations in this region and provide added confidence in the analysis technique. Traversing frequency can be varied to identify both the uncompensated and compensated response limits of the probe.

Another test sequence should utilize a turbulent premixed flame, stabilized on a laboratory combustor. The probe would be stationary, relying on the turbulent fluctuations of the flame front for the transient component of the temperature signal. Evaluation of the thermocouple response can be based on observed response from the first test series and analysis of mean calculated temperatures and temperature distribution functions.

A third test series should then be conducted in a high velocity, ducted burner. Tests in this combustor will provide information on probe response and survivability in practical propulsion research systems.

After full characterization of this approach to real-time thermal data acquisition, it is recommended that a direct comparison with available non-intrusive diagnostic methods be conducted. For dynamic measurements, this would require a very clean flame and Rayleigh scattering apparatus; CARS data could be compared at low data rates.

REFERENCES

Brum, R. D., E. T. Seiler, J. C. LaRue, and G. S. Samuelson, "Instantaneous Two-Component Laser Anemometry and Temperature Measurements in a Complex Flow Model Combustor," AIAA-83-0334, 1983.

Chandran, S. B. S., N. M. Komerath, and W. C. Strahle, "Scalar-Velocity Correlation in a Turbulent Premixed Flame," 20th Symposium on Combustion, The Combustion Institute, pp. 429-435, 1984.

Chandran, S. B. S., N. M. Komerath, W. M. Grissom, J. I. Jagoda, and W. C. Strahle, "Time Resolved Thermometry by Simultaneous Thermocouple and Rayleigh Scattering Measurements in a Turbulent Flame," Comb. Sci. and Tech., vol. 44, pp. 47-60, 1985.

Elmore, D. L., W. W. Robinson, and W. B. Watkins, "Dynamic Gas Temperature Measurement System," NASA CR-168267, 1983.

Elmore, D. L., W. W. Robinson, and W. B. Watkins, " Further Development of the Dynamic Gas Temperature Measurement System," NASA CR-179513, 1986.

Farrow, R. L., P. L. Mattern, and L. A. Rahn, "Comparison Between CARS and Corrected Thermocouple Temperature Measurements in a Diffusion Flame," Appl. Optics, vol. 21, no. 17, pp. 3119-3125, 1982.

Fristrom, R. M., and A. A. Westenberg, Flame Structure, pp. 170-174, McGraw-Hill, 1965.

Heitor, M. V., A. M. K. P. Taylor, and J. H. Whitelaw, "Simultaneous Velocity and Temperature Measurements in a Premixed Flame," Exp. in Fluids, vol. 3, pp. 323-339, 1985.

Katsuki, M., Y. Mizutani, and Y. Matsumoto, "An Improved Thermocouple Technique for Measurement of Fluctuating Temperatures in Flames," Comb. and Flame, vol. 67, pp. 27-36, 1987.

Kent, J. H., "A Noncatalytic Coating for Platinum-Rhodium Thermo-couples," Comb. and Flame, vol. 14, pp. 279-281, 1970.

Lenz, W., and R. Gunther, "Measurements of Fluctuating Temperature in a Free-Jet Diffusion Flame," Comb. and Flame, vol. 37, pp. 63-70, 1980.

Lockwood, F. C., and H. A. Moneib, "Fluctuating Temperature Measurements in a Heated Round Free Jet," Comb. Sci. and Tech., vol. 22, pp. 63-81, 1980.

Lockwood, F. C., and A. O. Odidi, "Measurements of Mean, Root-Mean-Square, and Energy Spectrum of Fluctuating Temperature in a Round, Free, Turbulent Diffusion Flame", Comb. Inst. European Symp., The Combustion Institute, pp. 507-512, 1973.

Moffat, R. J., Temperature, Its Measurement and Control in Science and Industry, vol. 3, part 2, pp. 553-571, Reinhold, 1962.

Ozem, H. L., and F. C. Gouldin, "Temperature Time-Series Measurements in Premixed V-Shaped Flames," Paper 89-36, Eastern Section of The Combustion Institute, 1989.

Yamazaki, M., and M. Ohya, "Measurement of Fluctuating Temperature in Turbulent Flame by Dual Thermocouple Method," Nenryo Kyokai-Shi, vol. 62, no. 673, pp. 318-326, 1983.

Yoshida, A., and H. Tsuji, "Measurements of Fluctuating Temperature and Velocity in a Turbulent Premixed Flame," 17th Symposium on Combustion, The Combustion Institute, pp. 945-956, 1978.

Yule, A. J., D. S. Taylor, and N. A. Chigier, "Thermocouple Signal Processing and On-Line Digital Compensation," J. Energy, vol. 2, no. 4, pp. 223-231, 1978.

1986 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/
GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

<u>FINAL REPORT</u>

<u>Hydrogen Permeation in Metals at Low Temperatures</u>

Prepared by:	Kaveh A. Tagavi
Academic Rank:	Associate Professor
Department and	Mechanical Engineering
University:	University of Kentucky
Research Location:	WRDC/POOS-3 Wright-Patterson Air Force Base Dayton, OH 45433
USAF Researcher:	Michael J. Morgan/Jerry E. Beam
Date:	30 September 90
Contract No.:	F49620-88-C-0053

Hydrogen Permeation in Metals at Low Temperatures

by

Kaveh A. Tagavi

ABSTRACT

Hydrogen permeation through metals at cryogenic to low temperatures is considered in this report. Unlike permeation in high temperatures, there are very few data on low temperature permeation. Supercritical hydrogen has been suggested as a prime candidate for cooling of vacuum tubes aboard spacecraft. The knowledge about hydrogen permeation at low temperatures, therefore, is essential in establishing feasibility of this idea. In this project, a comprehensive literature search is conducted in order to document the state of the art research efforts on hydrogen permeation in metals at low temperatures. The phenomenon of permeation is investigated and the relevant parameters affecting it are identified. An apparatus based on vacuum method is constructed, fabricated, and calibrated. Gathering of actual permeation data are planned for future activities. As a part of this effort usage of exotic material such as gold, diamond, or zinc plated metals; graphite copper compounds; and beryllium oxide will be investigated.

ACKNOWLEDGEMENTS

I would like to thank Dr. Jerry E. Beam for selecting me for and entrusting me with this exciting and challenging project. He was always available for consultation and giving me overall guidance. Michael J. Morgan was there on a daily basis to encourage and support this project. I always enjoyed and benefited from his positive attitude. My thanks also goes to Donald Reinmuller for his dedication and constant effort to put life into this project.

My experience at WPAFB this summer was rewarding. The staff of the Thermal Technology Group, although diverse in technical expertise, is united on one philosophy -- make the visitor feel like he is at home. They certainly succeeded at this.

I am also thankful to Air Force Office of Scientific Research and the Universal Energy Systems for sponsoring my appointment.

Finally, my visit was enriched by racquetball games, coffee break discussions, and other social activities.

I. INTRODUCTION:

Using hydrogen as a coolant in space applications seems both logical and economical. It's engineering feasibility, however, is not obvious and has to be investigated. It is proposed that liquid hydrogen be on board as fuels, which makes it attractive as a coolant. A hydrogen coolant has two major advantages: first, it does not add substantially to the overall mass of the spacecraft (since it already exists on board as fuel) and second, it does not have to be recirculated and cooled since upon absorbing the thermal load it will be burned as fuel.

The Power Technology Branch of the Aerospace Power Division at Wright Research and Development Center has been involved in assessing the feasibility of hydrogen as a coolant for several years. An example of a device that needs to be cooled is vacuum tubes. However, one of the major concerns regarding cooling of vacuum tubes by hydrogen is hydrogen mobility through metals. As hydrogen comes into contact with metal surfaces it starts permeating through them. If enough hydrogen is permeated across the tube wall and enters into the tube, the ultra high vacuum that exists in the tube will be compromised. Therefore, it is essential that the magnitude of hydrogen permeation be estimated. Data on hydrogen permeation are readily available for common metals at mostly moderate to high temperatures (500-1300 K). But at lower temperatures of supercritical hydrogen (cryogenic to room temperatures, 35-300 K), little or no data exist.

My formal education and research interests have been in the general area of heat and mass transfer. My particular background and interest in space applications and microgravity-related research contributed to my assignment at the Power Technology Branch.

II. OBJECTIVES OF THE RESEARCH EFFORT:

Presently, there are no extensive data on hydrogen permeation through metals at cryogenic to low temperatures. It is the objective of the efforts initiated at the Power Technology Branch (WRDC/POOS-3) to establish a data-base on hydrogen permeation at low temperatures and evaluate the feasibility of using supercritical hydrogen as coolant for removing heat from microwave vacuum tubes. The specific objectives are:

- Perform a literature search and establish the knowledge-base available on the theories and mechanisms of hydrogen permeation.
- To construct, install, and calibrate a set-up which would measure hydrogen permeation at cryogenic to low temperatures.
- To establish a base-line permeation data and compare it with the existing data.

- To examine whether hydrogen permeation levels at cryogenic to low temperatures are significant enough to interfere with the normal operation of vacuum tubes.
- To investigate whether exotic materials such as gold, zinc, or diamond plated metals; graphite cooper; and beryllium oxide will inhibit or slow down hydrogen permeation.
- To recommended directions for future research and development efforts in this subject.

III. LITERATURE SURVEY:

Hydrogen permeation at high temperatures has been the subject of extensive research in the past due to its applications in such areas as nuclear reactors (fission and fusion), metal processing, and electrochemical processes. But only recently research on low temperature hydrogen permeation has been emphasized due to its applications in superconductivity and cooling in space. It has been suggested by Chow (1988) that the already available hydrogen on board as a fuel be used to overcome the thermal load on spacecrafts. This idea has many potential advantages along with some disadvantages. The most serious disadvantage is hydrogen permeation.

One of the devices that is proposed to be cooled by supercritical hydrogen is a microwave tube kept at very high vacuum. While being cooled by hydrogen, some hydrogen molecules may permeate inside the tube and corrupt the vacuum inside. This will cause malfunctioning of the microwave tube. It is hoped that at low temperatures, the rate of hydrogen permeation is low enough that a reasonable vacuum can be maintained during the operation time of such tubes (Pais, 1990).

Hydrogen is very mobile in metals. In fact, its mobility is several orders of magnitude larger than other interstitials such as oxygen and nitrogen. In addition to moderate to high temperatures, hydrogen diffusion can also be observed at low temperatures due to quantum effects at the molecular level (Schaumann et al., 1968; Wipf and Alefeld, 1974; Rowe et al., 1971).

IV. MECHANISMS OF PERMEATION:

To understand permeation one needs to understand solubility in, diffusion through, and chemical reaction with the metal in question. Once the metal wall is exposed to hydrogen, an equilibrium condition between the metal surface and the adjacent hydrogen is achieved.

Depending on the type of the metal, five types of reactions between hydrogen and the metal may occur.

1. Metal groups IA and IIA. Hydrogen forms hydrides. Examples of these metals are: Na and Ca.

2. Metal groups IVB, VB, and VIB. Hydrogen forms covalent hydrides. Examples of these metals are: C, Si, S, Se, and As.

3. Hydrogen forms true solutions in some metals. Examples of these metals are: Cu, Ag, Cr, Mo, W, Fe, Co, Ni, Al, and Pt.

4. Metal groups IIIA, IVA, and VA. Hydrogen forms pseudo hydrides. Examples of these metals are: Ce, La, Ti, Zr, Th, V, Cb, and Ta.

5. Hydrogen is not absorbed by Au, Zn, Cd, In, or Tl.

The relationship between the adjacent hydrogen partial pressure and the concentration of hydrogen just inside the metal surface is governed by the well-known Henry's law (Johnson, 1988). The higher the hydrogen partial pressure, the higher the concentration of hydrogen in the solid. Once, a concentration gradient is established, hydrogen is diffused in the direction of decreasing concentration according to Fick's law.

Based on state Henry's law, one expects to observe a linear relationship between the partial pressure of the adjacent hydrogen and the permeation rate Winkelmann was the first to suspect that the rate of permeation of hydrogen through platinum is not proportional to pressure (as was originally expected) but rather is proportional to square root of pressure. His explanation was that the hydrogen molecule first dissociates into atoms and then diffuses into and through the platinum plate. This theory was later confirmed by the careful experiments of Richardson et al. (1904) and Richardson (1904).

Besides pressure, hydrogen permeation depends on many other parameters such as:

- Temperature
- Surface condition
- Stress
- Atomic structure and kinetics
- Chemical reaction between hydrogen and material

In the following sections parameters that affect hydrogen permeation in metals are described.

V. MATHEMATICAL FORMULATION:

Assume a thin metal membrane is exposed to hydrogen. The temperature and pressure at each side in general is assumed to be different. As was mentioned before, the hydrogen molecule enters a chemical reaction (dissociation) in the presence of the metal and breaks into two hydrogen atoms. The hydrogen atoms subsequently dissolve

in the metal. According to Henry's Law the concentration of hydrogen just inside the metal surface will be

$$C \propto P_H \tag{1}$$

where C is the concentration and P_H is the pressure of the hydrogen atom gas. But due to the kinetics of the chemical reaction (dissociation)

$$P_H \propto P_{H_2}{}^{1/2} \tag{2}$$

and therefore

$$C(P,T) = S(T) \, P_{H_2}{}^{1/2} \tag{3}$$

where P_{H_2} is the pressure of hydrogen molecule gas and S is the solubility. The above equation is also known as Sieverts' law. The parameter S(T) is the constant of proportionality (indicative of solubility) which is solely a function of temperature and is independent of pressure and concentration. It is reported by Smithells and Ransley (1935) that S depends on temperature as

$$S(T) \propto T^{1/2} \exp(-E_S/RT) \tag{4}$$

where T is the temperature, E_S is the energy of solution, and R is the gas constant. E_S is positive for endothermic reactions such as H_2 in copper and negative for exothermic reactions such as H_2 in vanadium or niobium. The square term in equation (4) is usually ignored. It is argued that the temperature dependence of the square term is weaker than that of the exponential term for moderate and high temperatures. (This might not be as accurate for low to cryogenic temperatures.) The S expression may now be rewritten as:

$$S(T) = S_0 \exp(-E_S/RT) \tag{5}$$

where S_0 is the solubility pre-exponential. Combining these two equations yields:

$$C(P,T) = S_0 \exp(-E_S/RT) \, P^{1/2} \tag{6}$$

Hydrogen diffusion (permeation) caused by concentration gradient for a steady state situation may be written as

$$J_\infty = D \nabla C \tag{7}$$

where J_∞ is the steady state flux and D is the diffusion coefficient. Assuming pressure P (which results in concentration C) on one side and vacuum (which results in zero concentration) on the other side

$$J_\infty = DC/L \tag{8}$$

where L is the length along the diffusion path. Substituting for C

$$J_\infty = (D/L) \, S_0 \exp(-E_S/RT) \, P^{1/2} \tag{9}$$

It should be mentioned here that the flux is not proportional to the square root of the pressure difference rather the difference of the square root of pressure. Permeability is defined as

$$J_\infty = \Phi(T) \, (P^{1/2}/L) \tag{10}$$

That is

$$\Phi(T) = DS \tag{11}$$

where Φ is called permeability. The dependence of diffusion coefficient on temperature is written in terms of the energy of migration

$$D = D_0 \exp(-E_m/RT) \tag{12}$$

Accordingly

$$\Phi(T) = \Phi_0 \exp(-E_p/RT) \tag{13}$$

where D_0 and Φ_0 are diffusion coefficient and permeability pre-exponentials, and E_m and E_p are energies of permeation and migration. The energy of permeation E_p may be expressed as

$$E_p = E_s + E_m \tag{14}$$

and

$$\Phi_0 = D_0 S_0 \tag{15}$$

It is important to note that Φ is only a function of temperature and not pressure. This is the result of separating the pressure dependence in equation (10). Accordingly Φ_0 is neither a function of pressure nor temperature and is considered a constant.

VI. FACTORS INFLUENCING HYDROGEN PERMEATION:

There are many parameters that influence permeation. The mechanisms of many of the dependencies are not well understood. It should be mentioned that most of our present knowledge is based on experimental observation or on phenomenological reasoning. This plus the fact that not many experiments have been conducted for low temperatures, makes the low temperature stipulations risky.

Below, the dependency on some of the more important parameters are discussed.

PRESSURE -- Ordinarily it is expected that permeation is proportional to pressure because of Henry's law. However, since the hydrogen molecule first dissociates into hydrogen atoms and then diffuses, permeation becomes proportional to the square root of pressure. The latter is more or less universally accepted for the case of hydrogen permeation in metals. One exception is permeation under very low pressures where the permeation becomes proportional to pressure.

TEMPERATURE -- It is widely believed that temperature dependency follows Arrhenius relationship. However, a single Arrhenius relation might not successfully represent permeation for a large temperature range of interest -- namely 10 to 1000 K. Experimental data are available for temperatures down to 0° C (Johnson, 1988). There are some theoretical work (e.g., Valone et al., 1985), however, that considers permeation of hydrogen in copper at low temperatures down to 100 K. It was shown that if quantum

mechanics effects are included, the diffusion coefficient at low temperatures will deviate appreciably from those calculated based on classical theory. For example, diffusion of hydrogen in Cu(100) at 100 K will be 7200 times higher than that calculated based on classical theory (Valone et al., 1985).

This deviation at low temperatures has also been observed experimentally for hydrogen permeation in tantalum (Kokkinidis, 1977) for temperatures less than 200 K and as low as 100 K and in niobium (Schaumann et al., 1968) for temperatures less than 250 K and as low as 120 K.

One other known effect of temperature in the classical theory (Arrhenius relation) that has been ignored by most studies is the so called square term shown in equation (4). Figure 1 shows the ratio of permeability with and without the square term. This figure is plotted based on preserving the permeability value at the room temperature, hence the value 1 at 300 K. As seen the effect of this square term in not very large for temperatures near room temperature. As the temperature decreases, however, this effect increases. At around 50 K the square term would cut the permeability more than 50%. It should be mentioned that the elimination of this term causes overestimation of the permeation levels.

The concentration level does not always increase with temperature. If the enthalpy of solution is negative (due to exothermic solution reaction) then the level of concentration for a constant pressure decreases as the temperature increases. Similarly, if enthalpy of permeation is negative, the permeation rate will decrease with temperature. Figure 2 shows the hydrogen solubility levels for various metals (Fast, 1965). Therefore, in selecting an optimum material, we should avoid materials that react with hydrogen exothermically, since the cooling with hydrogen will occur at very low temperatures (as low as cryogenic temperatures).

CONCENTRATION -- According to the simple theory presented here, the diffusion coefficient does not depend on the level of concentration while the permeation is proportional to it (see equation (11)). Experimentally, however, some dependence on concentration has been observed. According to Gissler and Alefeld (1970) values of up to three times smaller for diffusion coefficient may be observed for higher concentrations of hydrogen permeating in niobium.

At high concentrations of hydrogen there is an overshoot effect (Beck et al., 1966). For example for a hydrogen-iron system, the permeation rate was observed to reach a maximum value and then decay to a steady state value. The overshoot was as much as 100% of the steady state value and lasted for several minutes.

SURFACE CONDITION -- Hydrogen permeation depends strongly on the surface

condition. Both the macroscopic and microscopic structure of the surface as well as the presence of oxides affect hydrogen permeation. The most severe problem with gas phase charging is the presence of oxides at the surface which are presumed to hinder hydrogen entry or exit at temperatures about 300 K (Kumnick and Johnson, 1975). This explains the fact that most charging experiments are performed for temperatures well over 300 K. It has been shown (Kumnick and Johnson, 1975) that this difficulty can be overcome by deposition of a thin layer of palladium on each surface of the permeation membrane. The palladium layer acts as an excellent catalyst for hydrogen dissociation. With this technique gas phase permeation methods may be carried out at temperatures well below 300 K.

STRESS -- Stress has been shown not to affect the diffusion coefficient, but it increases the solubility of hydrogen in the lattice (Beck et al., 1966). As a result, permeability, which is equal to the product of diffusion coefficient and solubility, will be increased under stress. The mechanical behavior of metals in the presence of hydrogen is further investigated by Toriano (1960).

VII. EXPERIMENTAL METHOD:

There are several methods for measuring diffusion (permeation) of hydrogen in solids. Volkl and Alefeld (1979), for example, have given details of all these techniques. One direct way of measuring permeation is a technique whereby a thin membrane of the metal of interest is exposed to pure hydrogen on one side and vacuum on the other. As hydrogen permeates through the membrane, the pressure on the vacuum side increases. By measuring the pressure variation with time, information about the diffusion coefficient, solubility, and permeability is obtained.

Permeation measurements may be performed in two different ways: stationary and time-lag methods (Volkl and Alefeld, 1979). The stationary method reveals the product of the diffusion coefficient and the solubility. The time-lag method reveals both diffusion coefficient and solubility. The diffusion coefficient can be deduced from the relaxation time while the solubility data can be determined from the change in permeation rate.

At steady state, permeation rate is proportional to the product of concentration and diffusion coefficient (Barrer, 1951). While transient data reveals trapping parameters in addition to concentration and diffusion coefficient (Johnson, 1988). It is reported by Johnson (1988) that the time lag decreases with temperature and pressure. For hydrogen permeation in iron, time lags between 1000 to 10,000 seconds were observed for a sample of about 0.8 mm at standard temperature and pressure.

The present set-up utilizes the vacuum technique. Figure 3 shows the schematic of the set-up. It essentially consists of two chambers. One contains hydrogen at slightly above atmospheric pressure while the other is maintained at ultra high vacuum (between 1×10^{-5} and 1×10^{-9} torr). The two chambers are separated by a copper disk, 0.080 inch (about 2.03 mm) thick, which also serves as the test specimen. All joints are sealed by rotatable flanges that utilize knife-edge technology. This will provide seals that are appropriate for maintaining ultra high vacuums. The desirable vacuum is reached by two pumping units -- a molecular pump and an ion pump. The molecular pump evacuates the chamber at pressures near the atmospheric levels down to 1×10^{-3} torr levels. At this point, the pumping action is switched to the ion pump which is capable of creating vacuums of up to 1×10^{-10} torr range. The pressure in the vacuum side is measured with a Convectron gauge (for low vacuum) and a Nude gauge (for ultra high vacuum).

The hydrogen and vacuum chambers are wrapped with heating tapes and cooling coils. Two options may be used as a liquid coolant inside the coils -- liquid nitrogen or coolant derived from an ultra low temperature (down to -90° C) circulating bath. The combination of the heating tapes and the cooling coils allows us to maintain any desired specimen temperature in the range -150° C to 250° C. The whole assembly including the pump is wrapped with thick insulation to facilitate temperature control.

VIII. EXPERIMENTAL PROCEDURE:

Before a permeation test can be performed, there are a couple preparatory tests that must be conducted: 1) leak detection, 2) system bake-out.

First, the vacuum chamber must be tested against leaks. One way this test is conducted is by turning off the vacuum pump after reaching an ultra high vacuum and observing the system pressure. A typical pressure-time history is shown in Figure 4. After the pump is turned off, the pressure starts to rise for two distinct reasons -- leaks and outgassing. (Outgassing is the diffusion of dissolved gasses out of the chamber inner surfaces and into the vacuumed space. The driving force for outgassing is the difference between the chemical potentials of the dissolved gas and of the same gas in the chamber. That is, the lower chamber pressure the higher the rate of outgassing.) If there is a leak, the pressure ultimately equals that of the surrounding atmosphere. In the absence of any significant leak, the pressure in the vacuum chamber rises to a unique pressure (equilibrium pressure) corresponding to the amount of gases dissolved at the inner surfaces of the vacuum chamber. After this pressure is reached, the rate of outgassing quickly drops to zero and no further significant increase in the chamber pressure is observed. In order to lower the equilibrium pressure, a procedure called baking-out is

employed.

The chemical potential of the dissolved gases increases drastically with temperature. Therefore if the chamber is heated and vacuumed simultaneously, considerable amounts of dissolved gases may diffuse out and be taken out through the vacuum system. After the system is cooled, a lower level of equilibrium pressure may be attained due to the decreased level of dissolved gases. Baking periods of up to several days at elevated temperatures of 150-250° C are recommended in order to obtain ultra high vacuums.

After the leak detection and baking out the set-up is ready for permeation tests. The principle tenet of the experiment is to measure the time (Δt) it takes for the chamber pressure to rise from the initial pressure of P_i (it is best that this pressure be greater or equal to the equilibrium pressure) to the final pressure of P_f. Assuming that the pressure of the hydrogen chamber (P_{H_2}) is much greater than P_f and that it is held constant during the experiment, and using equation (10), the following relationship between the diffusion rate and the pressure buildup inside the vacuum chamber may be written:

$$\Phi = [V(P_f - P_i)L]/[RT\, P_{H_2}{}^{1/2}\, A\Delta t\, \rho_{stp}] \tag{16}$$

where V is the volume of the vacuum chamber, A is the area of the test specimen available for permeation and ρ_{stp} is the density of hydrogen at standard temperature and pressure. Figure 5 shows the amount of time necessary for the pressure rise from 1×10^{-5} to 1×10^{-4} torr for our present set-up with copper as the specimen. As seen, at room temperature, it takes about 2.15 hours for the pressure buildup. However, for lower temperatures, the time required increases rapidly. For example, at 200 K, it takes about 100 million hours for the same pressure buildup. It should be mentioned that this figure is based on Arrhenius relation and that this relationship might not hold for lower temperatures.

IX. MATERIALS OF INTEREST:

A few materials have been selected for testing in this work each for a specific reason.

- Copper is selected for the baseline material. The set-up will be checked and calibrated against data available in the literature.
- Diamond is selected with hopes that its compact and uniform molecular structure would inhibit hydrogen permeation.
- BeO is selected because of its excellent thermal, structural, and electrical properties.

- Graphite compounds such as GrCu is selected for their excellent structural properties.

In the following sections the available data for each of these materials are listed.

COPPER -- There are several references dealing with hydrogen permeation through copper. One of the earliest, yet widely referred, is by Smithells and Ransley (1935). Table 1 lists the diffusion data for hydrogen-copper system as well as other permeation data. More results are shown in Table 2 taken from Dushman (1962). This table lists the hydrogen solubility data for various temperatures. Alloys of copper have received some attention too. It has been shown that the hydrogen solubility varies as a function of the amount of alloying metal. It is seen that the effect of the alloying covers a wide range of solubility enhancement for Cu-Ni to decrease in solubility for Cu-Al. Figure 2 shows the solubility data vs. temperature for Cu as well as other metals.

The diffusivity of hydrogen in copper at low temperatures was studied by Ishikawa and McLellan (1985). The data presented are for temperatures as low as 300 K. Figure 6 show these results.

DIAMOND -- No data was found for solubility, diffusion, or permeation of hydrogen through diamond. However, a point of concern should be any possible reaction of hydrogen and diamond (carbon) which may result in methane (CH_4). The intensity and severeness of such reaction (if any) should increase with temperature. This in turn increases the importance of predicting hot spots.

BeO -- No data was found for solubility, diffusion, or permeation of hydrogen in BeO.

GRAPHITE COMPOUNDS -- No data was found for solubility, diffusion, or permeation of hydrogen through graphite compounds such as Gr-Cu. However, a few reports are available for graphite (e.g. Kiyoshi et al., 1988; Karimi and Vidali, 1989).

X. CONCLUDING REMARKS/FUTURE DIRECTIONS:

Creating, maintaining, and measuring ultra high vacuums (e.g., 1×10^{-9} torr) are by no account a trivial matter. The ten week period fell just short of obtaining accurate and repeatable hydrogen permeation data. As a result, this report is focusing on theoretical background, explaining the physical phenomena, and exploring future directions. Plans are undergoing for obtaining low temperature permeation data in the future. Below, some specific remarks and suggestions are listed.

- Since we are dealing with very large time intervals (days or even weeks), the design should be modified so that larger hydrogen pressures and thinner test specimen could be used.
- The experimental set-up and procedure should be tested against known data.

Choices of pure copper or palladium (a precious metal) at or above room temperature seem appropriate.

- In choosing the material, we should exercise caution in selecting those which react with (or dissolve) hydrogen exothermically. The solubility of these materials increase with decreasing temperatures.

- It is not clear whether BeO reacts with hydrogen endothermically or exothermically.

- It is reported that certain metals (i.e., Au, Zn, Cd, In, and Tl) do not absorb hydrogen. Therefore, these metals might be used as barriers to hydrogen permeation. For example, gold plating a sample might inhibit hydrogen permeation to a considerable degree.

- Use of diamond coating has been suggested as a hydrogen permeation barrier. For testing purposes, a metal with a large hydrogen absorption should be coated with diamond. Since the metal virtually presents no resistance to hydrogen permeation, the experimental data would reveal permeability of the diamond coatings.

- The literature is scarce when it comes to permeation data for temperatures lower that 300 K. This plus the fact that the mechanism of hydrogen absorption and permeation is not fully understood (even for higher temperatures) suggest that we need data for the actual temperature of operation. In short, data for the range of room temperature down to 250 K might not necessarily predict data for 200 K or 100 K. One purely theoretical paper suggests that the diffusion coefficient of hydrogen in copper when including the quantum mechanics effects will be 7200 times higher than that calculated according to the classical (Arrhenius) relationships.

- Even if experimental data show that hydrogen permeation levels are not significant at supercritical hydrogen temperatures (about 40 K), we should try to design the system for higher temperatures. The supercritical hydrogen is cooling a surface that might be very hot (several hundred Kelvin). Any disruption or instability in the cooling mechanism might create temporary hot spots in the vacuum tube wall. These hot spots will act as open windows for hydrogen permeation. This scenario should encourage the search for material which are more impermeable to hydrogen (such as gold or diamond plated metals?).

REFERENCES

Barrer, R. M., Diffusion In and Through Solids, Cambridge University Press, Oxford University, England, 1951.

Beck, W., Bockris, J. O'M., McBreen, J., and Nanis, L., "Hydrogen Permeation in Metals as a Function of Stress, Temperature and Dissolved Hydrogen Concentration," Proceedings, Royal Society, London, Series A, Volume 290, 1966, pp. 220-235.

Chow, L. C., "Forced Convection Cooling of Microwave Tube Collector with Supercritical Hydrogen," Final Report SCEEE Contract F30602-81-C-0193, Task 0091, 1988.

Dushman, S., Scientific Foundations of Vacuum Technique, Second Edition, Edited by J. M. Lafferty, John Wiley, New York, 1962.

Fast, J. D., Interaction of Metals and Gases, Volume 1, Thermodynamic and Phase Relations, 1965.

Gissler, W., Alefeld, G., Springer, T., J. Phys. Chem. Sol., Volume 31, 1970, p. 2361.

Ishikawa, T. and McLellan, R. B., "The Diffusivity of Hydrogen in Copper at Low Temperatures," J. Phys. Chem. Solids, Volume 46, Number 4, 1985, pp. 445-447.

Johnson, H. H., "Hydrogen in Iron," Metallurgical Transactions B, Volume 19B, October 1988, pp. 691-707.

Karimi, M. and Vidali, G., "The Adsorption of H_2, D_2 and Ar on Graphite: New Theoretical Results," Surface Science, Volume 208, 1989, pp. L73-L79.

Kiyoshi, T., Namba, T., and Yamawaki, M., "Hydrogen Permeation through Graphite," Journal of Nuclear Materials, Volumes 155-157, 1988, pp. 230-233.

Kokkinidis, M., Diploma Thesis, Technical University Munchen, Germany, 1977.

Kumnick, A. J. and Johnson, H. H., "Steady State Hydrogen Transport through Zone Refined Irons," Metallurgical Transactions A, Volume 6A, 1975, pp. 1087-1091.

Pais, M. R., "Permeability of Hydrogen in Silicon," SBIR Proposal submitted to and funded by U.S. Department of Defense, 1990.

Richardson, O. W., "The Solubility and Diffusion in Solution of Dissociated Gases," Philosophical Magazine, Sixth Series, Volume 7, 1904, pp. 266-274.

Richardson, O. W., Nicol, J., and Parnell, T., "The Diffusion of Hydrogen through Hot Platinum," Philosophical Magazine, Sixth Series, Volume 8, Number 43, July 1904, pp. 1-29.

Rowe, J. M., Skold, K., Flotow, H. E., and Rush, J. J., J. Phys. Chem. Sol., Volume 32, 1971, pp. 41-54.

Schaumann, G., Volkl, J., and Alefeld, G., "Relaxation Process due to Long-Range Diffusion of Hydrogen and Deuterium in Niobium," Physical Review Letters, Volume 21, 1968, pp. 891-893.

Smithells, C. J. and Ransley, C. E., "The Diffusion of Gases through Metals," Proceedings, Royal Society, Series A, Volume 150, 1935, pp. 172-197.

Troiano, A. R., "The Role of Hydrogen and other Interstitials in the Mechanical Behavior of Metals," Transactions ASM, Volume 52, 1960, pp. 54-80.

Valone, S. M., Voter, A. F., and Doll, J. D., "The Isotope and Temperature Dependence for Self-Diffusion for Hydrogen, Deuterium, and Tritium on Cu(100) in the 100-1000 K Range," Surface Science, Volume 155, 1985, pp. 687-699.

Volkl, J. and Alefeld, G., "Hydrogen Diffusion in Metals," in Diffusion in Solids, edited by Nowick, A. S. and Burton, J. J., Academic Press, New York, 1979, pp. 231-302.

Winkelmann, A., Drude's Ann. Volume VIII, p. 388.

Wipf, H., and Alefeld, G., Phys. Stat. Sol. (a), Volume 23, 1974, pp. 175-186.
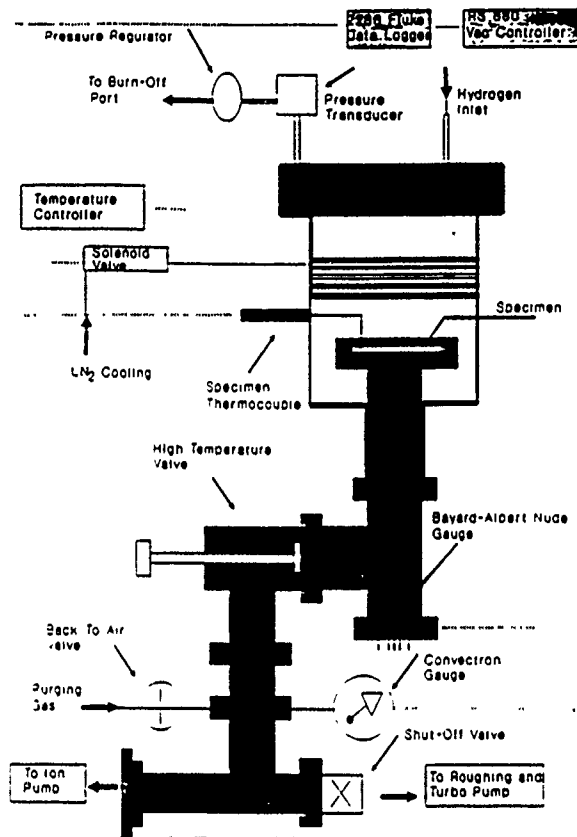
Figure 3.   Test set-up.
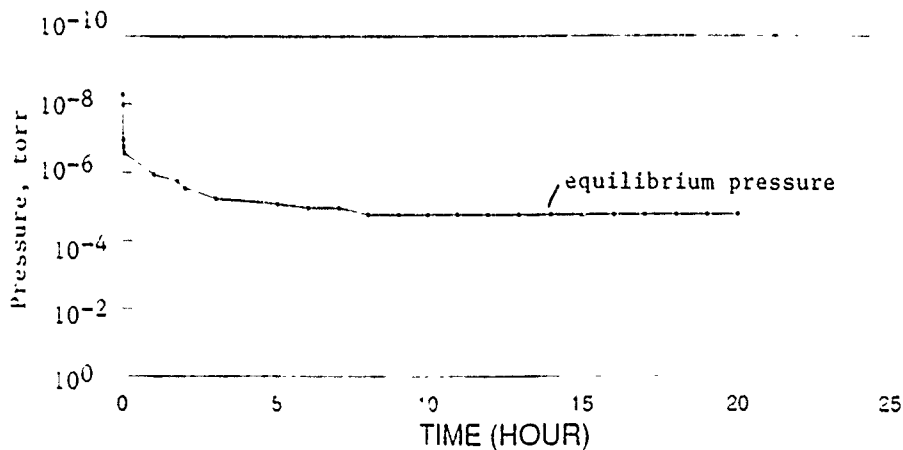
Figure 4.   Baseline test with no hydrogen permeation.
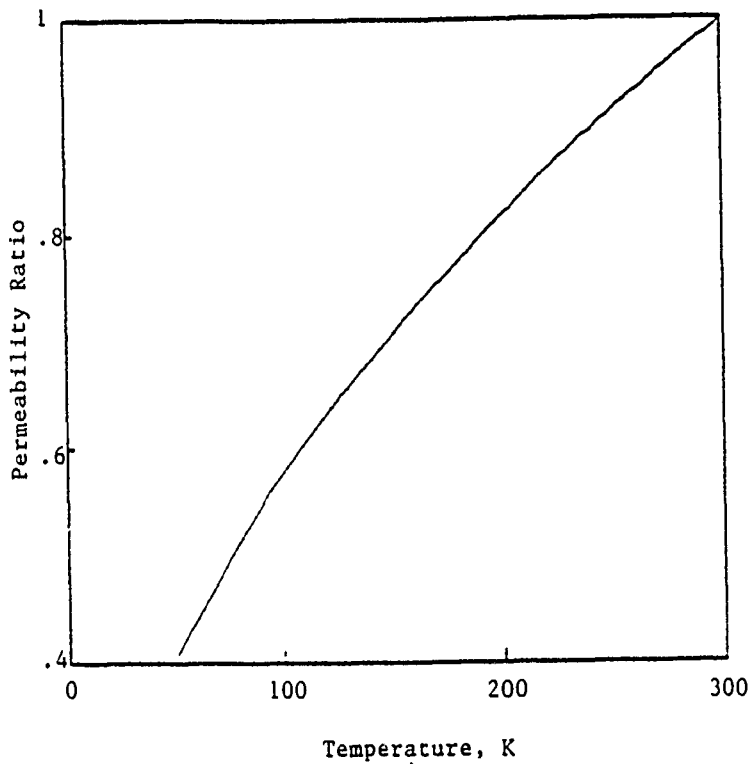
Figure 1. Effect of T^2 term on permeability
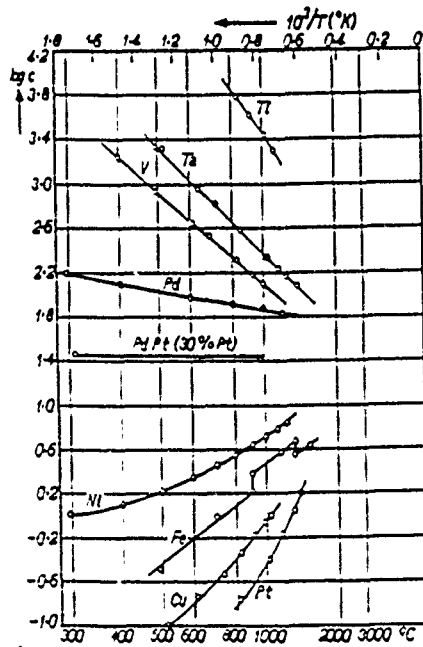
Figure 2. Solubility data for various
metals from Fast (1965).

Table 1. Diffusion data for hydrogen and
various metals from Smithells
and Ransley (1935). D is in
STP cm^3 per second through 1 cm^2
of surface and across 1 mm thickness,
p is in mmHg.

$$D = k \cdot \frac{1}{d} \cdot \sqrt{P} \cdot e^{-b/T}$$

(1) System		(2) b	(3) k	(4) Authors
H_2	Ni	7,710	$2 \cdot 3 \times 10^{-2}$	Lombard.
		6.930	$0 \cdot 85$	Deming and Hendricks.
		6,900	$1 \cdot 4$	Borelius and Lindblom.
		6,700	$1 \cdot 05$	Ham.
H_2	Pt	9,800	$1 \cdot 41 \times 10^{-2}$	Richardson.
		9,000	$1 \cdot 18$	Ham.
H_2	Mo	10,100	$0 \cdot 93 \times 10^{-2}$	Smithells and Ransley.
H_2	Pd	2,100	$4 \cdot 1 \times 10^{-2}$	Lombard and Eichner.
H_2	Cu	8,300	$0 \cdot 91 \times 10^{-3}$	Smithells and Ransley.
H_2	Fe	4,800	$1 \cdot 63 \times 10^{-3}$	Smithells and Ransley.
		4,700	$1 \cdot 60$	Borelius and Lindblom.
O_2	Ag	11,300	$3 \cdot 75 \times 10^{-2}$	Spencer.
		11,300	$2 \cdot 06$	Johnson and Larose.
N_2	Mo	22,500	$8 \cdot 3 \times 10^{-2}$	Smithells and Ransley.
N_2	Fe	11,900	$1 \cdot 5 \times 10^{-4}$	Ryder.
CO	Fe	9,300	$1 \cdot 3 \times 10^{-3}$	Ryder.

Table 2. Hydrogen solubility in copper from
Dushman (1962). V_0 = cm^3(STP) per 1 g
metal, s = cm^3(STP) per 100 g metal,
V_g = volume of gas (STP) per 1 volume
of metal, v = number of atoms per
atom of metal.

t° C:	400	500	600	700	800	900	1000
s:	0 06	0.16	0.30	0.49	0.72	1.08	1 58
V_g:	0 0054	0.0143	0.0268	0.0439	0.0644	0.0967	0 141
$10^4 v$:	0.053	0.0908	0.170	0.278	0.409	0.613	0.896

t° C.	1083 (mp)		1100	1200	1300	1400	1500
s:	2.10 (s)	6.00 (l)	6.3	8.1	10.0	11.8	13.6
V_g:	0.188	0 537	0.564	0.725	0.893	1.05	1.22
$10^4 v$:	1.19	3.40	3.58	4.60	5.67	6.70	7 71

* Values of V_g and v calculated by Dushman.
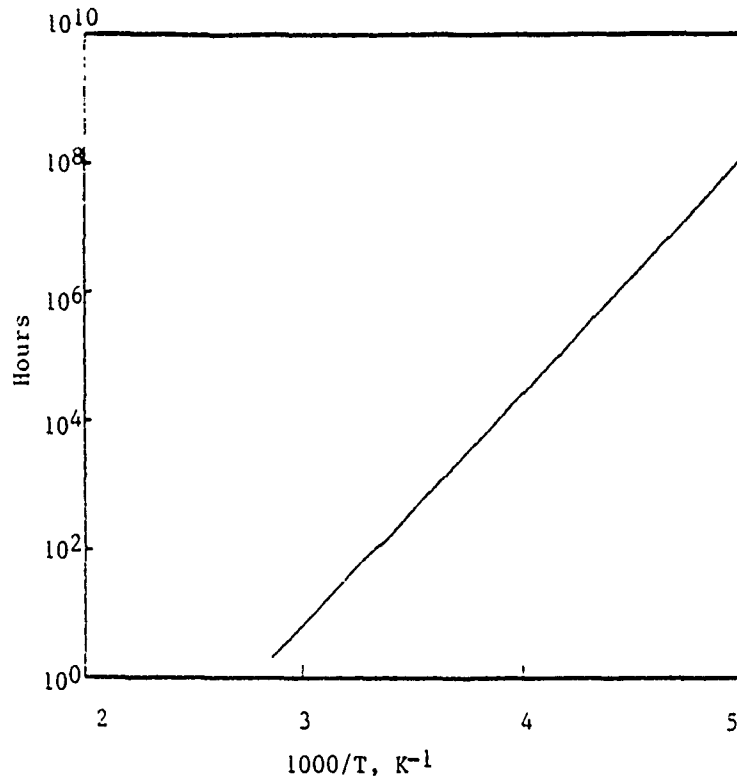
Figure 5.   Estimated time it takes for the present
set-up to go from 10^{-4} torr to 10^{-5} torr.
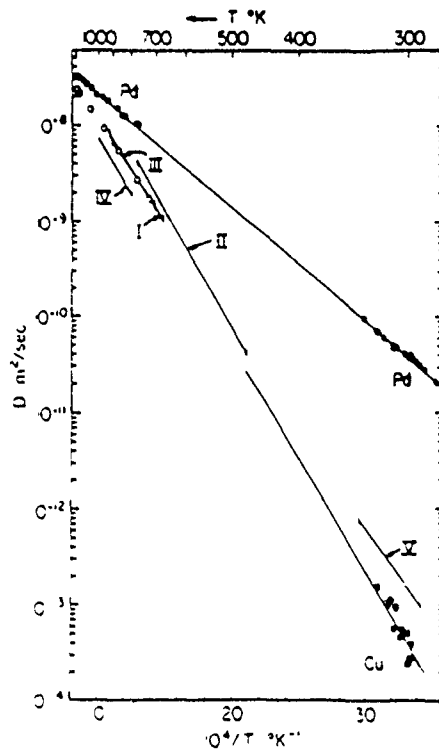
Figure 6.   Diffusivity of copper
at low temperatures
from Ishikausa and
McLellan (1985).

1990 USAF - UES SUMMER FACULTY RESEARCH PROGRAM

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

<u>FINAL REPORT</u>

<u>MEASUREMENTS OF DROPLET VELOCITY AND SIZE</u>

<u>DISTRIBUTIONS FOR A PRESSURE/AIR BLAST ATOMIZER</u>

Prepared by:	Richard S. Tankin
Academic Rank:	Professor
Department and	Mechanical Engineering Department
University:	Northwestern University
Research Location:	WRDC/POSF
	Wright Patterson Air Force Base
	Dayton, Ohio 454336563
USAF Researcher:	Thomas Jackson
Date:	21 August 1990
Contract No:	F49620-88-C-0053

MEASUREMENTS OF DROPLET VELOCITY AND SIZE

DISTRIBUTIONS FOR A PRESSURE/AIR BLAST ATOMIZER

by

Richard S. Tankin

Abstract

A phase doppler instrument was used to measure droplet sizes and velocities in a water spray. This nozzle consisted of a hollow cone water spray and two swirling, concentric air channels. Three different water flow rates were examined; and three different air flows. Horizontal traverses were made across the spray near the sheet break-up region. More than 200,000 samples were taken in each traverse. The results show that the spray is axially symmetric which is important for the planned theoretical analysis. The analysis of the data will follow the same procedure that was developed to analyze the data that was collected last summer.

Acknowledgements

I wish to thank WRDC/POSF at Wright Patterson Air Force Base and the Air Force Office of Scientific Research for the sponsorship of this research. I also want to thank Universal Energy Systems for their efficient handling of all administrative aspects of the program.

This summer's research was beneficial to me and hopefully will lead to joint research activities with Wright Patterson. Having worked before with Dr. W. M. Roquemore and Dr. T. Jackson, it was no surprise to find them very cooperative and encouraging. The experimental work could not have been accomplished without the expertise and experience of Dr. G. Switzer. The ease and readiness with which Jeff Stutrud handled complex computer problems was invaluable. I found the staff at Wright Patterson to be friendly, stimulating, and helpful.

I. INTRODUCTION

Gas turbines in nearly all cases utilize liquid fuels. These fuels must be delivered as small droplets to the combustion zone if efficient combustion is to occur. The process by which these small drops are formed is called atomization. The method of achieving atomization is varied and not a topic for consideration in this study. In this study the droplet distributions from a pressure/air blast atomizer are examined. The liquid spray discharges from the nozzle as a liquid sheet which breaks up downstream to form droplets.Surrounding this liquid sheet are two concentric, swirling air flows. Pictures of the spray from this nozzle were taken and a typical photograph is shown in Figure 1. The light source for this photograph was a strobe (pulse duration is about 100 μsec).

It is important to determine the spray characteristic if one hopes to correlate different sets of experimental data, compute numerical simulations, determine droplet trajectories, etc. In the past, these characteristics were limited to drop size distribution, patternation, cone angle, dispersion, and penetration. For example various techniques have been used to determine the drop size distribution - photographic, optical, collection devices, etc. Each of these techniques have very limited capabilities. The Malvern technique obtains a size distribution that is integrated over the optical path length. To determine radial distribution requires the use of Abel inversion. Dense sprays, or asymmetry can complicate this technique.Two situations have recently arisen - a vast improvement in instrumentation and a new approach for predicting droplet distributions.

Ia. INSTRUMENTATION

Recently a highly sophisticated instrument has been developed by Aerometrics which is capable of optically measuring the sizes and velocity of droplets. This method utilizes light scattering techniques, and consists of a transmitter, receiver, signal processor, computer an' laser light source. The transmitter has a beam expander which reduces the size of the measuring point - which in our case is

about 1 x $10^{-4}cm^2$. Thus excellent spatial resolution is achieved. Since the detectors in the receiver unit are photomultipliers (three), the response time of this instrument is very short. Signals from individual drops can be processed and the data transferred to computer memory in 20 μsec. For the operator, an important aspect of this system is the software program associated with the signal processor. As data are being collected and stored by the computer, they are presented in bar graph form for the operator to observe. After a selected number of droplets are collected, a listing of pertinent data - such as attempts, validations, corrected count, probe area, etc. - are displayed. The bar graphs consist of particle size counts, and velocity distributions. A typical example of such a display is shown in Figure 2.

Last summer at Wright Patterson Air Force Base , I collected similar data from a pressure atomizer. These results were presented at the 1990 Institute for Liquid Atomizarion and Spray Systems in Hartford, Conneticut (Li, Chin, Tankin, Jackson, Stutrud, and Switzer). This summer we extended the work to include the effects of high velocity, concentric, swirling air. To achieve our goal, nine different conditions were examined - three different water flow rates and three different air flow rates. I might add that this is a very complicated flow and if we can successfully predict the size and velocity distributions for this nozzle, I think we can analyze any nozzle.

Ib. THEORY

The concept of information entropy was developed by Claude Shannon (1948), and Jaynes (1957) later extended this concept into the well-known method of maximum entropy formalism. This formalism can be applied to problems involving probability, i.e., where insufficient information is available to obtain exact solutions. We have applied the maximum entropy formalism to liquid sprays in which we predicted the droplet size and velocity distributions. Since the application to this problem has been adequately discussed by several researchers - Kelly (1976), Sellens and Brzustowski (1985, 1986), Sellens (1989), and Li and Tankin (1987, 1988, 1989), it will not be necessary to develop the background material once again (for details, see Li 1989). The data collected in this study will be examined using the maximum entropy principle. There has only been two papers in which comparisons between theory and measurements exist. In one, Li and Tankin

(1988) used data that were obtained from holographic and photographic methods. Thus these results were limited to size distributions. The other paper is by Sellens (1989) in which he use a phase doppler instrument to obtain the data. However in this study there are some inconsistencies and the experiments are questionable (see Li and Tankin, 1989). Thus it was necessary to carefully obtain correct data with which to compare the theory. That is the purpose of this study.

II. OBJECTIVES OF THE RESEARCH EFFORT:

Currently, there are very few precise measurements of radially integrated droplet size and droplet velocity distributions for sprays. Our goal is to obtain such measurements under these various flow conditions and to present them in a two-dimensional matrix form - one axis being the droplet diameter, and the other being the droplet velocity. From such a two-dimensional matrix of the data, one can then obtain a variety of other representations - droplet size distribution; velocity distributions for different diameter particles; total number of droplets; flow rates; average velocity as a function of drop size, etc. These measurements should be made as close to the sheet break-up region as possible.

My task this summer was to work with the staff at Wright Patterson Air Force Base making radially-integrated spray measurements using a phase doppler particle analyzer and to obtain the data in a two-dimensional matrix form. Then to extract from this matrix the various probability distribution functions, average droplet velocity as a function of droplet diameter, flow rate, etc.

These experimental data are important because they can be used to extend the applicability of recently developed theoretical predictions. One problem in collecting such data is the requirement that samples be sufficiently large for the statistical quantities to be significant. In our case we collected more than 200,000 samples per run. Nine such runs were completed.

III. RESULTS

We measured the radial droplet population for three different water flow rates - 4.2 ml/sec, 7.6ml/sec and 8.33ml/sec. The three air flow rates, for each water flow rate, were zero, 1,833ml/sec, and 3,000 ml/sec. When the water flow rate is

4.2ml/sec, measurements were made in a horizontal plane located 10 mm from the nozzle exit; when the water flow rate is 7.6ml/sec, measurements were made 7mm from the nozzle; and when the water flow rate is 8.3 ml/sec, measurements were made 5mm from the nozzle. The radial population was obtained from measurements at 0.5mm increments over the radius of the spray. The velocity measurements and Sauter mean diameter measurements indicate the spray is axially symmetric (see Figure 3). It should be added that during each of these traverses it was necessary to shut down the spray at least once, refill the reservoir, and reset the flow rate. Although there may be some variations in reestablishing the flow, we did not detect it once the flow was restablished. Since there are two air passages, we needed to know the ratio of the air flow through each passage. To accomplish this, we blocked each air passage separately and measured the pressure drop for various air flow rates. The results are seen in Figure 4 - which indicate the air flow is equally divided between the two passages.

A joint size-axial velocity distribution function will constructed from the individual point measurements, weighing each measurement by their time of collection and the ratio of their optical probe area to the ring area represented at that location. This experimentally determined joint distribution function will then be compared to that predicted by the maximum entropy analysis.

The experimental data will be similarly handled as the data obtained last summer. It will be necessary to make some assumptions regarding the source terms in the theory. The mass source term - S_m - will be assumed to be zero. The other two source terms - momentum (S_{mv}) and energy (S_e) - will require estimates after examining the experimental data that were collected.

IV RECOMMENDATIONS

The necessary experimental data have been collected and stored on floppy disks. It may be necessary to obtain from Aerometrics additional information that will allow us to read these data and calculate the probe area for those runs in which this information is lacking. Then the integrated two-dimensional matrix will be formed. Once we have the experimental data in this form, the various distributions and mean values can be determined. An additional quantity that needs to be

compared is the flow rate obtained from the droplet measurements with the flow collected in a graduate over a period of time.

Since this nozzle was supplied to us by Allison Gas Turbine Division; we will report to them, through Wright Patterson, our results. Hopefully similar experiments will be conducted on a combusting flow using this nozzle. Once these data are collected, analysis can begin on the source terms for such a flow. In this flow there will be heat, mass, and momentum transfer - much different from the cold flow experiments conducted this summer.

VI. REFERENCES:

1. Shannon, C. E., "A Mathematical Theory of Communication ", The Bell System Technical Journal, 27, 1948, p.379.

2. Jaynes, E.T., "Information Theory and Statistical Mechanics", Physical Review, 106, 1957, p. 620.

3. Kelly, A. J., "Electrostatic Metallic Spray Theory", J. of Applied Physics, 47, 1976, p. 5264.

4. Sellens, R. W. and Brzustowski, T. A., "A Prediction of Drop Size Distribution in a Spray from First Principle", Atomization and Spray, 1, 1985, p. 89.

5. Sellens, R. W. and Brzustowski, T. A., "A Simplified Prediction of Droplet Velocity Distributions in a Spray", Combustion and Flame, 65, 1986, p. 273.

6. Li, X. and Tankin, R.S., "Droplet Size Distribution:A Derivation of a Nukiyama-Tanasawa Type Distribution Function", Combustion Science and Flame, 56, 1987, p. 65.

7. Li, X. and Tankin, R.S., "Derivation of Droplet Size Distribution in Sprays Using Information Theory", Combustion Science and Flame, 60, 1988, p. 345.

8. Li, X. and Tankin, R.S., "On Prediction of Droplet Size and Velocity Distributions in Sprays Through Maximum Entropy Formalism", <u>Combustion Science and Technology</u>, 68 , 1989, p. 147.

10 Li, X., PhD Thesis, Department of Mechanical Engineering, Northwestern University, June, 1989.

11. Li, X., L.P. Chin, R.S. Tankin, T. Jackson, J. Stutrud, and G. Switzer, "Comparison Betwwen Theory and Experiments for Sprays From a Pressure Atomizer", 1990 ILASS Meeting in Hartford, Conneticut.
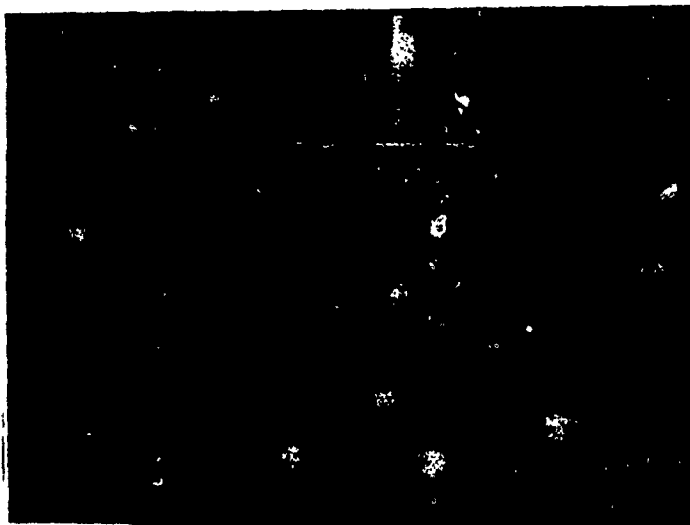
a
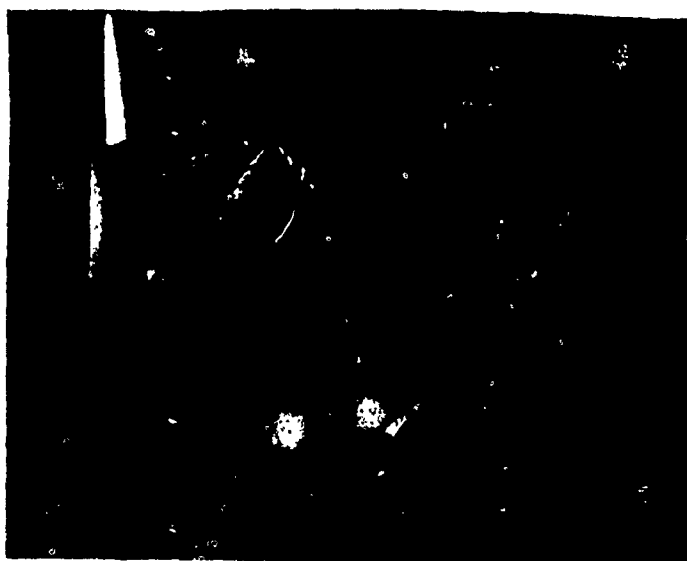
b

Figure 1a. This is a photograph of target, where each small division is 1mm. Figure 1b is a photograph of the spray when the water flow rate is 8.3ml/sec.
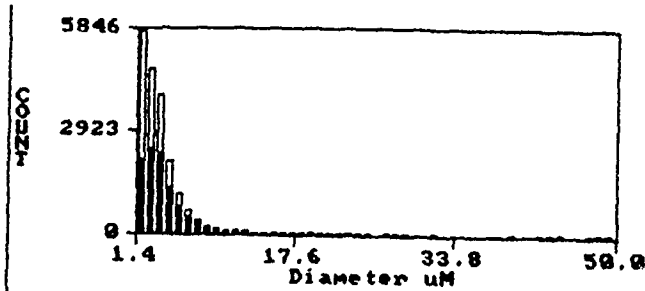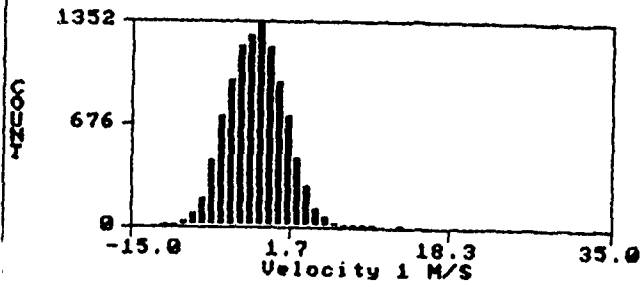
Figure 2a. These are typical bar graph outputs from the phase doppler analyzer. Along with the velocity and size distributions are mean velocity, mass mean diameter (D30), Sauter mean diameter (D32), etc. These values were obtained along centerline of spray.
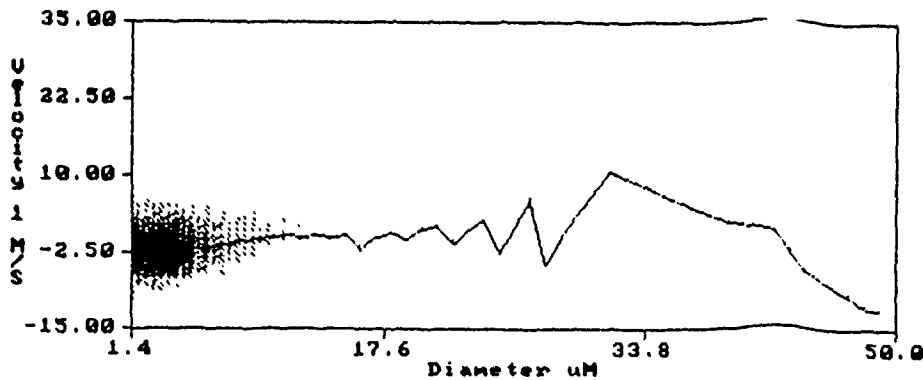
Figure 2b. A plot of the mean velocity as a function of droplet diameter.
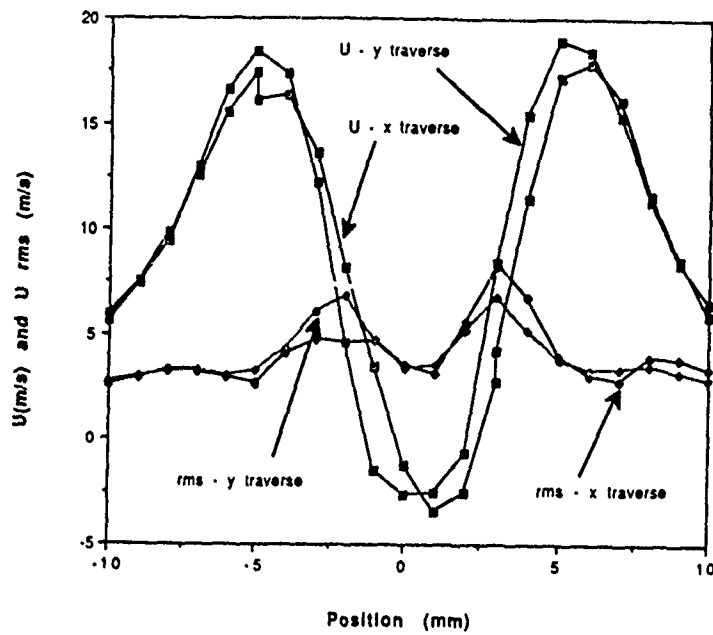
Figure 3a. Plot of mean velocity and rms values as a function of radial position. These data were taken 5mm from nozzle
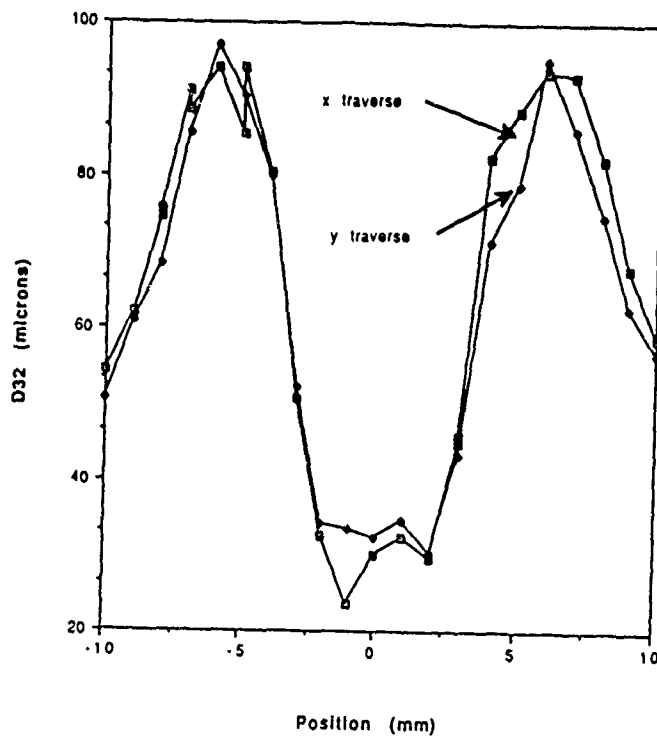
Figure 3b. Plots of the Sauter mean diameter (D_{32}) as a function of radial position.
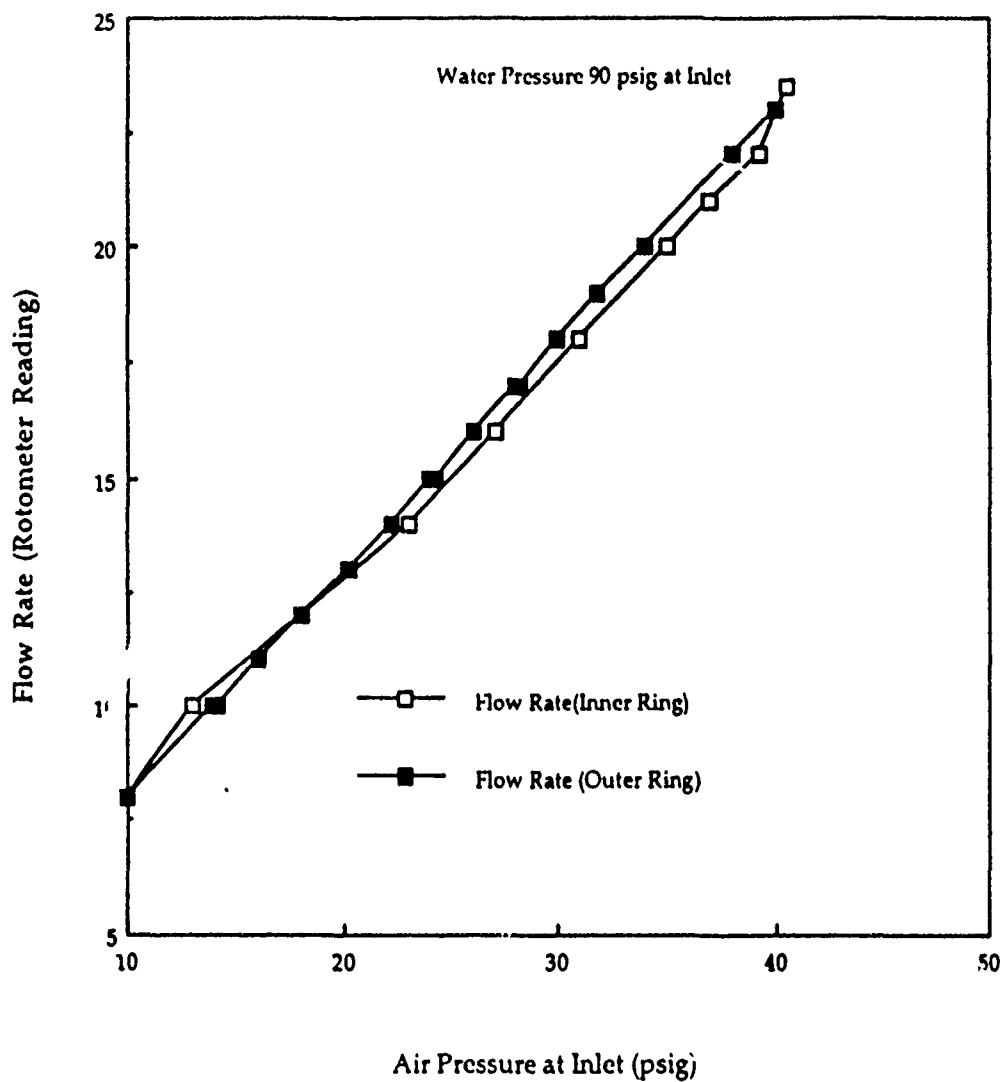
Figure 4. Plot showing the split of air between the inner ring and the outer ring. This plot indicates the air is evenly split between the two rings.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

<u>FINAL REPORT</u>

<u>Pattern Recognition: Machine vs. Man</u>

Prepared by: Thomas Abraham

Academic Rank: Instructor

Department: Natural Science & Mathematics

University: Saint Paul's College, Virginia

Research Location: USAF WRDC/AART-2

                                  Wright-Patterson AFB OH   45433-6543

USAF Researcher: Timothy D. Ross

Date: 27 July 90

Contract No: F49620-88-C-0053

PATTERN RECOGNITION: MACHINE VS. MAN

by

Thomas Abraham

ABSTRACT

Sixty-six images of varying degrees of pattern-ness were compiled and their AFD* ratings obtained. The respective human ratings were also gathered from 11 people. The data was consolidated and studied. Also, through the survey, we investigated to see whether there is a generally understood meaning for pattern-ness among the respondents.

* Ada Function Decomposition (program)

ACKNOWLEDGEMENTS

I wish to thank UES for giving me another great opportunity to tap my research potentials and to enable me to contribute my share in meeting the objectives of the AFOSR. Also, I would like to thank Les Lawrence and Paul Johnson for the part they played in realizing this activity.

John Jacobs and his crew of the AART-2 group were very cooperative in accommodating me and to share the facility with me. Dr Tim Ross deserves to be thanked for the outlining and the facilitating of the project. He should also be complimented for his publication which allowed me to measure pattern-ness of my images mechanically. Dr Tom Gearhart, another UES Fellow in our section, needs to be thanked for helping me obtain hard copies of my patterns and for his valuable counsel. Also, Lt. Tim Taylor is to be thanked for allowing me to use some of his patterns to add to the variety of my collection. Again, I like to extend my thanks to Mike Noviskey, Dave Gadd, Keith Graves, Mark Boeke, Michael Chabinye, Shannon Spittler, and Dr. Mike Breen for expending their time to rate the patterns. A special thanks goes to Kip Turner and Peggy Alltop for sharing their facilities with me many a time. Last, but not least I like to thank my home institution, Saint Paul's College, for the support rendered.

I.   INTRODUCTION:

Although this was my second opportunity to participate in this program,

I was still having apprehensions about the kind of work I will be doing,

as my exposure in research was not extensive. However, my meeting with

Dr. Tim Ross during my pre-summer visit enabled me to gather momentum.

My interests in the study of patterns greatly overlapped with those of

Dr. Ross.

The road to success was not quite easy. I had to overcome more than one

impediment to meet my objective. I launched a sojourn to computer graphics

as I never experimented on one before. With some effort, I managed to

produce patterns on the screen. Getting hard copies of the patterns so

obtained posed another threat to my efforts. Sympathizing in my

predicament, Dr. Tom Gearhart volunteered to help me get the desired hard

copies.

While recognizing the part played by others, I should also attribute my

success to my background education and my strong commitment to fulfilling

my obligations. This laboratory is now involved in Pattern Based Machine

Learning, and I wish to contribute my share in realizing this goal.

II.   OBJECTIVES OF THE RESEARCH EFFORT:

Our Objective is to find how closely the machine* ratings of certain

patterns are related to the corresponding human ratings of the same.

Also, we want to investigate the criteria for pattern-ness as evidenced by

people.

* Pattern Based Machine Learning (Final Technical Report)

  Published by Tim Ross et. al (1990)

III.

Pattern is a primitive concept experimented by mathematicians, philosophers, artists, and so forth. Often we think of patterns of numbers and objects. The very set of integers is achieved through pattern and order. As one mathematician (ref. 1) has said, "God created integers; everything else is the work of man". He must have meant that the set of integers is as patterned as the universe, a creation of God.

Identifying pattern-ness in objects or numbers may be called pattern recognition. People are good judges of patterns although there is some degree of latitude between one another. Of course the pattern-ness changes as the perspective is altered. The more we focus, the less patterned the object becomes.

Studies on pattern recognition produced considerable revolution in improving the respective disciplines. However, a general technique that is rigorous and mathematically oriented is still being experimented upon. Dr. Tim Ross, in his publication (ref. 2) makes an attempt to accomplish this by designing an algorithm for pattern recognition and refers it as 'recognizer design'.

Recognizer design is about designing machines that 'map large (possibly infinite) sets called 'measurements' into small (always finite) sets

called 'class-labels', when the dominant factor is the pattern-ness of this mapping. A mapping, of course, is a set of ordered pairs on a domain and a codomain such that every element in the domain has a unique paired element in the codomain as in a functional relationship. A class-label is also a straight forward concept which means simply the name of the class. A measurement, however, requires special explanation. It is, in the conventional sense, an intermediate internal state of the recognition machine and is considered an important part of the design process. Measurements are usually reckoned with lenses, microphones, and other transducers rather than with digital computers. In short, it is simply the set of recognizer inputs and are not under the designer's control. This measurement to class-label mapping concept led Dr. Ross to design his Function Decomposition algorithm.

Recognizer design is partitioned into two sub-problems and termed 'definition' and 'realization'. Definition is used to mean the problems involving knowledge acquisition, and realization is used to mean those involving hardware/execution. The generic term 'cost function' is being employed to reflect the performance and dollar cost considerations for the entire process including the cost considerations for errors and rejections due to knowledge acquisition or hardware/execution. Some of these costs may trade off against each other, and some may not. The net effect of all these constitute the cost function. Thus the recognizer design problem is translated in terms of cost function. It is this cost function that I was able to obtain through the Ada Function Decomposition program.

IV.

Two dimensional patterns of varying degrees of pattern-ness were generated by manipulating functions of six or eight variables. (Some sample images and their respective programs are included in the appendix.) 66 of them were compiled for the study. Upon obtaining their AFD ratings as a cost function, people were asked to rate them in terms of the pattern-ness on a scale 0 through 9. The arithmetic mean of these ratings from 11 people were found. After standardizing the AFD rating and the people rating to 3 significant figures, they were subjected to our study. (The consolidated table so obtained is included in the list of tables.) We have noticed that several pictures had one AFD rating but distinct people ratings. We took their mean and their standard error and represented them graphically showing 99.73% confidence level. Also we estimated the correlation coefficient of the regression line obtained from their scatter plots and found to be .78 approximately. Again, we classified the people ratings into classes of width 100 and recorded the corresponding mean AFD ratings in each class and plotted them against the mid values of the corresponding classes of human ratings. The correlation coefficient measured .96 here. (Both graphs are appended in the list of figures.) We also obtained the scatter plots of all the 66 pairs (graph attached) and their corresponding regression lines. The correlation coefficient was found to be .80.

Our experiment was also intended to obtaining a generally understood

meaning for pattern-ness. This was achieved through the questionnaire by

asking the evaluators to comment on their view of the term. (A copy of

the instructions to the evaluators is included in the appendix.)

V.   FINDINGS:

The wide angle between the regression lines is due to the distortion

caused by the gap between the two sets of ratings of low-patterned

pictures. Among the highly patterned ones, there is more consensus

between man and machine. The correlation coefficients we obtained is

fairly high. But for the abnormal cases, we would have obtained a better

figure. The human ratings are intuitive and spontaneous and hence have

considerable variations. The large number of patterns involved also

should have been responsible for the lack of precision. With all these

shortcomings, we noticed a steady trend of direct linear relationship

between the two sets of ratings as the correlation coefficient stayed in

the range .78 through .96.

It was observed from the responses to the questionnaire that the criteria

for pattern-ness centered around symmetry, repetition of the same shape in

an orderly fashion, and easy description. This is in agreement with the

author's perception on pattern-ness: simplicity. According to him,

"having to list all the elements (a brute table look-up) is the essence of

an unpatterred set." Symmetry implies simplicity because you can complete

the whole given its part. The same rule applies to repetition too. Easy

description also implies simplicity 86-9

VI.   RECOMMENDATIONS:

Our experiment is a sure indication that the AFD ratings are quite

reliable. However, there is still room for improvement. A rigorous

technique to measure pattern-ness of any image is still to be explored.

We hope that it will eventually be realized by the same token.

REFERENCES

1. Kronecker, Leopald (b. Dec. 7, 1823, Liegnitz, Prussia - d. Dec. 29, 1891, Berlin)

2. Pattern Based Machine Learning (Final Technical Report)
   by Tim Ross et. al (1990)

REGRESSION LINES

• AFD AGAINST HUMAN
○ HUMAN AGAINST AFD
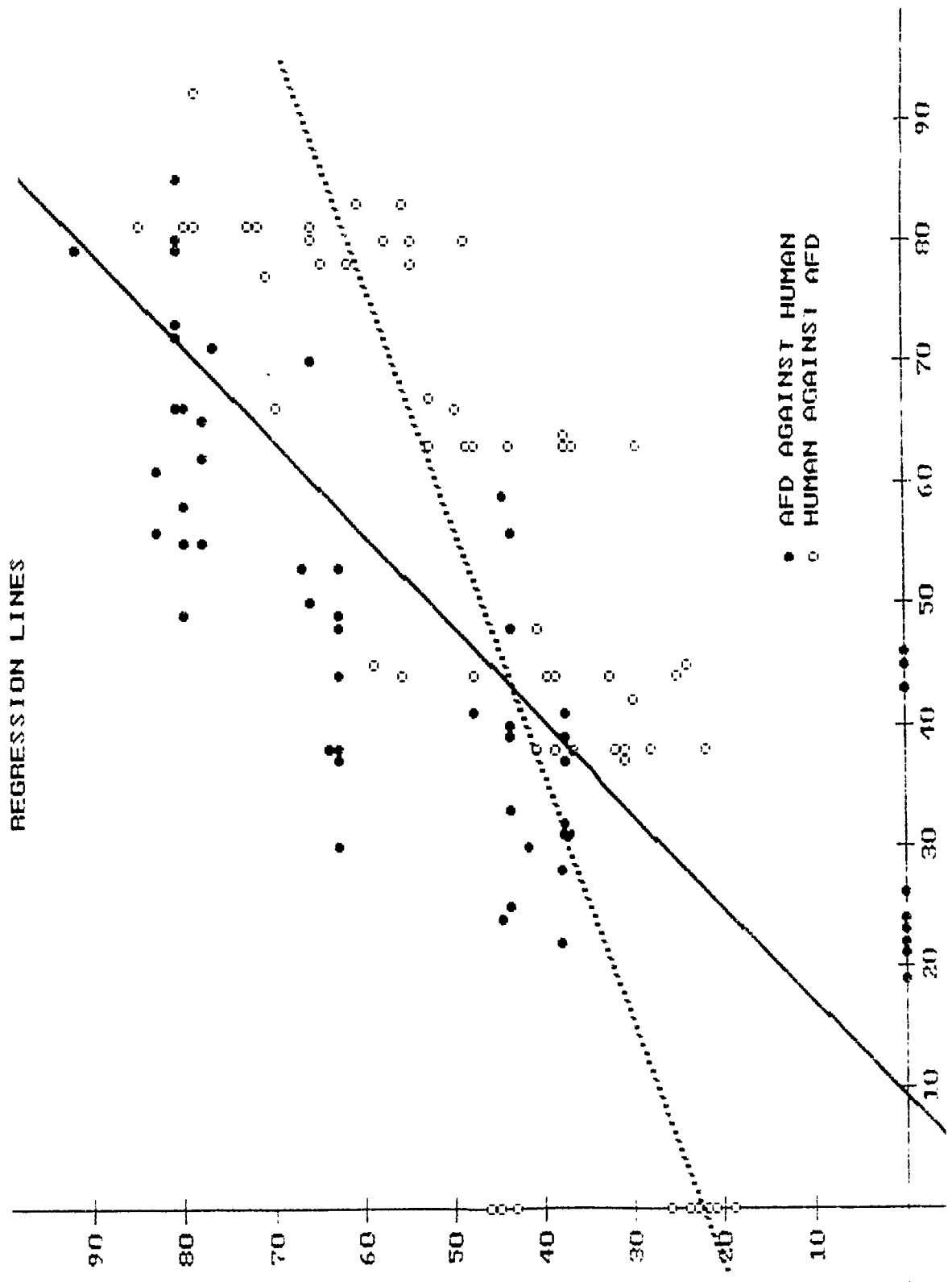
FIGURE 1
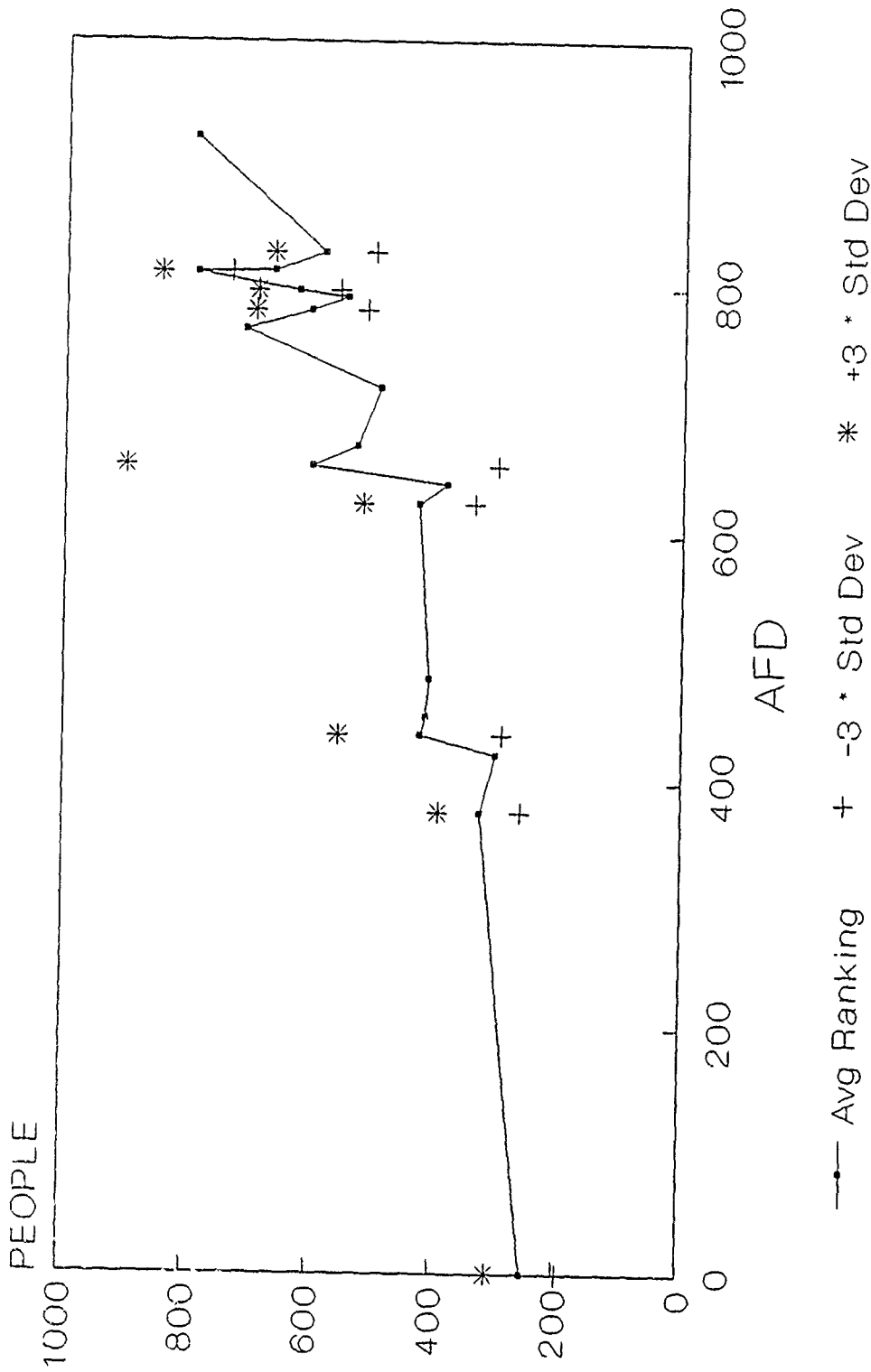
PATTERN-NESS RANKING EXPERIMENT

FIGURE 2

86-13

Tom Abraham 7/90

PATTERN-NESS RANKING EXPERIMENT

AFD

1000 800 600 400 200 0 -200

0 100 200 300 400 500 600 700 800 900 1000

PEOPLE

Avg Ranking   + +3 * Std Dev   * -3 * Std Dev

Tom Abraham 7/90
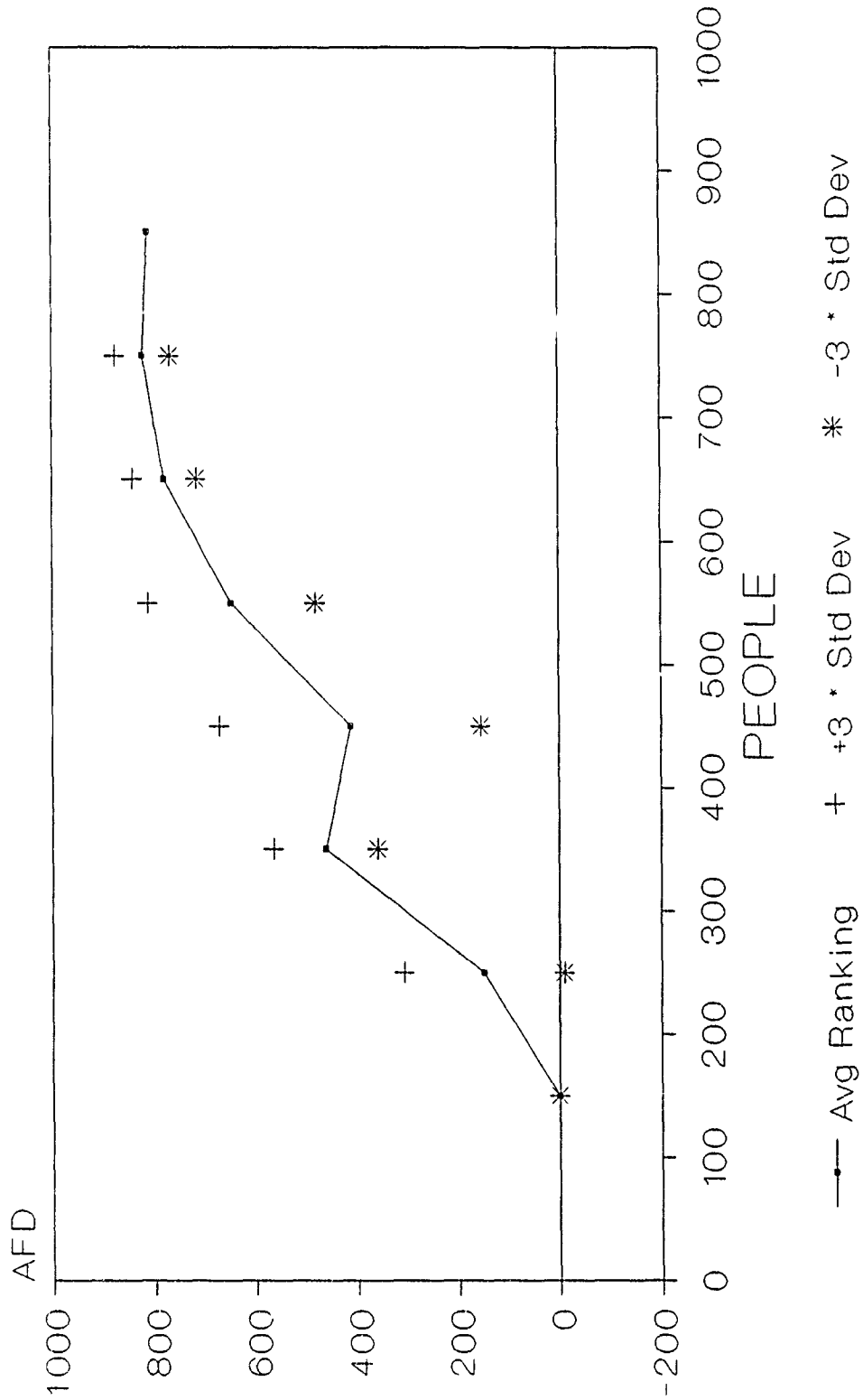
FIGURE 3

TABLE 1

Pattern-ness Ratings of Pictures
Approximated to 3 Sig. Figures

Picture	Human	AFD	Picture	Human	AFD	Picture	Human	AFD
01	373	625	23	709	766	45	300	625
02	236	000	24	591	453	46	227	000
03	236	453	25	664	813	47	245	438
04	445	000	26	455	000	48	491	719
05	718	812	27	227	000	49	227	000
06	382	641	28	382	625	50	845	812
07	282	375	29	482	438	51	700	656
08	555	438	30	800	812	52	618	781
09	255	000	31	527	625	53	482	625
10	427	000	32	655	797	54	582	797
11	655	797	33	218	375	55	491	625
12	309	375	34	545	792	56	409	375
13	800	812	35	727	812	57	409	484
14	555	438	36	400	438	58	318	375
15	209	000	37	300	422	59	791	922
16	500	656	38	227	000	60	373	375
17	545	781	39	609	828	61	645	781
18	218	000	40	527	672	62	227	000
19	791	812	41	845	812	63	236	000
20	391	375	42	209	000	64	191	000
21	391	438	43	436	625	65	309	375
22	218	000	44	327	438	66	555	828

APPENDIX MAY BE OBTAINED FROM

THE AUTHOR OR UES

1990 USAF-UES Summer Faculty Research Program/

Graduate Student Research Program

Sponsored by the

Air Force Office of Scientific Research

Conducted by the

Universal Energy Systems, Inc.

<u>Final Report</u>

<u>Some Results in Pattern-Based Machine Learning</u>

Prepared by:	Mike Breen, PhD
Academic Rank:	Assistant Professor
Department and	Mathematics Department
University:	Alfred University
Research Location:	WRDC/AART-2
	Wright-Patterson Air Force Base
	Ohio 45433
USAF Researcher:	Tim Ross, PhD
Date:	July 23, 1990
Contract No:	F49620-88-C-0053

Some Results in

Pattern-Based Machine Learning

by

Mike Breen

Abstract

The principle on which fire control hardware and software operate is that of a function. Our purpose is to find the fastest, most efficient form of the function possible. It is not feasible to have the function in (input,output)–form. We want to find the rule under which the function operates. Towards this end, we concentrate on machine–learning systems and function decomposition. This paper lists several results in each area seeking to establish a lower bound on training set size at which the machine–learning system is still performing well. There is no universal lower bound for each function and each machine–learning system, but we give instances where a convergence conjecture holds. Also, we give a proof of the main result in function decompostion.

Acknowledgements

I am of course extremely grateful to the Air Force Systems Command, the Air Force Office of Scientific Research, and to Universal Energy Systems, Inc. for this tremendous opportunity. I found this summer to be a very rewarding experience professionally.

As individuals, I would like to thank Dr. Tim Ross who gave me all the help and advice I needed as well as John Jacobs and Paul Johnson who were also very helpful. Many thanks to Dave Gadd and Devert Wicker who helped overcome the inflexibility of LaTeX and to Gary Lukens who generated the diagrams of Figure 1. All the people at AART-2 were very friendly and helpful.

I. INTRODUCTION:

In Avionics, both the amount of information ncecessary to operate effectively, and the speed with which that information must be processed and acted on to maintain high performance, require more than sophisticated hardware. What is required is a design to allow the hardware to perform. An algorithm to allow it to work its best.

The Systems Concept Group of the Mission Avionics Division is studying Pattern–Based Machine Learning in an attempt to have the machine learn as much as possible with a small amount of data. The idea is to allow the machine to train on the given data and find the pattern on this sample. Then the machine can apply that pattern to all the data. With this approach, a machine can perform its task with less training.

It is the belief of the Group that any function of interest will be patterned. Functions that most people are fam liar with always have a pattern in that they can be specified with a rule. In general, most functions are not patterned, however most — if not all — functions that are dealt with in real life are patterned.

Many of the functions discussed in machine–learning are Boolean functions. My research area is the area of Boolean Matrix Theory — a theory which has many applications. It can and has been used to analyze dominance relations and food web relations in the animal kingdom and to analyze cliques and the spread of information in human social structures. Familiarity with this theory contributed greatly to my understanding in this assignment.

II. OBJECTIVES OF THE RESEARCH EFFORT:

When most people speak of pattern theory, they speak of patterns present in individual objects. For example, they would like a machine to be able to recognize the letter "A". Our goal is to bring the idea of pattern to functions (which are correspondences between objects).

We would like our hardware and software to find the pattern in the assignment function. The more data the machine has, the better it will do. Yet we do not want to increase the data a hundred–fold if it causes only a 1/100 per cent increase in learning. In this regard we would like to know when it is not practical to continue training. During most of this research period, I worked on this idea. We wanted to see when a particular inequality was true and when it wasn't. Both experimental and theoretical work was necessary in this area.

At times we also worked in the area of function decomposition — also an investigation into the patternness of a function. We related this idea to the above one since function decomposition is one example of a machine learning system. We wanted to provide a different perspective on, and a different proof of, the Basic Decomposition Condition.

III. THE CONVERGENCE QUESTION

a. In (Ross,1989) we have the following situation. A function is given which we want our machine learning system (MLS) to learn. It is impossible to tell the MLS the whole function. However, we can give it small parts of the function, and allow it to train on that sample in an effort to learn the whole function. How big a sample should we give the MLS? This is the focus of this section.

b. We are given a function $f : X \longrightarrow Y, |X| = n$. Without loss of generality, assume $Y \subseteq R^m$. Denote a sample (called a training set) by T where $T \subseteq X$ and $|T| = k$. Let $A = \{$all functions learned on sets of size $k\}$. The symbols $T', k', A',$ denote the corresponding notions for larger training sets.

Define a measure of the difference between two functions as

$$D(f,g) = \sum_x < f(x) - g(x), f(x) - g(x) >$$

and define

$$D(f,A) = \sum_a D(f,a).$$

Let $e = \frac{D(f,A)}{|A|}$. We want to know at what point, if any, is it true that e can not be reduced significantly by training on larger sets. In practice, both k and k' are very small when compared to n.

In (Ross, 1989) it is shown that $D(.,A)$ is minimized by the function $\alpha = \frac{\sum_a a}{|A|}$. That is, the average of the a's minimizes the total distance. Thus $\frac{D(f,A)}{|A|} \geq \frac{D(\alpha,A)}{|A|}$. We call this last number d. Note that d gives us a lower bound on e. We use this to formulate the convergence question: Given f, is there a k such that for all A'

$$\frac{D(f,A)}{|A|} - \frac{D(\alpha,A)}{|A|} \leq \frac{D(f,A')}{|A'|} \leq \frac{D(f,A)}{|A|} + \frac{D(\alpha,A)}{|A|} ?$$

In this section, we investigate this question. The above inequality is often referred to as "the inequality". The inequality determined by the middle term and the sum on the right is called "the right side of the inequality" with a similar meaning for "the left side of the inequality."

c. Write a_T for a function in A trained on a set T. We call f monotone if whenever $\emptyset \subset T \subseteq T' \subseteq X$, then $D(a'_{T'}, f) \leq D(a_T, f)$.

Proposition 1 *If f is monotone, then $\frac{D(f,A')}{|A'|} \leq \frac{D(f,A)}{|A|}$ for all A and A'.*

The proof of this proposition is given by Tom Gearhart. It tells us that if f is monotone then the right half of the inequality will be true. In fact, there is no need for the d term.

Although monotonicity is a satisfying idea, we found few sophisticated learning systems which could be guaranteed to be monotone. So, without monotonicity, the d term is often needed for the right side of the inequality to hold.

d. Now let $f : X \longrightarrow \{0,1\}$. In this case we call f a Boolean function. For $x \in X$ define $\alpha_B(x)$ to have the value that a majority of the a's have at x (in case of a tie, choose 1). Note that an alternate way to find $\alpha_B(x)$ is to first find $\alpha(x)$ and then round. To make it easier, I will right α for α_B from now on, when speaking of Boolean functions. Define (for fixed x)

$$P_x = |\{a : a(x) = \alpha(x) \text{ and } a(x) = f(x)\}|, p_x = P_x/|A|;$$

$$Q_x = |\{a : a(x) = \alpha(x) \text{ and } a(x) \neq f(x)\}|, q_x = Q_x/|A|;$$

$$R_x = |\{a : a(x) \neq \alpha(x) \text{ and } a(x) = f(x)\}|, r_x = R_x/|A|;$$

$$S_x = |\{a : a(x) \neq \alpha(x) \text{ and } a(x) \neq f(x)\}|, s_x = S_x/|A|;$$

and $E = \{x : f(x) = \alpha(x)\}, U = \{x : f(x) \neq \alpha(x)\}$. Note that $D(\beta, A) \geq D(\alpha, A)$ for all Boolean β; and when $x \in E, q_x = r_x = 0$, when $x \in U, p_x = s_x = 0$. Primes appended to these symbols carry the corresponding notions to A'.

Define $p = \frac{\sum_{x \in E} p_x}{|E|}, q = \frac{\sum_{x \in U} q_x}{|U|}, r = \frac{\sum_{x \in U} q_x}{|U|}, s = \frac{\sum_{x \in E} s_x}{|E|}.$

Proposition 2 *If for all A' i) $\alpha' = \alpha$ and ii) $s' \leq s$; and if iii) $r \geq 1/4$; then the inequality holds for all A'.*

Proof Since $\alpha' = \alpha, E' = E$ and $U' = U$. The inequality becomes

$$|U|(q - r) \leq |U|q' + |E|s' \leq |U|(q + r) + 2|E|s$$

which simplifies to

$$|U|(q - r) \leq |U|q' + |E|s' \leq |U| + 2|E|s.$$

The right side is true because of assumption ii) and the fact that $q' \leq 1$. As for the left side, since $r \geq 1/4, q - r \leq 1/2$. Because of the definition of $\alpha, q' \geq 1/2$. Hence, the left side holds as well. The proposition is true.

e. Perhaps the first assumption in the last proposition is too harsh. Let us eliminate that requirement, for indeed the Boolean average may change as the training set size increases. Now the inequality becomes

$$|U|(q - r) \leq |U'|q' + |E'|s' \leq 2|E|s + |U|.$$

To focus on the important quantities, let $k = \frac{|E|}{|U|}$ (where $U \neq \emptyset$) and $\triangle = (|U| - |U'|)/|U|$. Note that $\triangle \leq 1$. Using these terms and the fact that $r = 1 - q$, we can write the inequality as

$$2q - 1 \leq (1 - \triangle)q' + (k + \triangle)s' \leq 2ks + 1.$$

As is often the case with any MLS, the right side of the inequality will be true under certain reasonable assumptions. We endeavor to show what proportion of the time the left side holds given that the right side holds.

We assume that both $\triangle \geq 0$ and $k \geq 1$. That is, we assume that as we take larger training sets, the Boolean average of these functions will be at least as close to f as the current Boolean average is. The latter assumption means that $|E| \geq |U|$.

We graph the lines in the q'-s' plane whose equations are $(1 - \triangle)q' + (k + \triangle)s' = 2q - 1$ and $(1 - \triangle)q' + (k + \triangle)s' = 2ks + 1$. Observe that the inequality is valid in the region bounded by the lines just mentioned (call them CD and AB, respectively) and the lines with equations $q' = 1/2, q' = 1, s' = 0$ and $s' = 1/2$. We determine the proportion we are interested in by finding the areas of the corresponding regions. The slope of both CD and AB is $\frac{\triangle - 1}{\triangle + k}$. Since $0 \leq \triangle \leq 1$ and $k > 0$, the slope of these lines is nonpositive and for large k, close to 0.

We are interested in the points of intersection of AB and CD with the rectangle. The s'-coordinate of the point of intersection of CD with $q' = 1/2$ is

$$\frac{(2q - 1) - ([1 - \triangle]/2)}{k + \triangle} \leq \frac{1 - ([1 - \triangle]/2)}{k + \triangle} \leq \frac{1 + \triangle}{2(1 + \triangle)} = 1/2.$$

The s'-coordinate of the point of intersection of AB with $q' = 1$ is $\frac{2ks + \triangle}{k + \triangle} \geq 0$. Therefore, the shape of the region under consideration can have one of nine forms. These forms are given in Figure 1 along with the quantities that determine those forms. The reader should not be deterred by some of the proportions in the following proposition as we will estimate them later.
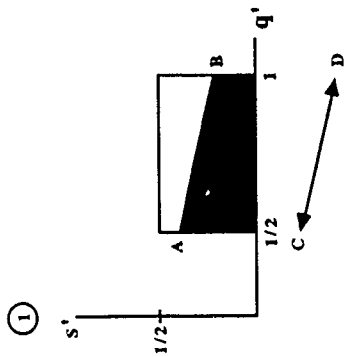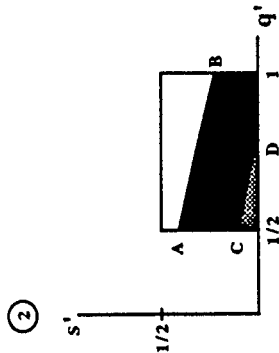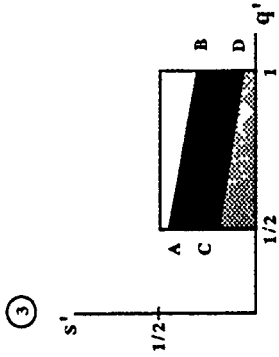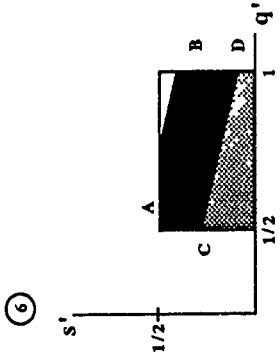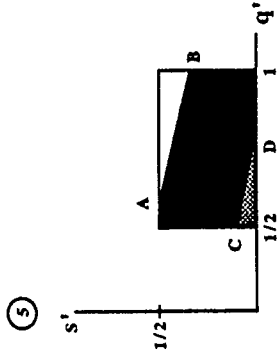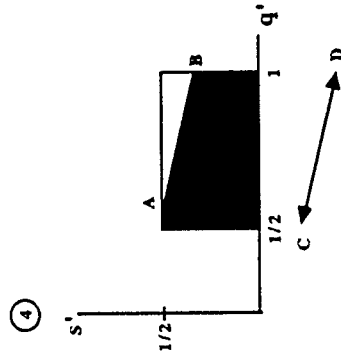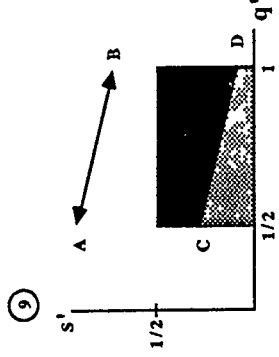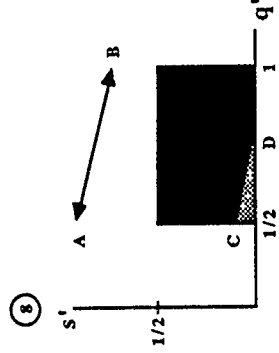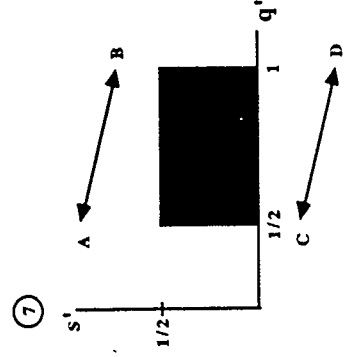
FIGURE 1

Proposition 3 *If* $\triangle \geq 0$ *and* $k \geq 1$, *then* $\dfrac{\text{the area of the region where the inequality holds}}{\text{the area of the region where the right inequality holds}}$ *is given below where the numbered proportion corresponds to the region number in Figure 1.*

1. 1 ,

2. $1 - \dfrac{(\triangle + 4q - 3)^2}{(1 - \triangle)(8ks + 3\triangle + 1)}$,

3. $\dfrac{8ks + 8(1 - q)}{8ks + 3\triangle + 1}$,

4. 1 ,

5. $1 - \dfrac{\triangle + 4q - 3)^2}{2(k + \triangle)(1 - \triangle) - (k - 4s - \triangle)^2}$,

6. $1 - \dfrac{(1 - \triangle)(3\triangle + 8q - 7)}{2(k + \triangle)(1 - \triangle) - (k - 4ks - \triangle)^2}$,

7. 1 ,

8. $1 - \dfrac{\triangle + 4q - 3)^2}{2(k + \triangle)(1 - \triangle)}$,

9. $\dfrac{2k - \triangle + 7 - 8q}{2(k + \triangle)}$.

Proof Proportions 1, 4, and 7 are obvious. Observe that in the remaining regions, the s'-coordinate of C is $\frac{(2q-1)-((1-\triangle)/2)}{k+\triangle}$. In all other regions where B is pictured, its s'-coordinate is $\frac{2ks+\triangle}{k+\triangle}$. In regions 2 and 3, the s'-coordinate of A is $\frac{2ks+1-((1-\triangle)/2)}{k+\triangle}$. In regions 5 and 6, A's q'-coordinate is $\frac{2ks+1-((k+\triangle)/2)}{1-\triangle}$. For D, when it is the q'-intercept (regions 2, 5, and 8) its q'-coordinate is $\frac{2q-1}{1-\triangle}$. In regions 3, 6, and 9, the s'-coordinate of D is $\frac{2q-2+\triangle}{k+\triangle}$.

The areas of the regions follow from these facts.

We now get a lower bound on the more cumbersome of these terms. In region 2, $\triangle \leq 2(1 - q)$ so,

$$1 - \frac{(\triangle + 4q - 3)^2}{(1 - \triangle)(8ks + 3\triangle + 1)} \geq 1 - \frac{(1 - \triangle)^2}{(1 - \triangle)(8ks + 3\triangle + 1)} = \frac{8ks + 4\triangle}{8ks + 3\triangle + 1}.$$

In region 5, the proportion is larger than $1 - \frac{(1-\triangle)^2}{2(k+\triangle)(1-\triangle)-(1-\triangle)^2} = \frac{2k+4\triangle-2}{2k+3\triangle-1}$ (since $k - 4ks < 1$). In region 6, again $k - 4ks < 1$, so the proportion given is at least as large as $\frac{2k-8q+6}{2k+3\triangle-1} \geq \frac{2k-2}{2k+2} = \frac{k-1}{k+1}$. Using the same fact as we did with region 5, we find that the proportion corresponding to region 8 is at least $\frac{2k+3\triangle-1}{2(k+\triangle)}$.

Summarizing this, we have

Proposition 4 *If $\triangle \geq 0$ and $k \geq 1$ then the proportions in the numbered regions are at least*

$\frac{8ks+4\triangle}{8ks+3\triangle+1}$ *(region 2),*

$\frac{2k+4\triangle-2}{2k+3\triangle-1}$ *(region 5),*

$\frac{k-1}{k+1}$ *(region 6), and*

$\frac{2k+3\triangle-1}{2(k+\triangle)}$ *(region 8).*

We believe that k should be the dominant term in these expressions. These results are intended to be a guide as to when much improvement is not likely (when the proportion is high), thus making further training undesirable.

f. Because of the wide range of possibilities for f and for the MLS, the inequality may or may not be true. In this section and the next we maintain the assumption that f is Boolean and look at two different MLS's. For the first MLS, the inequality holds. For the second, it doesn't.

Define the first MLS as follows. For $x \in X$, if $x \in T, a_T(x) = f(x)$; if $x \notin X, P(a_T(x) = f(x)) = \mathfrak{z}$. That is, we have an MLS which assigns function values to the learned functions in a random, independent way for elements not in the training set. We call this a random MLS.

Proposition 5 *Given a random MLS with the property that $P(a_T(x) = f(x)) = P(a_T(x) \neq f(x)) = 1/2$ for $x \notin T$, then $P(f = \alpha) > 1 - 1/n$ whenever $(k - 2)900^{1/k-3} \leq n \leq 10^{12}$ (where $k > 3$) or more simply $k \geq 13$ and $2k \leq n \leq 10^{12}$.*

Proof By the definition of α, the probability that $\alpha(x) = f(x)$ is the probability that $|\{a \in A : a(x) = f(x)\}| > |\{a \in A : a(x) \neq f(x)\}|$. Thus our desired probability is the probability that the number of functions which are trained at x plus the number of functions which are not trained at x but which are equal to $f(x)$ is greater than the number of functions which are not trained at x and which are not equal to $f(x)$. There are C(n,k) functions in A, C(n-1,k-1) of them are trained at x. Therefore, $P(\alpha(x) = f(x))$ is the probability that more than $\frac{n!-2k(n-1)!}{2(n-k)!k!}$ of the C(n-1,k) functions

not trained at x are equal to $f(x)$. Let F represent the random variable whose value is equal to the number of functions in A not trained at x which are equal to $f(x)$. Note that the values of F are binomially distributed with mean $C(n-1,k)/2$ and standard deviation $\sqrt{C(n-1,k)}/2$. Since $p = 1/2$, we can get a good approximation to the desired probability by using a normal distribution. The usual adjustment of the F-values by $1/2$ is negligible in this case.

So, $P(\alpha(x) = f(x)) \approx$

$$P\left(z > \frac{\frac{n!-2k(n-1)!}{2(n-k)!k!} - \frac{C(n-1,k)}{2}}{\frac{\sqrt{C(n-1,k)}}{2}}\right) = P\left(z > -\sqrt{\frac{kC(n-1,k-1)}{n-k}}\right).$$

If $n \geq (k-2)900^{1/(k-3)} + 1$, then using rough estimates, we get the above to be at least as large as $P(z > -30)$ which is at least $1 - 2.78 \times 10^{-86}$ (Smirnov,1965). Hence, we can be fairly certain that given $x, \alpha(x) = f(x)$. If $k \geq 13$ and $n \geq 2k$, then $n > 2(k-2) + 1 > (k-2)900^{1/(k-3)} + 1$ and our conclusion is satisfied.

Because of independence, $P(\alpha = f) = [P(\alpha(x) = f(x))]^n > [1 - 2.78 \times 10^{-86}]^n$. We can use a Taylor polynomial (for $g(y) = y^n$) to get an approximation. Let $\varepsilon = 2.78 \times 10^{-86}$. Then $(1 - \varepsilon)^n \approx 1 - n\varepsilon$, with an error of at most $\frac{n(n-1)\varepsilon^2}{2}$. By assumption, $n \leq 10^{12}$, so the error term is at most 3×10^{-88} and

$$P(\alpha = f) \geq 1 - 2.78 \times 10^{-14} - 3 \times 10^{-88} > 1 - 10^{-42} \geq 1 - \frac{1}{n}.$$

If we have a random MLS for which $P(a_T(x) = f(x)) > 1/2$ when $x \notin T$, then the mean number of learned functions equal to $f(x)$ will increase, and the standard deviation will decrease. So $P(\alpha = f)$ will be larger. This gives us the following corollary.

Corollary 6 *If a random MLS has the property that for* $x \notin T, P(a_T(x) = f(x)) >$ *$1/2$, then with n as in proposition 5, $P(\alpha = f) > 1 - \frac{1}{n}$.*

So if the function the MLS is attempting to learn is Boolean, and if the MLS is random doing at least as well as chance, then for almost all domains and training sets, we can be fairly certain that α and f are equal.

Once we have a domain size and training set sizes which satisfy the conditions of proposition 5, then $U' = U = \emptyset$. Therefore, the inequality reduces to $0 \leq s' \leq 2s$. When we neglect the d term on the right in the original inequality we get the stronger statement $0 \leq s' \leq s$. We now set about to find when these inequalities are true, naturally concentrating on the right side of each.

Proposition 7 *If the MLS is random and $P(a_T(x) = f(x)) = 1/2$ for $x \notin T$, then $P(s' \leq s) \approx 1$ for each A' where $k \leq k' \leq 2n$ and n satisifies the conditions of proposition 5.*

Proof If $k = k'$, the conclusion follows.

Otherwise, $\mu_{s'} = \frac{n-k'}{2n}, \mu_s = \frac{n-k}{2n}$. So, $\mu_{s'-s} = \mu_{s'} - \mu_s = \frac{k-k'}{2n}$. Also, $\sigma_{s'}^2 = \frac{n-k'}{4n^2 C(n,k')}, \sigma_s^2 = \frac{n-k}{4n^2 C(n,k)}$. Hence, $\sigma_{s'-s}^2 = \sigma_{s'}^2 + \sigma_s^2 = \frac{n-k'}{4n^2 C(n,k')} + \frac{n-k}{4n^2 C(n,k)}$ and $P(s' \leq s) = P(s' - s \leq 0) >$

$$P\left(z \leq \frac{0 - \frac{k'-k}{2n}}{2\sqrt{\frac{n-k'}{n^2 C(n,k')} + \frac{n-k}{n^2 C(n,k)}}}\right) = P\left(z \leq \frac{k'-k}{\sqrt{\frac{n-k'}{C(n,k')} + \frac{n-k}{C(n,k)}}}\right)$$

(the strict greater-than sign arises because we are calculating a smaller area than the one we would calculate usually). Since $n \geq 2k$ and $k \geq 13 > 5, \frac{n-k}{C(n,k)} \leq 1/32$. Because $k < k' \leq n/2, C(n,k) < C(n,k')$ and the entire denominator which contains these two terms is at most $1/4$. Thus, $P(s' \leq s) > P(z \leq \frac{k'-k}{1/4} \geq P(z \leq 4) \approx 1$.

If for each individual x, the probability that each trained function agrees with $f(x)$ when x is not in the function's training set is more than $1/2$, then the z-score above would increase, and the stronger version of the inequality would again be true.

Corollary 8 *With a random MLS for which $P(a_T(x) = f(x)) \geq 1/2$, we have that $P(s' \leq s) \approx 1$ for each A' where $k \leq k' \leq n/2$ and n satisfies the condition in proposition 5.*

We summarize these results below.

Theorem 9 *Suppose*

1. $f : X \longrightarrow \{0, 1\}$,

2. *our MLS is random,*

3. *for each $x \notin T$ and for each $a_T, P(a_T(x) = f(x)) \geq 1/2$,*

4. $|T| \geq 13$ *and* $2k \leq n \leq 10^{42}$;

then

$$\frac{D(f, A)}{|A|} - \frac{D(\alpha, A)}{|A|} \leq \frac{D(f, A')}{|A'|} \leq \frac{D(f, A)}{|A|}$$

for all A' where $|T'| \leq n/2$.

Corollary 10 *The inequality in the convergence question is true under the same assumptions that appear in the theorem.*

Proof Obvious.

g. Let us now look at another MLS for a Boolean function. Let m be the function in A such that $D(m, f) \leq D(a, f)$ for all $a \in A$. That is, m is the closest function in A to f — sometimes called the last best guess. Let our learning system learn in this manner. Given $T' \subseteq X$ and $a' \in A'$,

$$a'(x) = \left\{ \begin{array}{ll} f(x) & \text{if } x \in T' \\ m(x) & \text{if } x \notin T' \end{array} \right\}.$$

This MLS is monotone by definition, hence the right side of the inequality holds because of proposition 1. The left side is more difficult — it need not hold. First note that, at a given x, calculating the real average of the a'-values at x, gives us

$$\alpha'(x) = \frac{C(n - 1, k' - 1)f(x) + [C(n, k') - C(n - 1, k' - 1)]m(x)}{C(n, k')} =$$

$\frac{k'f(x)+(n-k')m(x)}{n}$. Since we assume that $n > 2k'$, we see that $\alpha'(x) = m(x)$ for all x. So for $x \in U, Q_r = |\{a \in A : x \notin T_a\}| = C(n - 1, k)$ (where T_a means the training set

that a is trained on), $R_x = |\{a \in A : x \in T_a\}| = C(n-1, k-1)$. For all $x, S_x = 0$. The left side of the inequality becomes

$$\sum_{x \in U} \frac{C(n-1, k) - C(n-1, k-1)}{C(n, k)} \leq \sum_{x \in U'} \frac{C(n-1, k')}{C(n, k')}$$

which is equivalent to $|U|(n - 2k) \leq |U'|(n - k')$.

If our last training was done on sets of size $k, |U'| = |U| - k'$ (assuming the last best guess misses f by more than k' points). The inequality we are interested in becomes $|U|(n - 2k) \leq (|U| - k')(n - k')$ or $k'(n - k') \leq |U|(2k - k')$. We regard k and k' as being of the same magnitude, so that the $n - k'$ term is dominant and this inequality fails. If we train on all sets of size $k + 1, k + 2, \ldots, k'$, then $|U'|$ will be even smaller and again the left side of the inequality fails.

IV. FUNCTION DECOMPOSITION

a. To connect this area with the last, we now make note of the fact that we wrote a Turbo Pascal program which acts on our function decomposition database to gather information about the truth of the convergence question for that learning system. It is just now being run. The program calculates both the real average and the Boolean average, compensating for any undefined points and finds d for each average. Dr. Tim Ross wrote most of this program.

We also point out that of all the MLS's we investigated, the ADA function decomposition algorithm is the top performer when $f \colon \mathbb{N} \longrightarrow \{0, 1\}$. We now look at function decomposition.

b. Our work here was done first in the mechanics of the ADA function decomposition algorithm. In order to learn a function (f), this algorithm chooses the least complex function which agrees with the original function on the training set. We wanted to know the expected size of the set of functions which were both no more complex than f and agreed with f on the training set.

If $|X| = n$, and $G = \{g | g : X \longrightarrow \{0, 1\}\}$, then $|G| = 2^{(2^n)}$. If the training set $T \subseteq X$ has size k and $I = \{g \in G : g(x) = f(x) \text{ for all } x \in T\}$, then $|I| = 2^{(2^n - k)}$. For a randomly chosen subset of $G, H, P(H$ contains exactly r elements of

$I) = \frac{C(|H|,r)C(n-|H|,|H|-r)}{C(n,|H|)}$. The random variable r has a hypergeometric distribution. Therefore $E(I \cap H) = \frac{|I||H|}{|G|}$. In our case, $E(I \cap H) = \frac{2^{(2^n-k)}|H|}{2^{(2^n)}} = \frac{|H|}{2^k}$.

c. Our idea of patternness of a function is based on its complexity cost — the lower the cost, the higher the patternness. The cost is related to how the function can be broken into simpler pieces. That is, how it decomposes. The principal result here is the Basic Decomposition Condition. To prove it in the most general case, we must define a relation on the columns of the function's partition matrix. The definition of that relation follows.

To modify the semantics of the problem, we define a total function $\hat{f} : X_1 \times X_2 \times \ldots \times X_n \to Y \cup \{u\}$ such that $\hat{f}(x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n)$ when f is defined and $\hat{f}(x_1, x_2, \ldots, x_n) = u$ when f is undefined.

A "column" $(C_{v_1 v_2})$ for some fixed v_1 and v_2 is defined as the sequence: $C_{v_1 v_2} = (f(v_1, v_2, b(1)), f(v_1, v_2, b(2)), f(v_1, v_2, b(3)), \ldots, f(v_1, v_2, b(|V_3|)))$ where b is a bijection from V_3 into $\{1, 2, \ldots, |V_3|\}$. Columns form a set of columns S_{v_2} for a fixed $v_2, S_{v_2} = \{C_{v_1 v_2} | v_1 \in V_1\}$. When f is total, the column multiplicity (ν) is the maximum over v_2 in V_2 of $\{[S_{v2}]\}$; however, when f is not total, we need an extra step.

Call two columns compatible if the only coordinates in which they differ are those where either is undefined. Consider $\hat{S}_{v_2} = \{C_{v_1 v_2} | f(v_1, v_2, v_3)$ is defined for each $v_1, v_3\}$. If $\hat{S}_{v_2}$ is empty, we can go to later steps. If not, define $\hat{R}_{v_2}$ on $\hat{S}_{v_2}$ using prefix notation by $\hat{R}_{v_2}(C_{v_1 v_2}, C_{v_{1'} v_2})$ if $f(v_1, v_2, v_3) = f(v_{1'}, v_2, v_3)$ for all $v_3 \in V_3$. This is an equivalence relation on $\hat{S}_{v_2}$. Enumerate the resulting equivalence classes $E_1, E_2, \ldots, E_{p_{v_2}}$ calling representatives $e_1, e_2, \ldots, e_{p_{v_2}}$. Also enumerate the elements of the set $S_{v_2} \backslash \hat{S}_{v_2}$: $C_1, C_2, \ldots, C_{m_{v_2}}$. Choose the first class such that C_1 is compatible with the representative of that class and adjoin C_1 to that class. If no such class is found, C_1 will belong to its own class $E_{p_{v_2}+1}$. If C_1 creates a new class, it is the representative - called $e^1_{p_{v_2}+1}$. Otherwise, define the new representative of the class that C_1 is in, to be e^1_k where $(e^1_k)_m$ is the value of $(C_1)_m$ or $(e_k)_m$ if either or both are defined (if both, they must be equal), and $(e^1_k)_m$ is undefined if both $(C_1)_m$ and $(e_k)_m$ are undefined. Here, $m = 1$ to $|V_3|$ and m stands for the "coordinate" of the

column vectors. Representatives of other classes remain the same but are denoted individually as e_t^1. Continue in this manner with the other elements of $S_{v_2} \backslash \hat{S}_{v_2}$. That is, if a column C_a is not compatible with any of the existing representatives, it will create a new class whose number is one more than the number of the last class: $l_{v_2} + 1$. It will be the representative, denoted $e_{l_{v_2}+1}^a$. If a column C_a is compatible with an existing representative, we choose the first occurrence of this, adjoin C_a to that class and proceed as with C_1 except in the above steps (when we dealt with C_1) replace e_k with $e_k^{(a-1)}$, e_k^1 with e_k^a, and e_t^1 with e_t^a. Finally, given v_2, we have the set of classes $\mathbf{E_{v_2}} = \{E_{i_{v_2}} | 1 \leq i_{v_2} \leq \nu_{v_2}\}$ which partitions the set of columns. The classes' representatives are $(e^{m_{v_2}})_1, (e^{m_{v_2}})_2, \dots, (e^{m_{v_2}})_{\nu_{v_2}}$. Call the equivalence relation determined by this partition of S_{v_2}, R_{v_2}.

Now we define $\nu_{v_2} = |E_{v_2}|$. When V_2 is empty, there is only one E_{v_2} since v_2 can not occur in the expression. Finally, we define ν as the maximum over all v_2 in V_2 of ν_{v_2}. This definition relies only on elementary set theory for background.

d. Now that the relation is defined we use it to prove the following theorem. The theorem is not new, but the proof is.

Theorem 11 *The Basic Decomposition Condition: Let $f : X_1 \times X_2 \times \dots \times X_n \longrightarrow Y$ be a partial function where each X_t is finite. Let V_1, V_2, V_3, be a partition of the domain. There exists a set V and functions $\phi : V_1 \times V_2 \longrightarrow Z$ and $F : Z \times V_2 \times V_3$ such that when f is defined $f(v_1, v_2, v_3) = F(\phi(v_1, v_2), v_2, v_3)$ if and only if $\nu \leq |Z|$.*

Proof First show that $\nu \leq |Z|$ implies the desired existence.

Step 1. Define $\phi : V_1 \times V_2 \to Z$ by $\phi(v_1, v_2) = i$ if $C_{v_1 v_2} \in E_{i_{v_2}}$.

Step 2. Define $F : Z \times V_2 \times V_3 \to Y$ as follows:

i. if $z = i$ for some i, then $F(i, v_2, v_3) = (e_{i_{v_2}}^{j_{v_2}})_{b(v_3)}$ where b is the bijection from V_3 into $\{1, 2, \dots, |V_3|\}$,

ii. otherwise, $F(z, v_2, v_3)$ is undefined.

Once in place, defined coordinates of the representatives of the equivalence classes do not change, so when defined $f(v_1, v_2, v_3) = (e_{\phi(v_1, v_2)}^{j_{v_2}})_{b(v_3)} = F(\phi(v_1, v_2), v_2, v_3)$.

Now we prove that the existence of functions $f : V_1 \times V_2 \to Z$ and $F : Z \times V_2 \times V_3 \to Y$ such that whenever $f(v_1, v_2, v_3)$ is defined, $f(v_1, v_2, v_3) = F(f(v_1, v_2), v_2, v_3)$ implies that $\nu \le |Z|$. First observe that $\nu \le |Z|$ is logically equivalent to $\nu_{v_2} \le |Z|$ for all $v_2 \in V_2$. Assume to the contrary that there exists $v_2 \in V_2$ such that $\nu_{v_2} > |Z|$.

1) We can assume that equivalent columns correspond to ordered pairs which have the same inverse image under ϕ. When $C_{v_1 v_2}$ and $C_{v_{1'} v_2}$ are R_{v_2}-related, then when defined $f(v_1, v_2, v_3) = f(v_{1'}, v_2, v_3)$ for each v_3. Thus, $F(\phi(v_1, v_2), v_2, v_3) = F(\phi(v_{1'}, v_2), v_2, v_3)$. No harm is done to the relationship between f and F and the range of ϕ is no larger than before. Hence, we assume that the inverse image of an element of Z contains ordered pairs which correspond to an equivalence class (under R_{v_2}) of columns.

2) If $\nu_{v_2} > |Z|$, then there must be two non-equivalent columns whose corresponding ordered pairs have the same image under ϕ. These columns can not come from $\hat{S}_{v_2}$ since R_{v_2} is the equality relation there because if two columns from $\hat{S}_{v_2}$ are not equivalent, there must exist a v_3 such that $f(v_1, v_2, v_3) = f(v_{1'}, v_2, v_3)$. Hence, $f(v_1, v_2, v_3) = F(\phi(v_1, v_2), v_2, v_3) = F(\phi(v_{1'}, v_2), v_2, v_3) = f(v_{1'}, v_2, v_3) \ne f(v_1, v_2, v_3)$, a contradiction.

3) For $S_{v_2} \setminus \hat{S}_{v_2}$, use the ordering developed in defining R_{v_2}. Choose the first element of $S_{v_2} \setminus \hat{S}_{v_2}$, C_j, such that the number of equivalence classes in $\hat{S}_{v_2} \cup \{C_1, \dots, C_j\}$ is larger than $|\phi(\hat{S}_{v_2} \cup \{C_1, \dots, C_j\})|$. If no such element exists, $\nu_{v_2} \le |Z|$. Because of how C_j was chosen, C_j must create its own class. That is, C_j is not compatible with any representative of the equivalence classes at that time. We can write $C_{v_1 v_2}$ for C_j and see that there exists $C_{v_{1'} v_2} \in \hat{S}_{v_2} \cup \{C_1, \dots, C_j\}$ such that $\phi(v_1, v_2) = \phi(v_{1'}, v_2)$. Call the representative of the class of $C_{v_{1'} v_2}$, e^a. We know that $C_{v_1 v_2}$ is not compatible with e^a. Therefore, there is a $C_{v_{1''} v_2}$ currently in the class with $C_{v_{1'} v_2}$ such that for some $v_3, f(v_{1''}, v_2, v_3)$ and $f(v_1, v_2, v_3)$ are defined and unequal. This last statement follows from the fact that a column is compatible with a representative of a class if and only if it is compatible with each element in the class. So we have $R_{v_2}(C_{v_{1''} v_2}, C_{v_{1'} v_2})$ which implies that $\phi(v_{1''}, v_2) = \phi(v_{1'}, v_2) = \phi(v_1, v_2)$ and that $F(\phi(v_{1''}, v_2), v_2, v_3) =$

$F(\phi(v_1, v_2), v_2, v_3)$ for each v_3. Yet for some $v_3, f(v_{1''}, v_2, v_3) \neq f(v_1, v_2, v_3)$ and both are defined. As before, this contradicts the assumption about f and F. Hence, no such non-equivalent columns exist. The theorem is proven.

V. RECOMMENDATIONS:

a. The convergence question is a very broad one. We have seen that it can not always be answered in the affirmative. Because machine-learning is an imprecise area now, no ground rules are given for an MLS. In order to be able to frame the question in a more specific way, such ground rules are needed. It may just be a matter of time before such rules are given.

b. We have given some results on machine-learning when f is a Boolean function. Some complicated functions can be reduced to the Boolean case, but others can't. Investigation into other types of functions is a good idea.

c. Function decomposition is our criterion for patternness. How do other experts in the field feel about this criterion? Some effort should be given to sounding out those experts whether by publishing our results and waiting for feedback, or by attending conferences and meetings. An encouraging sign that function decomposition is a valid criterion for patternness is the behavior of the ADA function decomposition algorithm (FDA) in the area of machine-learning. It may be 'he paramount practical justificiation for the group's definition of a patterned function.

d. Why does FDA work so well? Much more sophisticated MLS's (neural nets, for example) are outperformed by FDA. Research should be done into the reasons behind the success of FDA. Such research could unlock secrets in both areas: machine learning and function decomposition.

REFERENCES

1. Ross, Timothy D. "A Convergence Method for Evaluating Whether or not the Training Set for a Machine Learning System Contains an Appropriate Number of Samples". December, 1989.

2. Smirnov, N.V., Tables of the Normal Probability Integral, the Normal Density and its Normalized Derivatives, New York, The Macmillan Company, 1965.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

<u>FINAL REPORT</u>

<u>PROBABILISTIC IR EVIDENCE ACCUMULATION</u>

Prepared by:	Dr. R. H. Cofer and Jim Perry
Academic Rank:	Associate Professor and Research Assistant
Department and	Electrical Engineering and Computer Engineering
University:	Florida Institute of Technology
Research Location:	WPAFB, WRDC/AARA, ATR Branch
USAF Researcher:	Jim Leonard
Date:	September 22, 1990
Contract No:	F49620-88-C-0053

Probabilistic IR Evidence Accumulation

by

Dr. R. H. Cofer and Jim Perry

ABSTRACT

The work reported here is a continued outgrowth of Bayesian Target Recognition research started in the 1989 Summer Faculty Research Program. During the current 1990 research effort, the emphasis has been on probabilistic evidence accumulation in the IR ATR problem. Two important and fundamental types of probabilities were found: underlying target temperatures, and spatial homogeneity of target temperatures. The first is important for target to decoy discrimination, while the second can overcome unavoidable lack of the target's thermal history. Correctly used, these two probabilities will result in overall consistency of IR evidence accumulation. Also shown is the general robustness of probabilistic evidence accumulation to practical considerations of uncertainty, ignorance, and functional approximation.

Acknowledgments

1.0 INTRODUCTION

The central problem in ATR research is how to accumulate modeling and imagery evidence in a consistent and meaningful form. Earlier research by Dr. Cofer has abstractly showed that a probabilistic approach is both consistent and capable of the highest recognition performance - provided that the required conditional probabilities can be found.

Two types of probabilities are required in a complete probabilistic evidence accumulation process:

- The probability that a specific region in the image could have arisen from the target, and

- The probability that the same region could have arisen from non target background.

The first type of probability can be immediately utilized in a wide variety of ATR designs and is thus of more immediate interest to AARA. Last summer's research showed methodology of how to obtain this probability for LADAR imagery. This summer's research shows methodology of how to obtain this probability for IR imagery, even under ignorance of the target's thermal environment.

2.0 OBJECTIVES OF THE RESEARCH EFFORT

For Dr Cofer, the goals of this research effort were to: extend the theory of probabilistic evidence accumulation, develop probabilistic modeling approaches to the IR target problem, and supervise the efforts of his research assistant, Jim Perry. For Jim Perry, the goals of this research effort were to: convert an AARA IR modeling program written for a personal computer to the considerably higher performance UNIX / C / SunView environment, learn the rudiments of IR target modeling in preparation for Master's research, and to participate in an advanced research team.

3.0 IR MODELING AND EVIDENCE ACCUMULATION

As planned, the majority of the summer's research effort went into the IR target modeling and evidence accumulation arena. An IR target imagery base was placed on the SUN

workstation for further consideration throughout the upcoming academic year. Dr. Cofer and Jim Perry greatly expanded their knowledge of IR target modeling in preparation for future continued interaction with Lori Westerkamp and Bill Foley. Jim Perry successfully converted a non-trivial IR modeling problem from the BASIC IBM AT/EGA environment to the much more powerful C SUN SPARC workstation SunView environment for a 20:1 improvement in speed, graphic clarity, and expandable problem size. Dr. Cofer successfully determined the fundamentals of splitting the probabilistic IR evidence accumulation problem along a thermal history / temperature homogeneity cleavage. The research methods used and results obtained are further discussed below.

Jim Perry's coding is documented in the final program documentation left at AARA. Dr. Cofer's results are further documented in a 40 viewgraph Probabilistic Evidence Accumulation package also left at AARA.

3.1 Physical IR Target Modeling

The physics of IR target modeling, Figure 1, has been repeatedly investigated over the years, but not with the view of developing ATR evidence accumulation probabilities. Thus several points are reasonable to note as they have a bearing on the development of the required probabilities. The IR image is formed as a result of seeing the temperatures, T's, of the target. Much if not most of the information of the target's IR appearance derives from time varying heat flows, q_i's, to and from the target's environment. If these flows were to go to zero, the target would become constant temperatured throughout and thus detail within its silhouette would be difficult to detect. Fortunately, heat flows within and outside the target are non-deterministic making the target's surface temperatures visible but, unfortunately, hard to analyze. This non-deterministic behavior then is the genesis of the need to place a probabilistic flavor to the modeling of the IR target, its temperatures, and heat flows.

Physically, IR modeling inevitably begins with the equation

$$\frac{dT}{dt} = \frac{\partial^2 T}{\partial z^2} + \sum_i \frac{\partial q_i}{\partial z} \qquad \text{(EQ 1)}$$

which continuously relates target temperatures and heat flows across space and time. Eq. 1
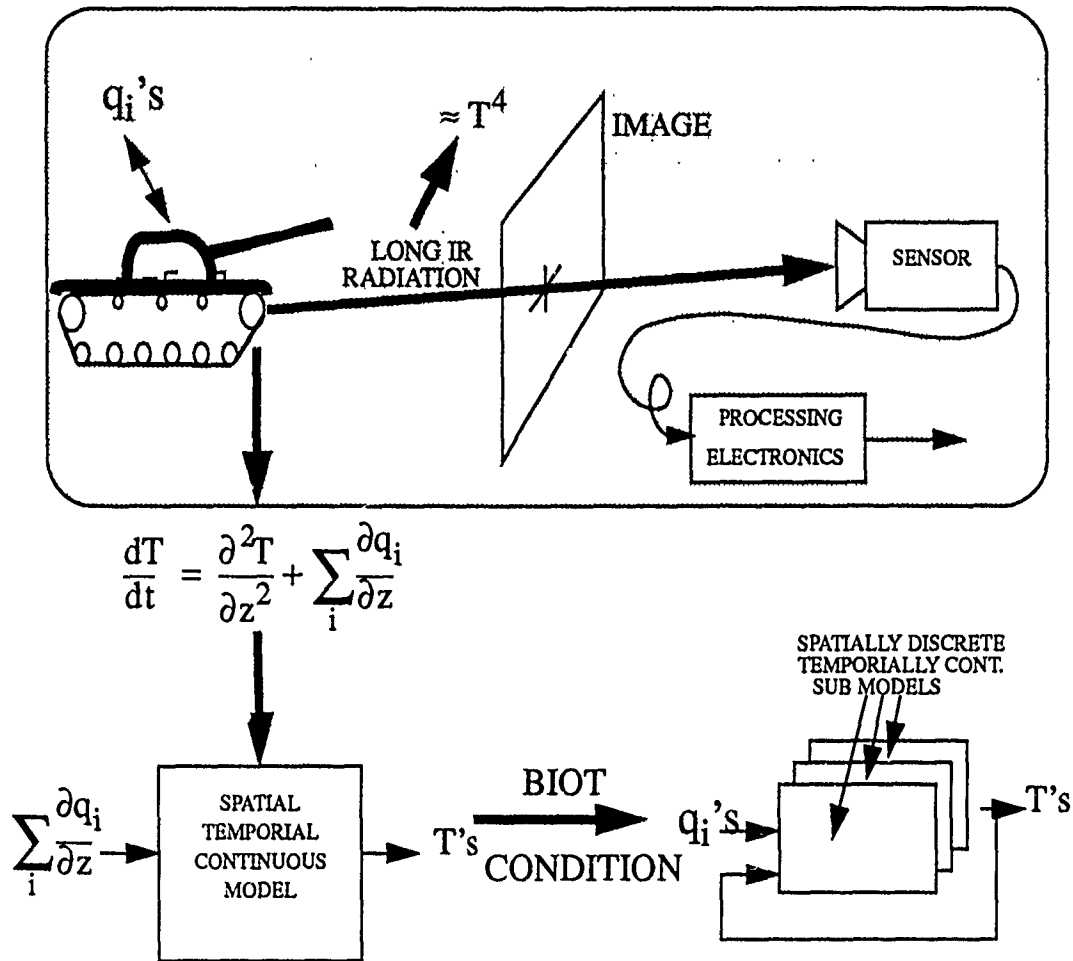
Figure 1. The IR Physical Modeling Problem

thus leads directly to the resulting spatial temporal model of Figure 1 lower left. Although such a model is quite general, it provides few easily accessible handles with which to attack the probabilistic IR problem.

3.2 The BIOT Condition in IR imagery

One of the best methods of obtaining clues to the nature of data is to examine it with an open mind. While looking at a chronological IR image history of a stationary M60 tank,

Dr. Cofer became struck by the fact that he could see the tank wheels, hubs, webs, tracks and body. Even more striking was the fact that various tank components were appearing and disappearing from visibility as whole entities. These observations lead to further study of heat transfer in solids. This study resurfaced the BIOT condition - objects (or subobjects) which have a high internal to external heat flow rates can be essentially considered constant temperatured across their mass. More interestingly, BIOT conditioning was quite evidently also being seen in the IR image. Wheels, webs, tracks and body of the tank all had high internal conductivity relative to their external heat flows. Thus each had relatively constant although non-deterministic temperatures across their mass. When their non-deterministic temperatures were sufficiently different, parts could be distinguished one with respect to another. When the non-deterministic temperatures were not sufficiently different, the parts remained visually indistinguishable.

Recognizing that BIOT isolated parts exist within a target allows the model of lower left Figure 1 to be reduced to the more spatially discrete model at the lower right of Figure 1. Although investigators have often discretized target thermal models, this current research specifically notes that the BIOT discretization also shows up in the IR imagery.

3.3 Target Reduction to an Equivalent Circuit

Further study of past IR modeling efforts revealed that the thermal heat transfer properties of metallic solids, on either a macro discrete or micro infinitesimal level, can be modeled as an electrical circuit composed of resistors and capacitors. The resistors serve as flow paths, the capacitors serve as heat sinks within the solid, electrical currents into and out of the circuit serve as environmental heat flows to and from the solid and finally electrical currents within the circuit serve as heat flows within the solid. As a result, the model of lower right Figure 1 and upper right Figure 2 can be replaced by a resistor capacitor circuit, represented graphically in the upper right of Figure 2. As a result the pixels of the image from an IR sensor are equivalent to voltage readings taken from points in a resistor

capacitor circuit being driven by environmental heat flows - the q_i's at the top of the circuit[1] of Figure 2.
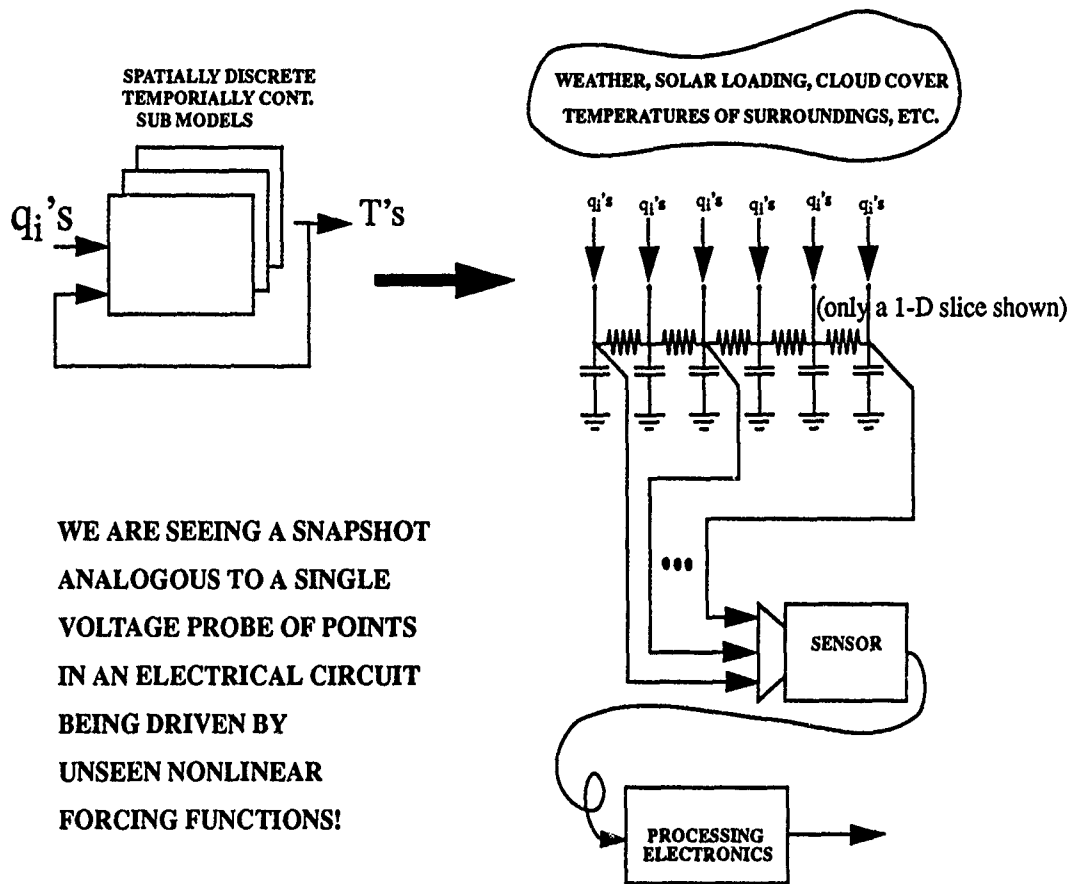
Figure 2. An Equivalent Model of What Is Being Sensed

3.4 Probabilistic ATR Requirements

To understand the precise nature of the probabilistic characterization needed in the IR ATR problem, consider Figure 3. Upon receipt of the incoming IR image, R_{IR}, the recognizer must determine the probabilities, $P(R_{IR} \mid I, H_x)$ and $P(R_{IR} \mid I, H_y)$, of R_{IR} conditioned on

---

1. This problem should not be confused with a similar problem that has long been addressed in electronics. In electronics, one is interested in the time characterization of the (usually single) output of a circuit which is driven by a (usually stationary ergodic) input. In the IR modeling problem, we do not have multiple time looks at a single output. Rather we have a single time look at multiple outputs where the inputs are driven by nonstationary, nonlinear nonergodic driving functions. This is a very different type of characterization problem.

indexing information, I, and target type hypotheses, H_x and H_y. To do this, the recognizer
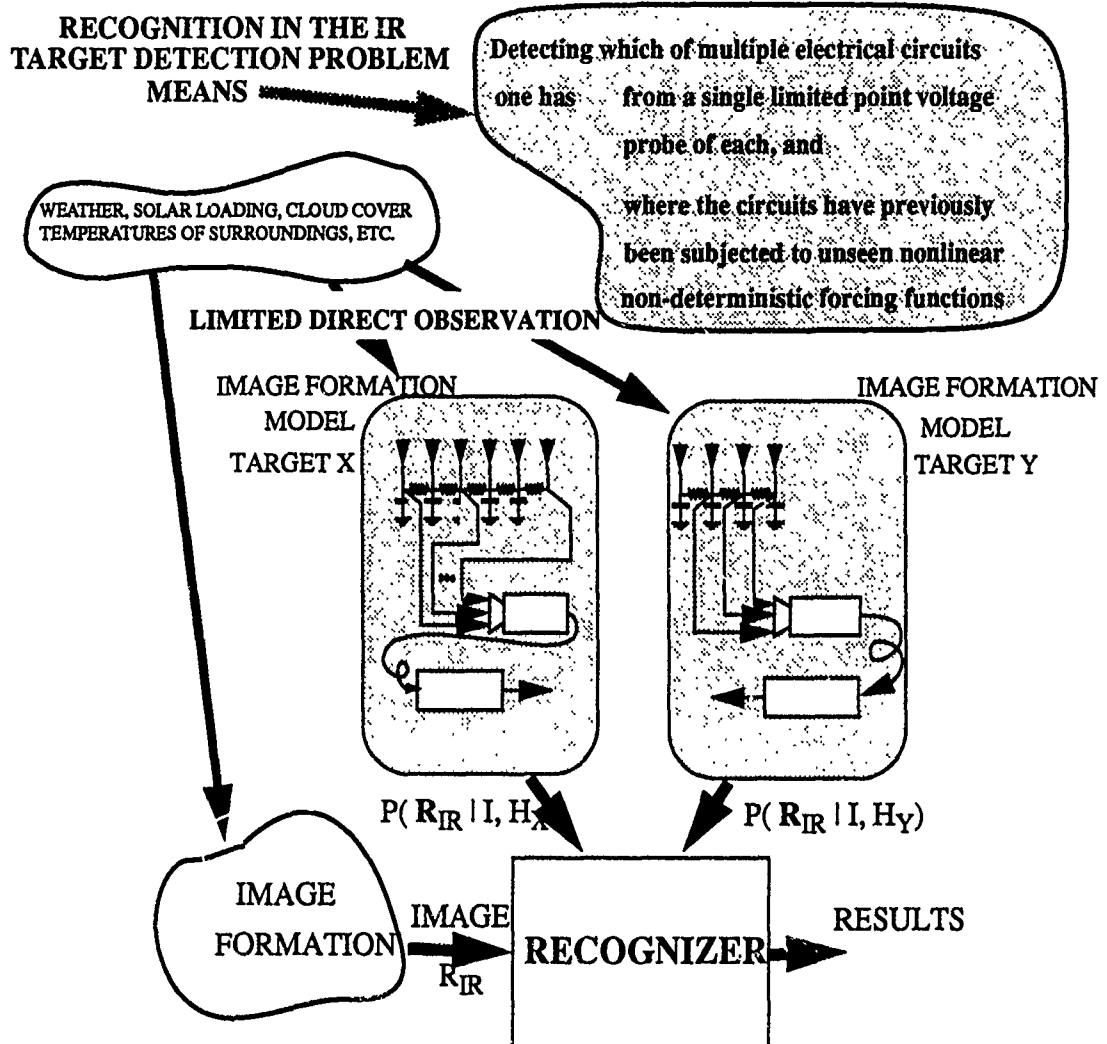
Figure 3. The ATR Scenario

needs the functional forms, $P(R_{IR} | I, H_x)$ and $P(R_{IR} | I, H_y)$, of the two probabilities, $P(R_{IR} | I, H_x)$ and $P(R_{IR} | I, H_y)$. Given $P(R_{IR} | I, H_x)$ and $P(R_{IR} | I, H_y)$, the recognizer will then insert R_{IR} to obtain the two scalar values $P(R_{IR} | I, H_x)$ and $P(R_{IR} | I, H_y)$ of interest to the final decision. Thus the forms, $P(R_{IR} | I, H_x)$ and $P(R_{IR} | I, H_y)$, are the probabilistic characterizations the ATR needs from the IR modeling effort.

3.5 Probabilistic IR Modeling

The problem which now arises is "How to determine in detail the forms, $P(R_{IR} | I, H_x)$ and $P(R_{IR} | I, H_y)$?"

Figure 3 shows the conditional probabilities, P(R_{IR} I I, H_x) and P(R_{IR} I I, H_y), are each simultaneously:

- a deterministic function of the modeling circuit of the respective targets, and

- a non-deterministic function of the thermal environment within which the target is immersed.
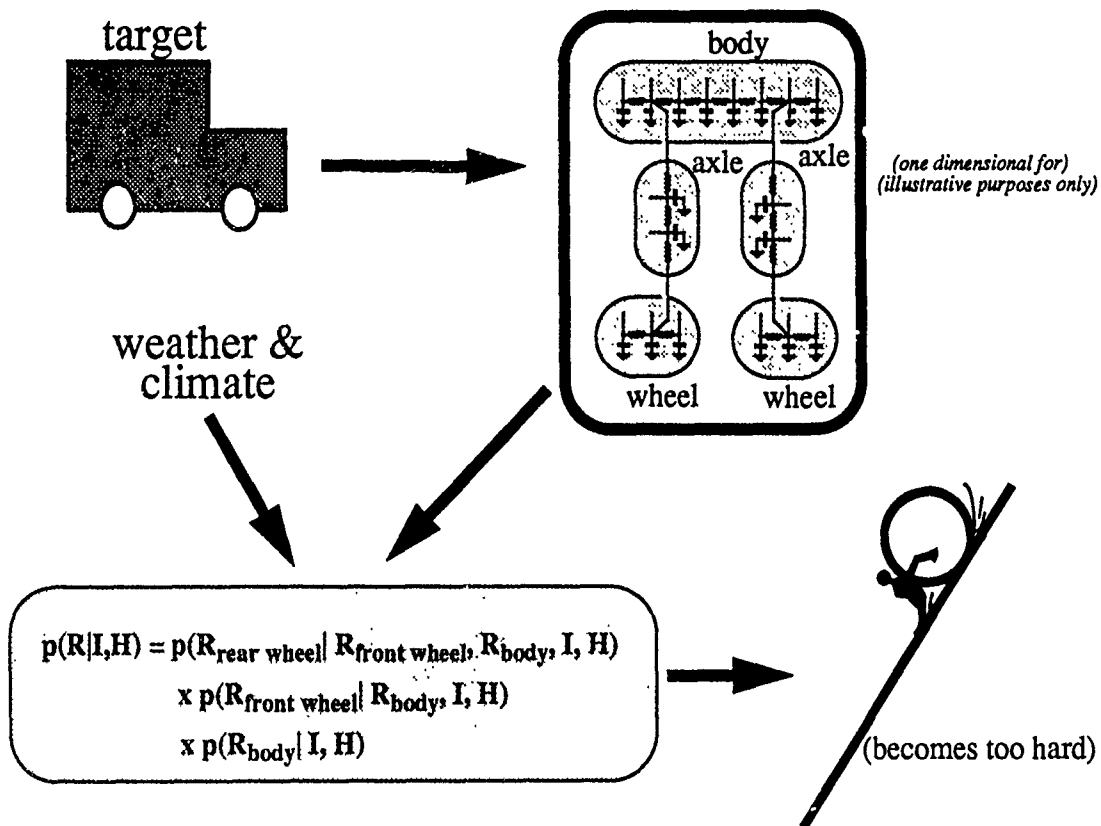
Figure 4. A First Probabilistic Modeling Approach

The second point is crucial. Although the non-deterministic thermal environment acts directly on the image, R_{IR}, received by the recognizer, the modeling formation process will have very limited observation of the thermal environment. Basically the reason is that in a battlefield engagement, it will be impossible to obtain a time history set of observations of important thermal environmental parameters such as solar loading, wind and target motion. Thus the solution must be probabilistic.

3.5.1 An Early IR Probabilistic Problem Decomposition

Figure 4 shows the first approach that was taken toward calculation of IR modeling proba-
bilities. It was based upon the concept that the physical target model can be reasonably
"chunked", due to the BIOT condition, to allow the probability statement to be written in
the conditional form shown. The concept, as then envisioned, was that consideration of
smaller target components could be somewhat isolated from major thermal environmental
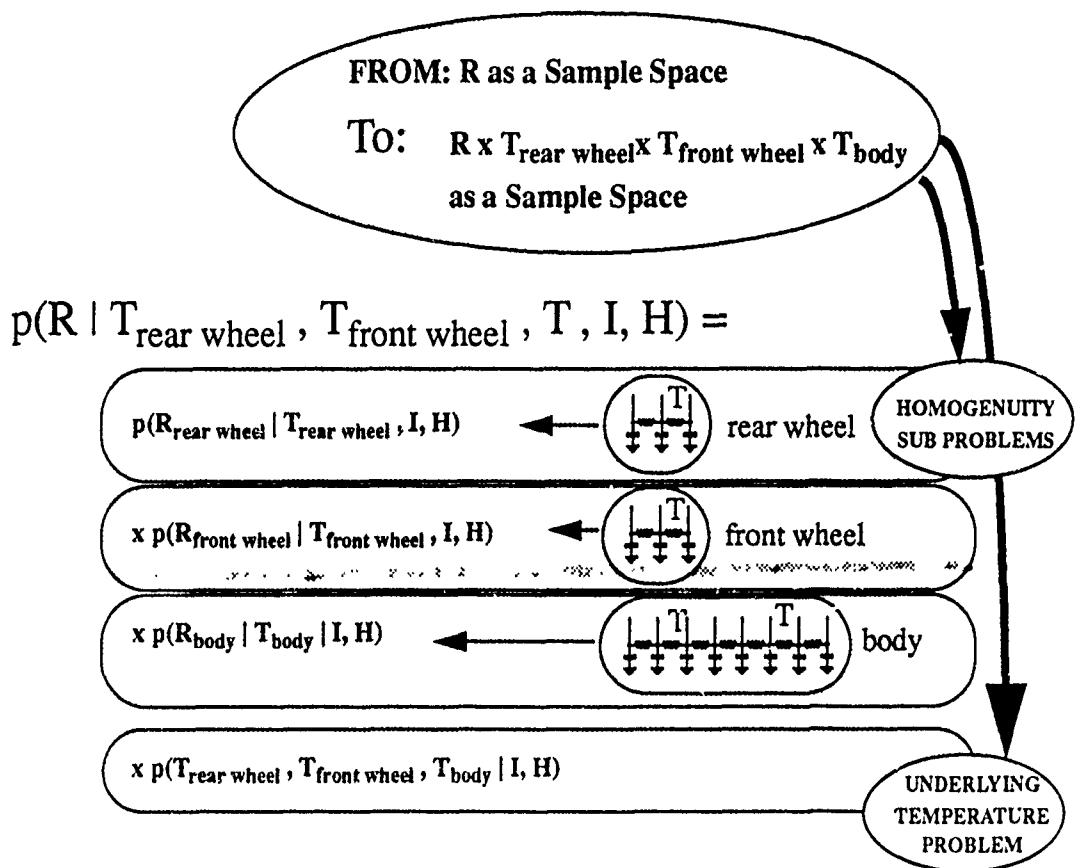
Figure 5. Transforming The Problem By Switching Sample Spaces

effects by conditioning with respect to the larger target body. While a step in the correct di-
rection of breaking down the probability calculation in terms of more elemental quantities,
this formulation could not explicitly break out the effects of the weather and climatic ther-
mal environmental influences. As a result the detailed modeling determination of $P(R_{IR} \mid I, H_x)$ and $P(R_{IR} \mid I, H_y)$ would still remain a Herculean task.

3.5.2 A More Satisfactory IR Probabilistic Problem Breakdown

After considerable further effort, the much more satisfactory probability breakdown of Figure 5 was found. The main utility of this breakdown is a separation of the probabilistic effects of space and time in the thermal target behavior dictated by Eq. 1.

By conditioning on the underlying temperatures of the object, $P(R_{IR} \mid I, H)$, for each H, can be conditionally partitioned. The first element of the partition is a probability $P(T_{rear\ wheel}, T_{front\ wheel}, T_{body} \mid I, H)$ that depends only upon the surrounding thermal environment. The other elements of the conditional partition are probabilities, $P(R_{rear\ wheel} \mid T_{rear\ wheel}, I, H)$, $P(R_{front\ wheel} \mid T_{front\ wheel}, I, H)$, $P(R_{body} \mid T_{body} \mid I, H)$, that depend only on the underlying temperatures of their own individual target body components. Thus the probabilistic calculation is cleanly split into distinct subproblems: the underlying temperature and the homogeneity problems. Further, Figure 5 points out that the modeling circuit can also be split into individual components for separate simplified consideration. In each, the underlying temperature of the part is introduced as a voltage forcing function.

The joint probability of the underlying target temperatures do not depend in any way on such non-deterministic variables as sensor noise, atmospheric effects, and emissivity variations. This of course is a further indication of the clean split. The determination of the joint probability of underlying target temperatures is taken up in the next section.

The set of IR observational probabilities conditioned on the underlying target temperatures[2] do not depend upon the surrounding thermal environment or each other[3] - giving further evidence of the clean split of the original problem. Their determination is further considered in Section 3.7.

---

2. In the example of Figure 5, these were $P(R_{rear\ wheel} \mid T_{rear\ wheel}, I, H)$, $P(R_{front\ wheel} \mid T_{front\ wheel}, I, H)$ and $P(R_{body} \mid T_{body} \mid I, H)$.

3. It is realized that there are other usually second order mutual couplings between target body components, chiefly radiation. Often such interaction will be small and even if not, further research should be able to extend the degree of decoupling.

3.6 The Underlying Part Temperature Subproblem

In summary form, Figure 6 shows the research work that went on toward the characteriza-
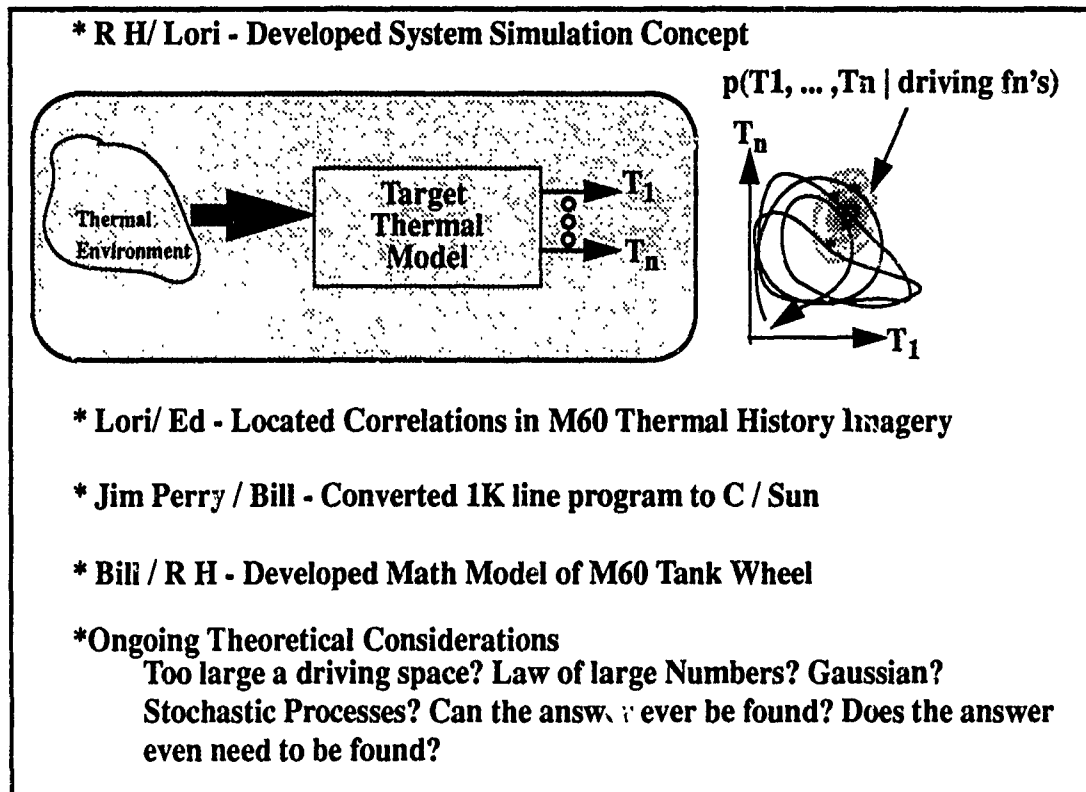
Figure 6. Consideration of The Underlying Temperatures of a Target

tion of the underlying joint target temperature probability. Starting with a system level concept, there was a desire to develop a corresponding simulation as an aid and cueing mechanism for further thought. Key to this is the state variable graph of Figure 6. If a real (or simulated) non-deterministic thermal time environment is input to the deterministic target thermal model as shown in the accompanying system block diagram to the left, one will obtain the representative target temperatures T_1 through T_n - also as a function of time. These target temperatures can then be parametrically plotted as shown by the coiled solid line of the graph. Each coiled loop represents the joint underlying temperatures of a day.

The fact that the loops, or equivalently joint daily temperature cycles, do not deterministically repeat are due to incidental daily non-deterministic variations of cloud cover, season, etc. This was confirmed by work done by Lori Westerkamp. Working from real M60 tank IR imagery data taken over a period of time, she plotted actual diurnal cycles which did not, in fact, repeat.

As the state variable temperatures track out their daily variations, they repeatedly cross any specific time of day that an image might be taken, as represented by *'s on the graph. At any one time of day, sufficient numbers of these points will begin to define a joint probability density function of the underlying target temperatures - such as the $P(T_{rear\ wheel}, T_{front\ wheel}, T_{body} | I, H)$ of Figure 6. Note that the resulting joint temperatures are not just dependent upon conditions at the time of the image, but also upon conditions for many hours past. Obviously this dependance slowly dies out as one goes backwards in time. Note then that the underlying diurnal temperatures of the target are really just filtered observations of the limit cycle of the strange attractor of Nature's thermal cycling[4]. This gives hope that the joint probability density function of the underlying target temperatures can be reasonably stated in terms of fractal theory.

A first prototype of a graphical interface was developed during this summer to allow direct simulation and visualization of the diurnal state variable cycling. Provisions were made:

- to allow changes to environmental parameters,

- to allow seeing the sweep of the cyclic behavior from one to another limit cycle, and

- to gather up time points visually showing the development of the underlying target joint pdf.

---

4. Initial reviews of fractal theory were made in the 1989 Summer FollowOn Research effort on the conviction that fractal theory plays a role in the domain of ATRs.

Due to the shortness of time, the capability was demonstrated using a different fractal than the thermal environment/target model - since the thermal environment/target model was still being simultaneously developed.

The Academic/AARA research team made many strides to provide reasonable thermal target models during the summer. Key factors here were speed of computation combined with reasonable accuracy since a large number of daily cycles must be computed before one begin to see the emerging pdfs. Jim Perry working with Bill Foley, an AARA IR expert, took one of Mr. Foley's IR computer models written in BASIC for the IBM AT color graphics environment and converted it for use in the much higher performance SUN workstation environment. R H Cofer working with Mr. Foley helped to develop a BIOT conditioned heat transfer model of a lower wheel assembly of the M60 tank from physical measurements in the field. Mr. Foley constructed the first draft of the model in BASIC computer code. Dr. Cofer took this data, reconverted it back to a theoretical form and suggested several changes to improve certain aspects of its accuracy. After further verification, he then modified it for traceability of heat flows, and was in the process of developing detailed documentation at the end of the research period. The next two steps which should be taken are to code the model within the SUN environment and to develop means of simulating a chronological thermal environment of months in duration, most likely along fractal lines. Upon completion of these two steps, an actual simulation can be run that will begin to give a indication of the nature of the joint pdf of the underlying target temperatures. This indication can then be used as a cueing mechanism for guiding analytic research toward a more mathematical derivation of the joint pdf.

Figure 6 also shows several of the ongoing positive and negative theoretical considerations that are currently considered likely to play to a role in future heat transfer simulations. Of these the theory of large numbers is most likely to provide a means by which the pdf can be analytically found. Once major periodicities of the diurnal cycle have been removed, it is also hypothesized that there will remain large numbers of relatively independent non-

deterministic variables. These will be summed by the thermal filtering properties of the target to give, upon proper probabilistic conditioning, a limiting pdf - perhaps even with Gaussian tendencies.

Figure 6 then poses the question, "Can the answer ever be found?" Obviously the joint pdfs, if obtainable, will be of use and particularly so in the case of target-decoy discrimination where geometries can be the same but thermal properties will differ. On the other hand, it is certain that the underlying target part temperatures are functions of a seemingly endless list of time and space dependent non-deterministic variables: cloud cover, shadowing, rain, wind, target orientation, target position, terrain, mud plastering, variations of paint emissivity, internal heating, target manufacturing variabilities, etc. It will not only be a potentially difficult problem to characterize these non-deterministic driving functions, but in a realistic battlefield situation one may not even know the full list. Additionally, it can be challenging to develop multi-dimensional pdfs due to the sheer size of the mathematical spaces. In summary, the fundamental limitation of developing joint target temperature pdfs is ultimately dependent upon the nature of the thermal environment. This nature is still in question and unverified. If it can be characterized sufficiently well, along fractal or other lines of thought, then joint target temperature pdfs may be indeed be obtainable.

The deeper question posed by Figure 6 is "Does the answer (to the underlying target part temperature question) even need to be found?" Dr. Cofer's continued inspection of actual IR imagery and related study indicates that the answer may very well lie in the area discussed in Section 3.8.

3.7 The Homogeneity Subproblem

The homogeneity probabilistic characterization subproblem is defined as determining the form of the probability of the image regions of target components conditioned on underlying target component temperatures. In Figure 5, these probabilities are $P(R_{\text{rear wheel}} \mid T_{\text{rear wheel}}, I, H)$, $P(R_{\text{front wheel}} \mid T_{\text{front wheel}}, I, H)$ and $P(R_{\text{body}} \mid T_{\text{body}}, I, H)$.

In the case where the BIOT condition reigns and emissivity holds constant across the surface, homogeneity probability calculations will be easily developable as a function of the sensor image chain. For other less common cases where BIOT conditioning does not hold and/or emissivity does not remain constant across the surface, further research will be required. A portion of the summer was devoted to amassing the necessary research materials to continue such an investigation though the ensuing academic year.

3.8 Robustness Given Lack of Knowledge of Underlying Temperatures

Jim Leonard, as responsive and responsible AARA technical contact for this research effort, was particularly adroit at setting one squarely on the horns of the dilemma: "How do you expect to complete the probabilistic characterization of IR targets, given the manifest problems in determining underlying target temperatures?" A fair question to which much thought was given and the following answer was developed: "Estimate the underlying temperatures from the image data itself and then use this estimate to developing the final probabilities, $P(R_{IR}|I,H)$, for each H."

Figure 7 portrays the idea. Figure 7 (a) shows the normal breakdown of $P(R_{IR}|I,H)$ and individual sources of information. Figure 7 (b) shows an alternate breakdown where the underlying temperatures are taken from estimates on the image itself. Two questions can be raised of this alternate breakdown:

- "What is the effect of errors in the functional form of the probabilities?", and

- "What is the error introduced by the approximation of the underlying temperatures themselves."

Research during the period indicates that increasing the image resolution can overcome both concerns. The obvious question is "How much must the image resolution be raised?" By example, an answer of not much. Shown in Figure 8 is a feature extraction problem: "Is the image region a rear wheel of a truck or background?" The domain knowledge and rec-

ognition setup information is also shown. The wheel temperature, T_W, is estimated by its mean taken from the image. The resulting recognition probability, P_W, is plotted as a fam-

IMAGE   MODEL

$$P(R_{IR}|I,H) = \prod_i P(R_1(i)|T_1,I,H) \prod_j P(R_2|T_2,I,H) P(T_1|I,H) P(T_2|I,H)$$

*the homogeneity
subproblem*

*the underlying temper-
ature subproblem*
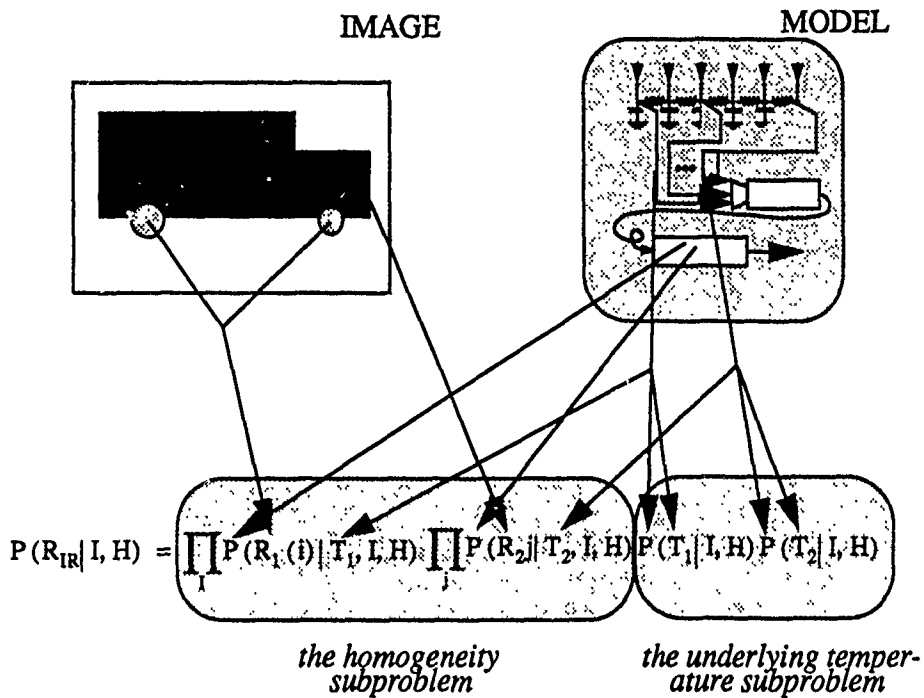
(a) Normal Probabilistic Breakdown

IMAGE   MODEL

*max
likelihood*   *max
likelihood*

$$P(R_{IR}|I,H) = \prod_i P(R_1(i)|T_1,I,H) P(T_1|I,H) \times \prod_j P(R_2(j)|T_2,I,H) P(T_2|I,H)$$
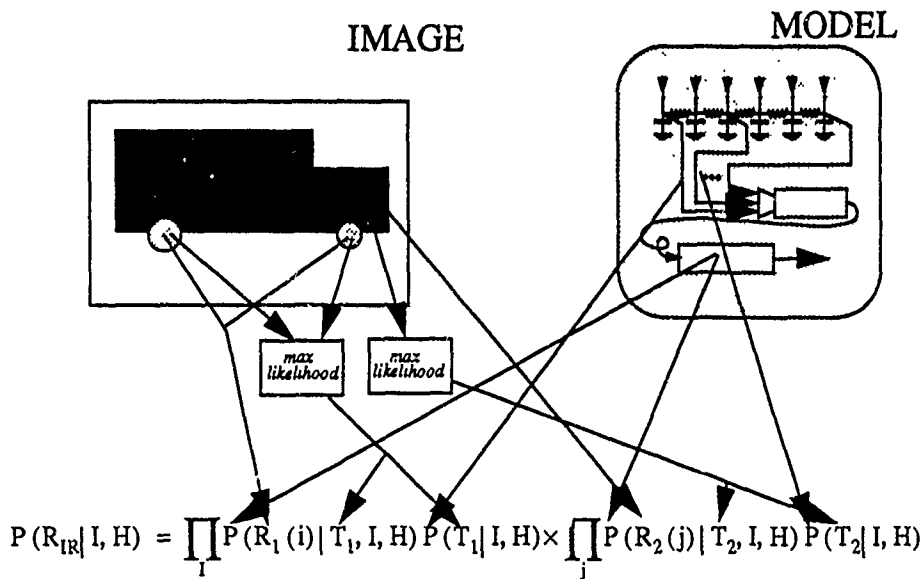
(b) Maximum Likelihood Probabilistic Breakdown

Figure 7. Alternate Homogeneity Subproblem Breakdowns

ily over a range of probabilities of misrecognition of background, P_f. The abscissa is a normalized function of the number of pixels on the wheel, N; the true difference of temperature of wheel from background, M; and the standard deviation of the sensor
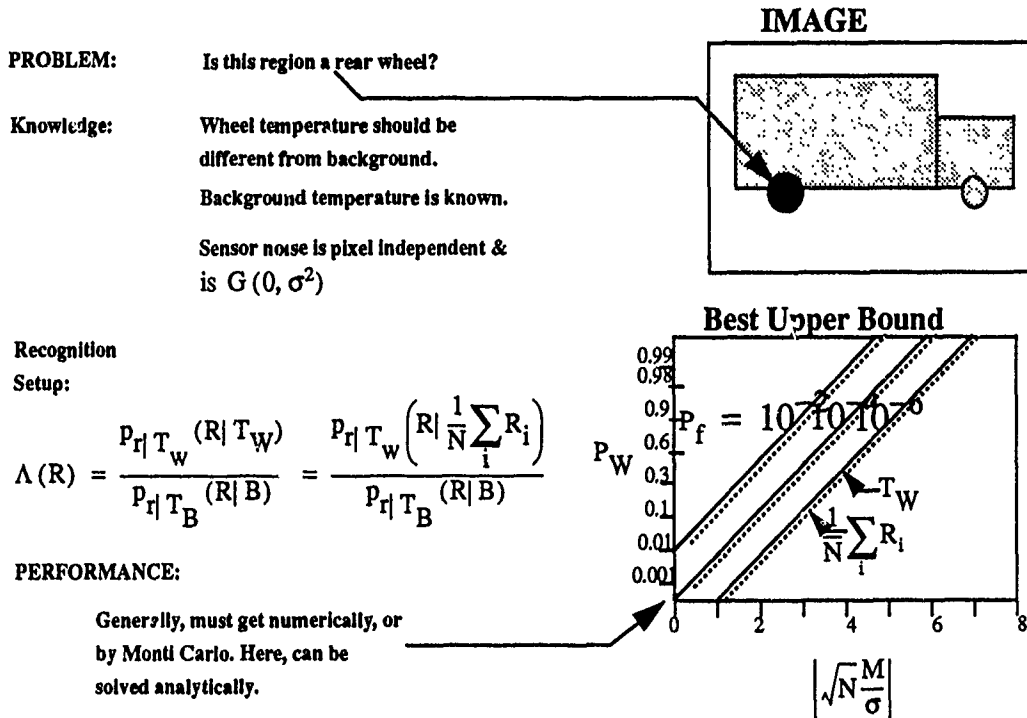
Figure 8. Robustness Against Unknown Underlying Temperature

noise, σ. The solid line shows the accuracy if the true underlying temperature were perfectly known. The dotted line gives the recognition performance when the temperature is estimated from the image itself. Obviously, the drop of performance is small and is correctable by a very small increase in image resolution. Note no modeling of underlying temperature probabilities was required, thus that problem was avoided.

The above example is representative of the robustness generally achievable by estimating the underlying target temperatures from the image itself. Of course, research should be continued to further extend the above results.

4.0 RECOMMENDATIONS FOR FURTHER RESEARCH

Probabilistic evidence accumulation is a most promising methodology for use in ATR design and implementation. It has withstood each theoretical recognition performance challenge so far in a consistent manner and has provided guidance for multisensor fusion, handling lack in target thermal history, allowing meaningful incorporation of modeling information, etc. Probabilistic evidence accumulation should also prove particularly useful in coping with the strong speckle/mutual interference probabilistic component of SAR imagery.

The Academic/AARA research team should continue its joint investigations into:

- Computer performance modeling and simulation of LADAR, IR and LADAR/IR target recognition based upon target and image chain modeling,

- In depth real world consideration of the variability of IR target skin effects including emissivity, thermal conductivity, and directional radiation characteristics under varying conditions of rain, humidity, splattered mud, cloud cover, etc.,

- The probabilistic effects of optical blur,

- A spatial probabilistic temperature representation of metallic surfaces under nearly BIOT conditions, and

- The probabilistic representation of background clutter.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

<u>FINAL REPORT</u>

<u>Investigations of a Lower Bound</u>

<u>on the Error in Learned Functions</u>

Prepared by: Thomas K. Gearhart, Ph.D.

Academic Rank: Associate Professor

Department and   Department of Mathematics, Computer
                      Science, and Physics

University: Capital University

Research Location: Avionics Laboratory
                      WRDC/AART-2
                      Wright-Patterson AFB   OH   45433-6543

USAF Researcher: Timothy D. Ross, Ph.D.

Date: 27 July 90

Contract No: F49620-88-C-0053

Investigations of a Lower Bound

on the Error in Learned Functions

by

Thomas K. Gearhart

ABSTRACT

An exact expression for the difference between the
average sum-of-squares error for a collection of learned
functions and a lower bound on that error is derived. A
bound on the difference of the average sum-of-squares
errors for two distinct collections of learned functions is
obtained. This bound can be computed without knowledge of
the desired function outside the training sets. A condition
is isolated which assures that average sum-of-squares error
will decrease as the size of the underlying training sets
increases. The lower bound on the average sum-of-squares
error is experimentally compared with the traditional
measure of error for specific machine learning systems.

ACKNOWLEDGMENTS

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research, whose sponsorship provided me this opportunity to participate in a stimulating research project. I also thank the staff of the SFRP Office at Universal Energy Systems for their careful and helpful handling of the administrative aspects of the program.

Numerous individuals in the Avionics Laboratory at Wright-Patterson AFB contributed to making my stay enjoyable and productive. John Jacobs and his staff in the System Concepts Group were most hospitable and provided a very pleasant work environment. I am especially grateful to Dr. Tim Ross for his assistance in identifying an interesting and challenging research problem, and for providing direction to my investigations throughout the Summer. Conversations with Mike Breen and Lt. Tim Taylor were helpful in clarifying certain issues related to my project. The assistance of Shannon Spittler and Peggy Alltop in analyzing experimental data and preparing graphs is also gratefully acknowledged.

I also wish to thank Dr. Harold Brockman, Chairman of the Department of Mathematics at Capital University, for his encouragement and support of my application to this program.

I. INTRODUCTION

Since January, 1989, Dr. Timothy Ross and associates in the Avionics Laboratory have been engaged in an extensive study of Pattern Based Machine Learning (PBML). A central finding of the study is the existence of a strong connection between pattern-ness and the concept of function decomposition. Techniques and tools from both mathematics and computer science have played an important role throughout this research effort.

My initial technical training was in pure mathematics with a concentration in general topology. More recently, I completed eleven months of graduate studies in computer science during a sabbatical leave granted by my university. These studies included courses in digital logic, pattern recognition, and automata theory. Each of these three fields has a direct connection to some aspect of the PBML study. My recent research interests in computer graphics are also complementary to the PBML study from the standpoint of developing programs for the visualization of the phenomena under study.

II. OBJECTIVES OF THE RESEARCH EFFORT

Machine learning systems (MLS) attempt to predict the global behavior of a function using information about the function on a small sample of its domain. An important issue is the determination of how large such a sample should be. In [4], Dr. Ross describes a method for determining a lower bound on the average sum-of-squares error resulting from samples of a given size. During a meeting prior to the start of my research period, it was decided that I would develop ideas introduced in this paper and attempt to resolve various conjectures advanced therein. My approaches to meeting these objectives involved the application of mathematical techniques and the development of computer programs for the simulation of various learning systems. These approaches are discussed in sections III and IV.

III. THEORETICAL INVESTIGATIONS

In this section mathematical techniques are applied to develop ideas introduced in [4].

Let $f:X \rightarrow R^n$ be a function from a finite set X into a Euclidean space. The notation $f|_T$ will be used to denote the restriction of f to subset T of X. A <u>Machine Learning System</u> (MLS) is a system which uses a restricted function $f|_T = \{ (x, f(x)) \mid x \in T \}$ to generate a function $a_T:X \rightarrow R^n$. A function a_T so generated by an MLS is called a <u>learned function</u> associated with the training set T. The

performance of an MLS can be measured by how closely such learned functions approximate the function f over the entire domain X.

To make more precise the notion of closeness of functions, define the sum-of-squares difference between functions f and g as:

$$d(f,g) = \sum_{x \in X} \| f(x) - g(x) \|^2$$

where the double bars indicate the usual Euclidean norm. If A is a collection of functions from X into R^n, define the sum-of-squares difference between function f and a set of functions A as:

$$d(f,A) = \sum_{a \in A} d(f,a)$$

Let $|A|$ denote the cardinality of A. The following theorem establishes an important relationship between d(f,A) and $d(\alpha,A)$ where α is the average of the functions in A.

__Theorem 1__ Let X be a finite set and suppose f:X -> R^n. Let A be a finite collection of functions from X into R^n. Define the function α by $\alpha(x) = (\sum_{a \in A} a(x)) / |A|$.

Then $d(f,A) - d(\alpha,A) = |A| \, d(f,\alpha)$.

__Proof__ In this proof the notation <V,W> will denote the usual inner product of vectors V and W in R^n.

$d(f,A) - d(\alpha,A)$

$$= \sum_{a \in A} \sum_{x \in X} \| f(x) - a(x) \|^2 - \sum_{a \in A} \sum_{x \in X} \| \alpha(x) - a(x) \|^2$$

$$= \sum_{a \in A} \sum_{x \in X} (\| f(x) - a(x) \|^2 - \| \alpha(x) - a(x) \|^2)$$

$$= \sum_{x \in X} \sum_{a \in A} (<f(x) - a(x), f(x) - a(x)> - <\alpha(x) - a(x), \alpha(x) - a(x)>)$$

$$= \sum_{x \in X} \sum_{a \in A} (<f(x), f(x)> - 2<a(x), f(x)> + <a(x), a(x)>$$
$$- <\alpha(x), \alpha(x)> + 2<\alpha(x), a(x)> - <a(x), a(x)>)$$

$$= \sum_{x \in X} \sum_{a \in A} (<f(x), f(x)> - 2<a(x), f(x)>$$
$$- <\alpha(x), \alpha(x)> + 2<\alpha(x), a(x)>)$$

Since the inner product is distributive over vector addition and $\sum_{a \in A} a(x) = |A| \alpha(x)$, the above double sum reduces to:

$$= \sum_{x \in X} (|A| <f(x), f(x)> - 2|A| <\alpha(x), f(x)>$$
$$- |A| <\alpha(x), \alpha(x)> + 2|A| <\alpha(x), \alpha(x)>)$$

$$= |A| \sum_{x \in X} (<f(x), f(x)> - 2<f(x), \alpha(x)> + <\alpha(x), \alpha(x)>)$$

$$= |A| \sum_{x \in X} <f(x) - \alpha(x), f(x) - \alpha(x)>$$

$$= |A| \sum_{x \in X} \| f(x) - \alpha(x) \|^2 = |A| \, d(f, \alpha). \qquad \blacksquare$$

Corollary 1 $d(f, A)/|A| - d(\alpha, A)/|A| = d(f, \alpha)$.

Proof   Divide both sides of the equality in the conclusion of the theorem by $|A|$. $\blacksquare$

Corollary 2 $d(f, A)/|A| \geq d(\alpha, A)/|A|$.

Proof   The right side of the equality in corollary 1 is nonnegative. $\blacksquare$

Corollary 2 above has important consequences in the context of machine learning systems. Suppose the MLS is

attempting to learn a function f. Let A be the collection
of learned functions output by the MLS when trained on a
collection of training sets of a given size. Then the
corollary guarantees that $d(\alpha,A)/|A|$ is a lower bound on
the average sum-of-squares error $d(f,A)/|A|$.

Since in practical applications of machine learning
systems the function f is known only on the training sets,
$d(f,A)/|A|$ is not computable. However, since each member
of A and the average α is known over the entire domain,
$d(\alpha,A)/|A|$ can be computed. Thus we have a computable
lower bound on the average sum-of-squares error.

In many instances the functions under study are
Boolean functions. A Boolean function $f:X \to Y$ has
domain $X = \{0,1\}^n = \{0,1\} \times \{0,1\} \times \ldots \times \{0,1\}$ and
codomain $Y = \{0,1\}$. Under these circumstances the
pointwise average α of a collection A of functions might
take on values other than 0 and 1 and hence might not be a
Boolean function. However, a Boolean function α' can be
defined from α by setting $\alpha'(x) = 1$ when $\alpha(x) \geq 1/2$
and setting $\alpha'(x) = 0$ when $\alpha(x) < 1/2$. Note that $\alpha'(x)$
= 1 when $a(x) = 1$ for at least half of the functions in the
collection A and $\alpha'(x) = 0$ otherwise. The question of the
existence of an analogue of Theorem 1 for Boolean functions
is answered affirmatively by the following.

Theorem 2 Let X be a finite set and suppose $f:X \to \{0,1\}$.
Let A be a finite collection of functions mapping X into

{0,1}. Define the function $\alpha: X \to [0,1]$ by

$$\alpha(x) = (\sum_{a \in A} a(x)) / |A| .$$

and define $\alpha': X \to \{0,1\}$ by $\alpha'(x) = 1$ if $\alpha(x) \geq 1/2$

and $\alpha'(x) = 0$ if $\alpha(x) < 1/2$.

Then $0 \leq d(f,A) - d(\alpha',A) \leq |A| \, d(f,\alpha')$.

<u>Proof</u> Define subsets S, T and U of X as follows:

$$S = \{ x \in X \mid f(x) = \alpha'(x) \}$$

$$T = \{ x \in X \mid f(x) = 0 \text{ and } \alpha'(x) = 1 \}$$

$$U = \{ x \in X \mid f(x) = 1 \text{ and } \alpha'(x) = 0 \}$$

Note that S, T, and U are disjoint and their union is X.

By definition, $d(f,A) - d(\alpha',A)$

$$= \sum_{a \in A} \sum_{x \in X} (f(x)-a(x))^2 - \sum_{a \in A} \sum_{x \in X} (\alpha'(x)-a(x))^2$$

$$= \sum_{a \in A} \sum_{x \in X} [(f(x)-a(x))^2 - (\alpha'(x)-a(x))^2] \qquad (1)$$

Since the bracketed expression in (1) is zero on S,

(1) can be rewritten as:

$$\sum_{a \in A} \sum_{x \in T} [(f(x)-a(x))^2 - (\alpha'(x)-a(x))^2]$$

$$+ \sum_{a \in A} \sum_{x \in U} [(f(x)-a(x))^2 - (\alpha'(x)-a(x))^2] \qquad (2)$$

Using the definitions of T and U, (2) is the same as:

$$\sum_{a \in A} \sum_{x \in T} [(-a(x))^2 - (1 - a(x))^2]$$

$$+ \sum_{a \in A} \sum_{x \in U} [(1 - a(x))^2 - (-a(x))^2]$$

$$= \sum_{a \in A} \sum_{x \in T} (2\, a(x) - 1) + \sum_{a \in A} \sum_{x \in U} (1 - 2\, a(x)) \qquad (3)$$

Since $\sum_{a \in A} a(x) = |A|\, \alpha(x)$, (3) is equal to:

$$|A| \left(\sum_{x \in T} (2 \, \alpha(x) - 1) + \sum_{x \in U} (1 - 2 \, \alpha(x)) \right) \qquad (4)$$

For $x \in T$, $1/2 \leq \alpha(x) \leq 1$ and so,

$$0 \leq 2 \, \alpha(x) - 1 \leq 1 \quad \text{for } x \in T.$$

For $x \in U$, $0 \leq \alpha(x) < 1/2$ and so,

$$0 < 1 - 2 \, \alpha(x) \leq 1 \quad \text{for } x \in U.$$

Thus it follows that (4) is bounded below by 0 and bounded above by $|A| (|T| + |U|)$. Finally, note that since f and α' take on only values from the set $\{0,1\}$,

$d(f,\alpha') = \sum_{x \in X} (f(x) - \alpha'(x))^2$ is simply the cardinality

of the subset of X on which f and α' disagree. But this cardirality is simply $|T| + |U|$. Since $d(f,A) - d(\alpha',A)$ is equal to (4), the theorem follows.   ∎

Suppose $d(f,A)/|A|$ is the average sum-of-squares error associated with a set A of functions generated by an MLS training on sets of a given fixed size. Let $d(f,B)/|B|$ be the average sum-of-squares error associated with set B of functions generated by the MLS training on sets of larger fixed size. Dr. Ross, in [4], conjectured that it might be possible to obtain a bound on $| d(f,A)/|A| - d(f,B)/|B| |$ which involved $d(\alpha,A)/|A|$. The following theorem is a step in this direction.

Theorem 3   Suppose $f: X \to \mathbf{R}^m$ where X is a finite set and $\| f(x) \| \leq M$ for every $x \in X$. Let A and B be two collections of functions mapping X into $\mathbf{R}^m$, and let α and β be the averages associated with A and B respectively.

That is, let $\alpha(x) = (\sum_{a \in A} a(x)) / |A|$ and let $\beta(x) = (\sum_{b \in B} b(x)) / |B|$.

Then $| d(f,A)/|A| - d(f,B)/|B| | \leq$

$$2M \sum_{x \in X} \|\alpha(x) - \beta(x)\| + | \sum_{x \in X} (\|\alpha(x)\|^2 - \|\beta(x)\|^2) |$$

$$+ | d(\alpha,A)/|A| - d(\beta,B)/|B| |.$$

<u>Proof</u>   Using Corollary 1 of Theorem 1 and the triangle inequality, $| d(f,A)/|A| - d(f,B)/|B| |$

$= | (d(\alpha,A)/|A| + d(f,\alpha)) - (d(\beta,B)/|B| + d(f,\beta)) |$

$\leq |d(f,\alpha) - d(f,\beta)| + | d(\alpha,A)/|A| - d(\beta,B)/|B| |.$

Now $d(f,\alpha) - d(f,\beta) = \sum_{x \in X} \|f(x) - \alpha(x)\|^2 - \sum_{x \in X} \|f(x) - \beta(x)\|^2$

$= \sum_{x \in X} (<f(x) - \alpha(x), f(x) - \alpha(x)> - <f(x) - \beta(x), f(x) - \beta(x)>)$

$= \sum_{x \in X} (\quad <f(x), f(x)> - 2<f(x), \alpha(x)> + <\alpha(x), \alpha(x)>$

$\quad\quad\quad - <f(x), f(x)> + 2<f(x), \beta(x)> - <\beta(x), \beta(x)> \quad)$

$= \sum_{x \in X} (-2 <f(x), \alpha(x) - \beta(x)> + \|\alpha(x)\|^2 - \|\beta(x)\|^2)$

$= -2 \sum_{x \in X} <f(x), \alpha(x) - \beta(x)> + \sum_{x \in X} (\|\alpha(x)\|^2 - \|\beta(x)\|^2).$

Thus, $| d(f,\alpha) - d(f,\beta) |$

$\leq 2 \sum_{x \in X} | < f(x), \alpha(x) - \beta(x) > | + | \sum_{x \in X} (\|\alpha(x)\|^2 - \|\beta(x)\|^2) |$

By the Cauchy-Schwarz inequality, $| < f(x), \alpha(x) - \beta(x) > |$

$\leq \|f(x)\| \|\alpha(x) - \beta(x)\| \leq M \| \alpha(x) - \beta(x) \|$ and

the theorem follows.   ∎

A desirable feature of an MLS would be that $d(f,A)/|A|$

decrease as the size of the underlying training sets

increases. The following discussion and theorem establish a condition under which this will occur.

Suppose X is a finite set and $f:X \to R^m$ and let S be a sample set of size n selected from X. Let $a_T:X \to R^m$ denote the function output by a given MLS when it is trained using $T \subseteq S$. Let us agree to call an MLS **monotone** with respect to f and S if $d(f,a_U) \leq d(f,a_T)$ for any two training sets T and U such that $S \supseteq U \supseteq T$. For each $k = 1, 2,\ldots,n$ let $A_k = \{ a_T \mid T$ is a k-member subset of S $\}$. Note that $|A_k| = C(n,k)$ where $C(n,k)$ denotes the number of k-member subsets of an n-member set.

Theorem 4 If an MLS is monotone with respect to function $f:X \to R^m$ and an n-member sample set $S \subseteq X$, then
$d(f,A_k)/|A_k| \leq d(f,A_j)/|A_j|$ whenever $k \geq j$.
Proof Let U be a k-member subset of S. Since the MLS is monotone, $d(f,a_U) \leq d(f,a_T)$ for any j-member subset T of U. Since there are $C(k,j)$ such j-member subsets of U,

$$C(k,j) \; d(f,a_U) \leq \sum_{\substack{T \subseteq U \\ |T| = j}} d(f,a_T).$$

Summing each side of this inequality over all k-member subsets of the sample set S yields:

$$C(k,j) \sum_{\substack{U \subseteq S \\ |U| = k}} d(f,a_U) \leq \sum_{\substack{U \subseteq S \\ |U| = k}} (\sum_{\substack{T \subseteq U \\ |T| = j}} d(f,a_T))$$

In the double sum on the right, a given $d(f,a_T)$ appears multiple times. This is because a given j-member set T is contained in several distinct k-member sets U. The number of occurrences of $d(f,a_T)$ in the double sum is $C(n-j,k-j)$,

because $C(n-j,k-j)$ is the number of ways of enlarging a fixed j-member subset of S to a k-member subset of S. So the double sum can be replaced by $C(n-j,k-j) \sum_{\substack{T \subseteq S \\ |T| = j}} d(f,a_T)$.

Hence, $C(k,j) \sum_{\substack{U \subseteq S \\ |U| = k}} d(f,a_U) \leq C(n-j,k-j) \sum_{\substack{T \subseteq S \\ |T| = j}} d(f,a_T)$.

That is, $C(k,j) \, d(f,A_k) \leq C(n-j,k-j) \, d(f,A_j)$.

Since $|A_k| = C(n,k)$ and $|A_j| = C(n,j)$ the above inequality is equivalent to:

$C(k,j) \, C(n,k) \, d(f,A_k)/|A_k| \leq C(n-j,k-j) \, C(n,j) \, d(f,A_j)/|A_j|$.

Since $C(k,j) \, C(n,k) = C(n-j,k-j) \, C(n,j)$, the proof is complete. ▌

IV. EXPERIMENTAL INVESTIGATIONS

To assist in assessing the practical value of the lower bound on the average sum-of-squares error, a computer simulation of an MLS appropriate for learning real valued functions was developed. Given a function f and a training set, this MLS generated an approximation of f by linearly interpolating the points of the training set.

The simulation was run for several real valued functions f defined on the domain X = (1, 2, ..., 1000). In each case a sample set S of size 70 was randomly chosen from the domain X. Next, for each of the values k = 10, 20, ... , 60, ten training sets of size k were randomly chosen from S. For each k, the graphs of the ten functions, the average α, and the target function f were

all displayed on the same axis. Also the following
measures of error were computed:

$$e = \frac{\displaystyle\sum_{a\in A}\ \sum_{x\in X}\ \|f(x)-a(x)\|^2}{|A|\ |X|}$$

$$d = \frac{\displaystyle\sum_{a\in A}\ \sum_{x\in X}\ \|\alpha(x)-a(x)\|^2}{|A|\ |X|}$$

$$\hat{e} = \frac{\displaystyle\sum_{a\in A}\ \sum_{x\in S\setminus T_a}\ \|f(x)-a(x)\|^2}{|A|\ (\ |S|-k\ )}$$

where T_a is the training set inducing a.

The measure e is a normalized version of the average
sum-of-squares error and d is the theoretically guaranteed
lower bound on e. The measure $\hat{e}$ is a measure of error
which has traditionally been used as a figure-of-merit for
machine learning systems.

For this MLS based on linear interpolation, results of
the simulations suggest that $\hat{e}$ tends to be a better
estimate of e than d is.

A second MLS appropriate for use with Boolean
functions was also designed and implemented in software.
This MLS incorporated a three-layer neural network. For an
n-variable Boolean function, the input layer had n nodes,
the middle layer had 2n+1 nodes, and the output layer had a
single node. Weights for the input layer nodes were
permanently fixed at 1. Weights for the middle and output
layer nodes were initialized randomly and then updated
using the backtracking strategy described in [3].

The effectiveness of the network as an MLS varied greatly with the function being realized. It did well on functions with many vacuous variables. For example, on a seven variable x1 XOR x2 function, the error rate was only two percent with training sets of size 40 from the domain of size 128. At the other extreme, the network did no better than chance attempting to learn a seven variable parity function.

Studies of e, d, and ê were also made for the neural network MLS. The only trend that could be discerned was that e was frequently between d and ê.

V. RECOMMENDATIONS

A. If the lower bound on the average sum-of-squares error is large, then the machine learning system is performing poorly and/or the training sets are not large enough. However the converse is not true, and so small values of the lower bound need to be interpreted with caution. They need not imply that the actual error is small.

B. Simulations for the machine learning system based on linear interpolation indicate that the traditional estimate of error is often closer to the actual error than is the lower bound on the error.

C. Comparison of the three measures of error (see section IV) for an existing machine learning system based

on function decomposition might be helpful in clarifying

the usefulness of the lower bound on the average

sum-of-squares error.

D. The neural network described in section IV appears

to find some feature of the function other than

decomposability. For a wide range of functions, the

network was outperformed by an existing MLS based on

function decomposition. A systematic study of other

network topologies and activation functions might reveal

ways of improving the network's performance as an MLS.

A hybrid MLS using a combination of neural network ideas

and ideas from function decomposition theory might also be

worthy of research.

REFERENCES

1. Caudill, Maureen, "Neural Networks Primer, Part I," *AI Expert*, December 1987, Vol. 2, No. 12, pp. 46-52.

2. Caudill, Maureen, "Neural Networks Primer, Part II," *AI Expert*, February 1988, Vol. 3, No. 2, pp. 55-61.

3. Caudill, Maureen, "Neural Networks Primer, Part III," *AI Expert*, June 1988, Vol. 3, No. 6, pp. 53-59.

4. Ross, Timothy D., "A Convergence Based Lower Bound on the Error in Learned Functions," December, 1989.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Machine Learning Applied to High Range Resolution Radar Returns

Prepared by:	Lawrence O. Hall and Steve G. Romaniuk
Academic Rank:	Assistant Professor and Research Assistant
Department and	Computer Science and Engineering
University:	University of South Florida
Research Location:	WPAFB, WRDC/AARA, ATR Branch
USAF Researcher:	Jim Leonard
Date:	August 6, 1990
Contract No:	F49620-88-C-0053

Machine Learning Applied to High Range Resolution Radar Returns

by

Lawrence O. Hall and Steve G. Romaniuk

ABSTRACT

This report examines the use of a neural network learning algorithm and a hybrid neural network, symbolic learning algorithm on the problem of recognizing airplanes from high resolution radar returns. Quickprop and SC-net are the techniques used. The intent of the study is to determine how to both recognize the planes and recover the aspect angle in an algorithm with small set up and good space/time characteristics. Three different representations of the radar returns to the learning algorithm were tried. The problem of representation is very important in this study. The first two representations were geometric hashing schemes. The last is a binning and averaging scheme. It has shown some invariance to aspect angle shifts, which is important in limiting the number of training times and examples. In both learning systems the third representation has been used to get 100% recognition for so e sets of aspect angles.

Acknowledgements

I Introduction

The returns from High Range Resolution Radar provide data that may be used to identify target airplanes. Specifically, air to air recognition is an object of this study. The angle of the target is also of interest. Hence, an examination of how machine learning can be used to provide recognition of the targets from this type of data is underway. This report will discuss the issues involved and progress to date.

One important problem is to understand the nature of the data and develop an appropriate representation for machine learning. A second problem of equal import is what type of machine learning algorithm will work the best. These two issues are discussed in the following.

Either a supervised learning or unsupervised learning scheme may be applied. Supervised learning works well, when the classes can be labeled beforehand and there is someone with domain information able to guide the learning process. When there are no labels for classes and/or there is no one available with a good understanding of the domain area in which learning is to be applied, unsupervised learning may be the best choice. The issues are more involved, but in this case the domain fits the criteria for supervised learning schemes.

At the current time there is tremendous interest in neural network learning algorithms. Neural networks (also called connectionist networks) are an attempt to model the function of the human brain. When it comes to learning algorithms, some are very closely tied to the way that the brain is thought to work and others are only loosely based on a biological model. Neural network techniques for learning compete with symbolic techniques, such as ID3 [8]. Recent studies have shown that for some domains performance can be almost identical [10].

Connectionist networks are said to have an advantage in that they require less set up. However, whether this advantage exists is debated. A connectionist network can be used without modification in different non-symbolic domains. The appear

to work well with data that is visual in nature, and patterns. The domain we are working in provides patterns that may be visually recognized by people (although this is not currently done in air to air encounters). There is no readily apparent symbolic information in the domain of this problem. Hence, connectionist models are a good choice for learning to recognize the planes from the radar returns.

Two different approaches to recognition are being used. One is a hybrid connectionist, symbolic system developed at USF, called SC-net. The other is Quickprop a feedforward network architecture which back propagates error and changes weights to minimize the error. Quickprop may be viewed as a fast version of the well known back-propagation algorithm. The learning algorithms are further discussed in the following sections.

The data used in this study was generated by the Srcrcs tool developed by the Syracuse Research Corp. The data used simulates a wide band or high range resolution radar (HRR) with a 20 meter range window, consisting of 256 discrete points. The magnitude of the return at each point corresponds to the radar cross section at a specific range on the aircraft. There are seven typical airplanes available for training and testing, the F4, F14, F15, F16, F18, T38 and Lear. The elevation and aspect angle may be chosen by the user when generating the return. In this study we have been working with no elevation, only changes in aspect angle.

II Objectives of the Research Effort

The goals of this research study are described in this brief section. The first is to be able to distinguish among airplanes based upon high resolution radar returns using some machine learning algorithm which requires minimal set up. Tied up with that objective is the necessity of determining an appropriate representation of the data for the learning algorithm.

A second objective is to determine the aspect angle of the plane that is in the

radar return. Another goal is to have training times and the storage required within reasonable ranges. A further goal is to determine what factors are important in the recognition process and gain some understanding of how it is being done by the algorithm.

III SC-net

This section briefly describes a hybrid connectionist, symbolic approach to learning the type of information that could be used in rule-based expert systems or sub-symbolic information. The system does its learning from examples. They are encoded in much the same way that examples to connectionist systems would be presented. The exceptions are due to our variable representation. The system can learn concepts where imprecision is involved. The network representation allows for variables in the form of attribute, value pairs to be used. Relational comparators are supported. Symbolic rules can be generated from the learned network configuration.

Learning may be simply described by the following. It uses a network structure, which is configured based on the distinct examples presented to the system. That is, the network structure is <u>grown</u> based on the distinctly new examples seen by the system. Initially, only input and output cells need to be specified. For examples which resemble others previously seen, bias values of cells in the network are adjusted. The system can learn incrementally. A complete description of learning in SC-net can be found in [1]. Rules may also be directly encoded in the network [9].

The system will explicitly indicate any contradictions, and patterns that have already been seen, which has proven useful in this research.

a SC-net – A Fuzzy Connectionist Expert System

A connectionist model is a network, which in its simplest format has no feedback loops. It consists of three types of cells (input, output, and hidden cells). Every cell has a bias associated with it, which lies on the real number scale. Cells are

connected through links which have weights associated with them. In the SC-net model of a connectionist network, each cell can take on an activation value within the range [0..1]. This corresponds to the fuzzy membership values of fuzzy sets. The uncertainty handling constructs come from fuzzy set theory.

In fuzzy logic one may define disjunction (fOR) as the maximum operation, conjunction (fAND) as the minimum operation and complement (fNOT) as strong negation. Since fOR and fAND are defined as maximum and minimum operations, we let certain cells act as max and min functions, in order to provide for the above operators. In order to be able to distinguish cells as modeling the min (fAND) or the max (fOR) function we use the sign of the bias of a cell to determine which of the two functions is to be modeled. Furthermore, we denote a bias value of zero to indicate when a cell should operate as an inverter (fNOT).

b The Network Structure

We can think of every cell in a network accommodating n inputs I_n with associated weights CW_n. Every cell contains a bias value, which indicates what type of fuzzy function a cell models, and its absolute value represents the rule-range. Every cell C_i with a cell activation of CA_i (except for input cells) computes its new cell activation according to the formula given in Figure 1. If cell C_i (with CA_i) and cell C_j (with CA_j) are connected then the weight of the connecting link is given as $CW_{i,j}$, otherwise $CW_{i,j} = 0$. Note, an activation value outside the given range is truncated. An activation of 0 indicates no presence, 0.5 indicates unknown and 1 indicates true. In the initial topology, an extra layer of two cells (denoted as the positive and the negative cell) is placed before every output cell. These two cells will be collecting information for (positive cell) and against the presence of a conclusion (negative cell). These collecting cells are connected to every output cell, and every concluding intermediate cell (these are cells defined by the user in the SC-net program specification). The final cell activation for the concluding cell is given as:

CA_i – cell activation for cell C_i, CA_i in $[0..1]$.

$CW_{i,j}$ – weight for connection between cell C_i and C_j, $CW_{i,j}$ in R.

CB_i – cell bias for cell C_i, CB_i in $[-1..+1]$.

$$CA'_i = \begin{cases} min_{j=0,...,i-1,i+1,...n}(CA_j * CW_{i,j}) * |CB_i| & CB_i < 0 \\ max_{j=0,...,i-1,i+1,...n}(CA_j * CW_{i,j}) * |CB_i| & CB_i > 0 \\ 1 - (CA_j * CW_{i,j}) & CB_i = 0 \text{ and } CW_{i,j} \neq 0 \end{cases}$$

Figure 1: Cell activation formula

CAoutput=CApositive_cell+CAnegative_cell-0.5 .

Note, the use of the cell labeled UK (unknown cell) in Figure 2. This cell always propagates a fixed activation of 0.5 and, therefore, acts on the positive and the negative cells as a threshold.

The positive cell will only propagate an activation $>=$ 0.5, whereas the negative cell will propagate an activation of $<=$ 0.5. Whenever there is a contradiction in the derivation of a conclusion, this fact will be represented in a final cell activation close to 0.5. For example, if CApositive_cell=0.9 and CAnegative_cell=0.1, then CAoutput=0.5, which means it is unknown. If either CApositive_cell or CAnegative_cell is equal to 0.5, then CAoutput will be equal to the others cell activation (indicating that no contradiction is present).

IV Quickprop

Quickprop [2] was developed by Fahlman at CMU. It is a back-propagation neural network architecture. Hence, it is a feedforward network with a learning scheme that back-propagates the error and then minimizes the error by modifying the weights. We have been using a 3-layer (one hidden layer) network representation. The system has been shown to be up to an order of magnitude faster than standard back-propagation [3]. In other tests that we have run, it has shown itself to also be

significantly faster than back-propagation with comparable accuracy after learning [2]. Quickprop has been the connectionist network of choice in this study due to its speed advantages.

There are several factors which make it fast. One is that it uses a small constant (0.1) added to the output derivative when back propagating error. This is useful because the derivative $o(1 - o)$ is 0 at $o \in \{0, 1\}$. It allows for error to propagate through these points. It also uses the hyberbolic arctangent function to propagate the error back. For small error this is linear, but for large error it approaches infinity. It is never allowed to be larger than plus or minus 17. This contributes to larger step sizes.

The algorithm uses information about the previous step size and direction. It also uses two risky assumptions. The first is that the error versus weight curve can be approximated by a parabola whose arms open upward. The other is that the change in the slope of the error curve, as seen by each weight is not affected by all the other weights that are changing at the same time [2]. In calculating new weights, for each weight the previous and current weight error slopes and the weight change between the points where the slopes were measured are used to determine a parabola. Then a jump is made directly to the minimum point of the parabola.

The above is the general idea behind the bigger step sizes in Quickprop. The technique clearly has some possible difficulties, like taking infinite step sizes. For complete information on how this problem and others are dealt with the reader is referred to Fahlman's paper [2].

It is necessary to note that like standard back-propagation algorithms, there is no guarantee that Quickprop will converge for any given problem. Convergence, for example on the exclusive or problem, will often depend on the choice of initial weights. In Fahlman's experiments it did not diverge significantly more than back-propagation.

V Representation

A key to be able to effectively apply machine learning techniques is the representation given to the problem [10]. Three different representations have been used in this study with mixed results. The first two methods are geometric hashing schemes [7]. They are designed to provide some invariance to the fact that the range profile will not always be aligned in any standard fashion, can be corrupted by noise and will be compressed and expanded due to changes in aspect angle. At this point in should be noted that this study has involved planes flying level. That is they have a zero elevation angle. Only changes in aspect angle are considered here.

The first representation scheme tried was a geometric hashing scheme. The 256 input points were **binned** into 64 bins each of 4 points. The bins were normalized so that the minimum bin served as 0. Next the peak points were found. Hashing is then done. Master and slave points are chosen. The master will be placed at 0 the slave at 1 and the other binned points placed in scale as appropriate. The output is then normalized so that the maximum height is 1. This provides 127 inputs, 63 to the left of 0 and 63 to its right.

The second representation scheme was a modification of the first. Now the data is binned into 51 bins of points each. A cosine curve is placed over each of the peaks. These curves are added together in order to capture the overlap of peaks. Hashing is done as before except that the points are not shifted to the left of the master. This provides us with 51 inputs.

Depending on the number of masters and slaves chosen, each of the above techniques will produce more than one hashed pattern for a given radar image. In fact the second scheme will often produce over 100 patterns, where the first might produce a few more than 20.

The third representation scheme initially normalizes the data to be between 0 and 1. The lowest value returns are set to 0 and the highest set to 1 with the others

appropriately scaled. Then bins of size 10 are made (currently ignoring the last six points which are usually zero, but must be used in general). An average is calculated for these bins and the 25 average values serve as the input to the learning algorithm. This scheme is clearly simpler than the first two and does not take into account the fact that the scatterers must be lined up. However, it is simply an attempt to find a representation which is a good discriminator. In the results section it will be shown that this representation has some very good characteristics.

VI Results

In this section, we will discuss the results from both SC-net and Quickprop. The results will be presented by the representation with which they were derived. This is also the chronological order in which this study was carried out.

a Results from the first Hashing Representation

The first geometric hashing method was initially used with data on a Lear and F16 around the 0° aspect angle. The Lear and F16 were chosen since it was felt they would be clearly distinguishable. It was necessary to determine the number of hidden layers and the number of cells in those layers. It is known [5] that one hidden layer is all that is necessary in order to approximate arbitrary functions in algorithms like Quickprop. Hence, we have used one hidden layer, three total layers, in all experiments. The results are summarized below in Table 1. Training is done at 0 degrees.

The training examples were clearly distinguishable. The error reported is the mean square error over the outputs. In this case, there are 12 outputs. They consist of an angle from 0 to 45 at 5 degree increments and the two possible planes. From the results it is clear that it was difficult to recognize the F16. The really bad part of this is that the F16 was identified strongly as a Lear (often with an output of 1, the maximum). It usually got the Lear correct. However, it is unclear if everything

Plane	Degree	Number recognized	hidden units	Train error	epochs
F16	1	$\frac{3}{16}$	19	.003	150
F16	1	$\frac{7}{16}$	19	.000695	250
F16	1	$\frac{0}{16}$	4	.055	281
F16	1	$\frac{9}{16}$	14	.000229	400
F16	1	$\frac{2}{16}$	14	.000174	600
Lear	1	$\frac{26}{30}$	14	.000217	500

Table 1: Results on testing Lear vs. F16 around 0 degree aspect angle

Plane	Degree	Number recognized	hidden units	Train error	epochs
F16	16	$\frac{14}{28}$	12	.000486	251
F16	16	$\frac{13}{28}$	12	.000442	351
F16	16	$\frac{14}{28}$	12	.000317	451
F16	16	$\frac{10}{28}$	8	.00428	350
F16	16	$\frac{6}{28}$	8	.00712	300
Lear	16	$\frac{25}{30}$	12	.00114	251
Lear	16	$\frac{26}{30}$	8	.00712	300

Table 2: Results on testing Lear vs. F16 around the 15 degree aspect angle

is being seen as a Lear with just 2 planes.

Since it was felt that around 0 degrees might be an unusually troublesome spot, we tried some experiments with the Lear and F16 at 15 and 16 degrees. The results are summarized in Table 2. Again the results are weak on the F16, but good on the Lear. It is also the case that misclassified F16's are strongly classified as Lears.

Another experiment was tried it which 3 hashed F16 and Lear patterns were randomly left out of the train set at 15 degrees. After training they were presented and all three Lears were recognized, but **two of three** F16's were misclassified (strongly). The other was barely recognized. This furthered our belief that the

hashing scheme might be a difficulty.

Next, noise was added to the 0 inputs of which there were quite a few. For any input below .0002, noise (a value from 0 to 0.05) was randomly added. This slowed convergence, as it was done each epoch. After 110 epochs with 8 intermediate nodes, training the Lear and F16 at 0 degrees, we found at 1 degree only 4 of 16 F16's could be recognized. Again they were for the most part strongly assigned as Lears.

With SC-net using the original binned and hashed data several experiments were conducted. Three different methods of representing the hashed data to SC-net were investigated. The results of each of the methods are listed below:

Plain Version (No variables):

- Lear and F16 jet trained at 0 degrees aspect angle

  Obtained (at 1 degree angle):

  92.9 % accuracy in Lear using Lear data. 62.5 % accuracy in F16 using F16 data.

- Lear and F16 jet trained a 15 degrees aspect angle.

  Obtained (at 16 degree angle):

  93.3 % accuracy in Lear using Lear data. 64.3 % accuracy in F16 using F16 data.

- Lear, F14, F15 and F16 jets trained at 15 degrees aspect angle.

  Obtained (at 16 degree angle):

  100% accuracy in F14 using F14 data.

  Obtained (at 17 degree angle):

  75 % accuracy in F14 using F14 data.

  Obtained (at 18 degree angle):

  100 % accuracy in F14 using F14 data.

Using Fuzzy Variables. This method was used to improve the recognition of the

F16 at 1 and 16 degrees aspect angle. The following partioning of the fuzzy variables was used:

very_low: 0.0 - 0.2 (0.0, 0.6)

low: 0.2 - 0.4 (0.0, 0.7)

medium: 0.4 - 0.6 (0.1, 0.9)

high: 0.6 - 0.8 (0.3, 1.0)

very_high: 0.8 - 1.0 (0.5, 1.0)

The above ranges are explained as follows by an example. Very_low is from 0 to 0.2, but it still has some belief up until 0.6 tapering off to 0 [9]. This representation allows for overlapping fuzziness in the recognition space. Then we trained the Lear and F16 jet at 15 degree angle:

We obtained **at 16 degree angle** a 95.8% accuracy in F16 using F16 data. We obtained **at 1 degree angle** 56% accuracy in F16 using F16 data.

The third method was to use discrete data to generate rules. By looking at discrete points of the data it is possible to generate rules. These rules will 100% correctly classify all the examples in the training set, by using only a small subset of all the data points. We will next identify some of the advantages in using this scheme.

- Give symbolic meaning to the radar returns, which may help identify possible reasons for the problems encountered with the F16 and the Lear jet.

- If the discrete data version gives results similar to the other two versions overall results and the results are good, rules can be easily ported onto other systems for practical use (both in speed and hardware requirements when designing an actual system).

- This method allows compression of the network to a more suitable size, since many connections and cells will be redundant and can be eliminated, again

yielding memory and speed improvements.

Performance was still weak, making the rules of no use. At this point, it appeared that some change in the representation was needed.

b   The second hashing scheme

In this case with the Quickprop system, we tried the F16 and Lear at 15 degree for the training set. In this case there were 113 Lear patterns and 26 F16 patterns. The network was unable to reduce the error to a reasonable level. It had most of its outputs in the train data incorrect well into the training. We then tried out the same problem with SC-net. It correctly classified the Lears, but mis-classified 5 of the F16s in the train data as Lears and could not get 15 of the others distinguished from the Lear. Based on these results and some plots of the hashed data it was clear that the patterns were being obscured by the hash process.

c   The revised representation

It has been the case that the results from Quickprop and SC-net have been very similar with SC-net providing slightly better results. Most of the results from this representation are from SC-net. However, we first report on the results from Quickprop.

First we trained with the F16 and Lear from 10 to 20 degrees on the odd degrees. Testing was then done with the even degree patterns. This resulted in a 100% correct recognition all 10 of the test patterns. The lowest recognition was 0.93.

The next experiment was with the F14, F15, F16, F18 and Lear. Training was done at 0, 5 and 10 degrees. Then all the planes at the other degrees made up the test set. Four hidden units were used and 101 epochs were trained for an error of .00054. Again recognition was 100% on the 40 test patterns.

The following are the extensive SC-net simulation results from this representation. Tests were again conducted using the Lear and the F16 data. Prior to this

they had presented lots of problems in obtaining correct classifications. The following experiments were conducted:

(a) The Lear and the F16 data was trained on all even points within the range of 10 and 20 degrees inclusive. Then, both planes were tested on all odd degrees.

Result: 100% accuracy on the unseen points was reached for both planes.

(b) In the second experiment we only trained for both planes at 10, 14, and 18 degrees. We then tested on the remaining degrees.

Result: Again 100% accuracy was obtained. In some cases the belief provided was lower, but the two planes were clearly distinguishable.

(c) In the next experiment we only trained at 10 and 20 degrees for both planes and tested on the remaining.

Result: Again 100% accuracy was achieved. In some occasions the certainty would again be lower, but still the planes were distinguishable.

The above results were extremely encouraging and prompted us to further investigate the behavior at the 0 - 10 degree range. In prior attempts it was impossible to distinguish an F16 from a Lear at 1 degree after it was trained at 0 degree. We conducted the following experiments:

(d) We trained the network for both planes at 0, 5, 10, 15, and 20 degrees. Then we tested on all remaining points.

Result: Like in the previous cases we were able to obtain 100% accuracy.

(e) In our final experiment we added the data of an F14. In this test we trained a Lear, an F16, and an F14 at 15 degrees and tested if the F14 could be correctly classified at 16, 17, and 18 degrees.

Result: Again we obtained 100% accuracy, in fact the results were extremely in favor of the F14, even though only one training example was provided for the F14.

In the final experiment a set of 7 planes were chosen for simulation. The Srcrcs simulator allows generation of radar return data for 8 planes one of these planes being the 747 Jumbo. We have excluded this plane from our study, since it appears

to be quite different from all the other planes. The following 7 planes were used for testing:

a) F4 Jet b) F14 Jet c) F15 Jet d) F16 Jet e) F18 Jet f) Lear Jet g) T38 Jet

To test how good the data was in allowing generalization, we decided to train for all 7 planes at 0, 5, 10, 15, 20, 25, 30, 35, 40, and 45 degrees. Testing was on all other degrees within the range [0,45]. Since there are a total of 322 instances of data and 70 were used for training, a test and training size of 79% and 21% respectively were used. In almost all instances a correct response was obtained if the best choice was selected. Problem points existed at 16 degrees were the system could not distinguish between a Lear and an F15 Jet given the Lear data for testing. At 17 degrees the problem was between a Lear and a T38 given the Lear data. The certainties were low in both instances, indicating uncertainty in making the proper classification. Two additional points of uncertainty were at 41 degrees. This time SC-net had difficulties to distinguish between an F15 and an F14 given the F15 data, and between an F16 and an F14 given the F16 data. Counting these four as miss-classifications SC-net obtained an overall classification accuracy of 98.5%. This is a favorable result, which is an improvement over results reported in [1].

VII Recommendations

The above results have shown that the wide-band data efficiently binned can be used for effective training with SC-net. It is also shown that with the new representation performance is good in Quickprop, also. It is expected that the results in Quickprop will continue to parallel those of SC-net. The new data is rather small (only 25 features) and there exists only one example of a plane for every degree. This stands in contrast to the prior method which sometimes generated dozens of hash patterns for every degree. Even though, we were able to generalize from rather different aspect angles, another problem still remains. The wide-band data provided by the

simulation tool shows the radar-return of the plane taken from an exact center spot. In reality the image will be shifted. In this case it is expected that the wide-band data will also be shifted. This will represent serious problems for the binning. One possible improvement of the current algorithm is to start the binning process from the highest peak of the return, by making it the center of the newly formed return. In this method a sort of ordering of the highest returns is achieved. This method could improve making the generated data be more invariant to the shifting of the peaks for different points. Additionally the relative magnitude of all the returns will be preserved, which could lend itself as another factor when deciding what plane is represented.

It is clear that the representation scheme is very important in the recognition process. With the good results in the artificial data, we feel that real data should be examined. It will have the problems of shifting and noise. Additionally, there will likely be other issues. Examination of it will help serve to guide the research in the proper direction. The current results are encouraging. They indicate that recognition **may** be done with relatively small networks. This will allow shorter training times and more patterns to be addressed.

We have begun to look at the ARTI data, which is actual data with the planes disguised. The results are very preliminary. It has been necessary to align the data. The results have been to get 17 of 19 planes correct in a 3 degree aspect arc around 180 degrees. One training example is used of each of the two planes in the set. Again a problem is representation and it clearly needs to be investigated further.

The symbolic features of SC-net have been helpful in explaining what is happening with the different representations. The ability to use variables allows one to try applying different meanings to ranges of input data in a simple fashion. Rule generation can show how inputs interact with outputs, even in cases where there is no (clear) symbolic meaning. It is felt that these types of features are important for connectionist models in this (and other) domains. Further investigation along this

line for both Quickprop and SC-net can be useful.

It can be a sort of luxury, due to time factors, to use more than one learning algorithm. However, we have found this valuable. Each has different characteristics which can bring to light distinct problems with representation, the data or type of algorithm. It is especially useful in a case where the characterization of the data is not clear. Of course, the characterization of many learning algorithms is incomplete, also. The comparisons can also shed light on the algorithms characteristics.

Learning in this domain can be very time-consuming in terms of CPU-time. It requires good, stable, fast computing facilities. Investigations into methods which are efficient in the learning process in terms of time and space are also of import to this research. Quickprop is faster than back-propagation schemes and SC-net is in turn faster than either of those. However, there is certainly much room for improvement and research in this area is well warranted.

Intuitively, it would seem that the aspect angle of the plane may be recovered at least somewhat. Training at every 5 degrees it is unknown where a plane at 3 degrees, for example, would be put. Investigation along these lines remains to be done. Due to early basic recognition problems it was ignored.

The results of this study provide us with optimism that connectionist models can be effectively used to solve the problem of recognizing airplanes from their radar profiles. However, there is significant research in the areas discussed here (and others that may arise) to be done.

References

[1] Atkins, R.G., Shin, R.T., Kong, J.A. (1989), A Neural Net Method for High Range Resolution Target Classification, Report from Dept. Of EE and CS and Research Lab. of Electronics, MIT, Cambridge, MA.

[2] Fahlman, S.E. (1988), Faster-Learning Variations on Back-Propagation: An Empirical Study, Proceedings of 1988 Connectionist Summer School, pp.38-51.

[3] Gorman, R.P. and Sejnowski, T.J. (1988), Analysis of Hidden Units in a Layered Network Trained to Classify Sonar Targets, Neural Networks, Vol. 1, pp. 75-89.

[4] Hall, L.O. and Romaniuk, S.G. (1990), A Hybrid, Connectionist, Symbolic Learning System, AAAI-90, Boston, Ma. August.

[5] Hecht-Nielsen, R. (1987), Kolmogorov's Mapping Neural Network Existence Theorem. Proceedings of IEEE First International Conference on Neural Networks, San Diego, CA., June.

[6] Hinton, G. (1987), Connectionist Learning Procedures, Artificial Intelligence.

[7] Lamdan, Yehezkel and Wolfson, Haim J., Geometric Hashing: A General and Efficient Model-Based Recognition Scheme, Robotics Research Lab, Courant Institute of Mathematical Sciences, NYU, New York.

[8] Michalski, R. S., Carbonell, J. G., Mitchell, T. M. 1983. Machine Learning: An Artificial Intelligence Approach. Palo Alto, Ca., Tioga Publishing.

[9] Romaniuk, S.G. and Hall, L.O. (1990), SC-net: A Hybrid Connectionist, Symbolic System, Tech. Report Dept. of CSE, USF, Tampa, Fl. (In review).

[10] Shavlik, J.W., Mooney, R.J. and Towell, G.G. (1989), Symbolic and Neural Learning Algorithms: An Experimental Comparison, Computer Sciences Technical Report #857, University of Wisconsin-Madison, Madison, Wisc.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

<u>FINAL REPORT</u>

<u>MODEL FOR CHARACTERIZING A DIRECTIONAL COUPLER BASED</u>
<u>OPTICAL HETERODYNE DETECTION SYSTEM</u>

Prepared by: Mohammad A. Karim

Academic Rank: Associate Professor

Department and   Department of Electrical Engineering &

                    The Center for Electro-Optics

    University: The University of Dayton

Research Location: WRDC/AARI-2

                    Wright-Patterson AFB; Dayton; Ohio

USAF Researcher: Lawrence E. Myers

Date: 28 September 1990

Contract Number: F49620-88-C-0053

MODEL FOR CHARACTERIZING A DIRECTIONAL COUPLER BASED OPTICAL HETERODYNE DETECTION SYSTEM

by

Mohammad A. Karim

ABSTRACT

This report summarizes the research performed during the USAF-UES Summer Faculty Research Program. The work involved developing an analytical model for characterizing a directional coupler based optical heterodyne (coherent) detection system. The coherent detection system in question consists of two fiber optic links carrying respectively optical signal and local oscillator beams which are then combined by means of a directional coupler. The directional coupler based heterodyning scheme is compared with that based on Y-coupler as well as that based on only beam splitter in terms of their signal-to-noise ratios. The current analytical and simulation results along with those expected to be generated through a follow-up mini-grant study would be able to dictate the design characteristics of the most optimum directional coupler based coherent detection system.

ACKNOWLEDGEMENTS

I. INTRODUCTION

Coherent (heterodyne) detection (Karim, 1990) is a very powerful technique for the sensing of optically narrowband radiation at the quantum noise limit. In particular, optical real-time beam steering devices such as those based on liquid crystals on nonlinear refracting material (Karim, 1988 and 1989) can be anticipated to bring in target signatures and thereafter the target signatures can be combined with certain laser transmitter output at a receiver. Such processing of signals at a receiver is accomplished by means of heterodyne schemes. The mixing of the two beams, however, can be carried out in two ways: (a) in free-space by means of a beam splitter; and (b) using fibers and either a directional coupler or a Y-coupler.

The WRDC/AARI-2 Division of the Wright-Patterson AFB is currently interested in developing a reliable heterodyne scheme for mixing steered beam and laser transmitter beam preferably using fiber based systems. As a first attempt, one may develop model for characterizing the performance of a directional coupler based heterodyne system where one may limit the analysis to only step-index fiber and single mode operation. Furthermore, in this study, the target signature is now assumed to be an ideal one for the sake of simplicity. It has already been shown that the field distribution of optical beams as well as detector uniformity play significant roles in determining the efficiency of the corresponding heterodyne system (Cohen, 1975; Fink, 1975; Fink and Vodopia, 1976). The problems associated with the multimode Y-coupler are definitely more cumbersome and, therefore, the corresponding modeling may involve more difficult analysis.

One of my current research interests at the University of Dayton has been directed towards exploring the various possibilities of designing and characterizing optical systems for generating, transforming, shifting, or modifying laser beams. In addition, I had also developed analytical as well as software tools for verifying liquid crystal based beam agility devices for WRDC/AARI-2 during my 1988 SFRP work as well as for nonlinear refracting material based E-O scanner for WRDC/AARI-3 during my 1989 SFRP work. These experiences have served as motivation for the current SFRP work that also consists of developing analytical tools but for characterizing a heterodyne system.

II. OBJECTIVES OF THE RESEARCH EFFORT

There has been an ever growing interest in achieving proper heterodyning of steered optical beams. Several vendors have already delivered prototype of several liquid crystal based beam steering devices to my sponsoring laboratory. One of the ways that such and other beam steering devices could be used would be to collect target signature and then mix it with laser transmitter beam in an optical heterodyne configuration. One of the means of heterodyne scheme may involve the use of a directional coupler for efficient mixing of the beams at the detector. My assignment as the SFRP Fellow was to perform basic research into the theory and characterizations of a directional coupler based heterodyne system shown in Fig. 1. The system consists of a combination of a the beam steering device, an acousto-optic modulator, beam splitters and a variable fiber coupler. The advantage of having such a set-up is that under suitable operating condition, the SNR may be

reasonably small.

The objectives of this research are three-fold: (a) to develop model for describing the performance of the aforementioned coherent detection system; (b) to develop theory to identify which of the two possible couplers (Y-coupler versus directional coupler) is more desirable; and (c) to develop software tool for studying the performance of the said system at least for a single mode.

III. MODEL DEVELOPMENT

In principle, heterodyning at light frequencies is the same as that at much lower frequencies. For the conventional receiver, voltages at two different frequencies are applied to a nonlinear circuit element, such as a diode, which acts as the translator. Current at the difference frequency is recovered from the output of the device and is amplified and processed. Similarly, at optical wavelengths two electric field at different frequencies are applied to a photodetector, and current at the difference frequency is recovered from the output. This technique provides a signal that is an exact replica of the original but moved down to some lower frequency range in which amplification, filtering, and detection are easier to achieve.

Fig. 1 shows the system wherein an input light of frequency f_i is mixed with a local oscillator beam of frequency f_1. The background light that may be present is represented by many small components, spread over certain band, and each having some amplitude A_b and frequency f_b. The local oscillator and input signal can be represented respectively as $E_1 = A_1 \cos w_o t$ and $E_i = A_i$

cos $w_i t$ where w_x represents the angular frequency of beam x. The total background light can be represented as $E_b = \sum_b A_b \cos w_b t$. A_b and w_b have values for every component of light present in the background. The current at the output of the square-law detector is thus given by

$$i = c (A_i \cos w_i t + A_1 \cos w_1 t + A_b \cos w_b t)^2. \qquad (1)$$

where c is a constant. As a result of the squaring operation, Eq. (1) will result into cross products between the various terms, and we shall obtain components involving the sums and differences of these frequencies. The sum frequencies are much too high to be passed by the detector and can, therefore, be neglected. The resulting expression for the current is given by

$$i = c [(A_i^2 + A_1^2 + \sum_b A_b^2)/2 + A_i A_1 \cos(w_i - w_1)t] \qquad (2)$$

In Eq. (2), all terms resulting from cross products, or beats, between the $\sum A_b \cos w_b t$ components and other components have been neglected. These beats are negligibly small except in unusual circumstances given that the heterodyne circuit can provide excellent frequency filtering as well as critical spatial discrimination. Eq. (2) consists of direct current (dc) terms resulting from the local oscillator, the signal and the total background light and is equivalent to $i_{dc} = c(P_i + P_1 + P_b)$ where P_1 is the local oscillator power, P_i is input signal power, and P_b is the total background power. The cross term represents the beat between the local oscillator and signal and is the desired intermediate-frequency (IF) carrier.

In addition to the dc and the desired IF current, there is the shot noise that accompanies the dc which is related to P_i, P_1, and P_b by

$$(i_{sn})2 = 2\ ce(P_i + P_1 + P_b)(BW)$$
$$= (2e^2(BW)/hf)(P_i + P_1 + P_b) \tag{3}$$

where e is the electronic charge, h is Planck's constant, f is the light frequency and BW is the IF bandwidth.

The thermal noise at the input to the IF amplifier is also important in the same way that such noise is important at the input to the signal amplifier when direct detection is employed. This noise power (Karim, 1990) is given by FkT(BW), where F is the noise figure of the amplifier, k is Boltzmann's constant, and T is the absolute temperature. Obviously, the best heterodyne receiver performance will be obtained only when the noise figure F is the lowest.

The IF signal-to-noise ratio SNR may be obtained to give:

$$SNR = 2\ c^2 P_i P_1 / [(2ce(P_i + P_1 + P_b) + Fkt)BW] \tag{4}$$

The noise includes total shot noise given by Eq. (3), and thermal noise. Furthermore, for practical receivers P_1 can usually be made much greater than $P_i + P_b$ and, in fact, can be large that the thermal noise may be insignificant. Under these conditions, the only noise of consequence is the shot noise produced by the local oscillator. Thus,

$$SNR = n^* P_i / (hf(BW)) \tag{5}$$

where n^* is an efficiency factor. The IF signal power is propor-

tional to the product of local-oscillator light power and signal light power. For the direct detector, however, the power of the recovered signal is proportional to the square of the signal light power only. Thus, when compared to the direct detection output, the heterodyne scheme provides a gain proportional to P_1/P_i.

From these considerations, we see that the heterodyne receiver provides several advantages. First, the conversion process provides gain so that the signal output of the detector may overwhelm both thermal and detector noise. Second, the heterodyne receiver provides excellent discrimination against background noises. And, finally, it allows for the possibility of recovering both phase-modulated or frequency-modulated signals.

Heterodyne conversion takes place over an area significantly larger than the order of a light wavelength. The maximum IF signal is obtained only when the signal and local-oscillator beams have the same phase relationship over the complete area of coincidence. Accordingly, the optical phase must be uniform over the complete wavefront of each beam. This requirement is met only under the following conditions:

1. The two beams must have the same mode structure (higher-order modes are undesirable, since they reduce the fiber bandwidth and introduce signal distortions);

2. The two beams must be coincident and, to maximize the SNR, their diameters must be equal;

3. The beams must propagate in the same direction;

4. The wavefronts must have the same curvature; and

5. The beams must be identically polarized.

The heterodyne receiver may be considered disadvantageous in that, for optimum performance, it must meet all of the above requirements we have set forth, some of which are very severe.

The SNR performance of the heterodyne system is influenced most particularly by n^* which may have contributions from several fiber optics consideration. Accordingly, n^* can be expressed as $n_1 n_2 n_3 n_4 n_5 n_6$ where n_1 is the fraction of the signal power that is actually coupled to the fiber, n_2 is the transmission of the waveguide, n_3 is the coupling efficiency into the photocurrent component oscillating at IF frequency, n_4 is the amount of output power actually coupled into the detector, n_5 is the bend loss factor and n_6 is the quantum efficiency of the photodetection process. However, n_1 is calculated from the product of another six launching efficiency components n_{11}, n_{12}, n_{13}, n_{14}, n_{15}, and n_{16}. The three spot-size related launching efficiencies (Marcuse 1970 and 1977; Kogelnik 1964) are: (a) n_{11} influenced by the variation of Gaussian spot size w_o, (b) n_{12} influenced by fundamental fiber mode spot size w_g, and (c) n_{13} influenced by the mismatch of two spot sizes while the three geometrical launching efficiencies (Karim, 1990) are (a) n_{14} caused by the transverse offset, (b) n_{15} caused by the angular offset, and (c) n_{16} caused by the longitudinal offset. It must be noted that n_6 is the same for all heterodyne systems irrespective of whether or not a directional coupler is used. Thus for proper comparison between different heterodyne systems, only the evaluation of n^*/n_6 is enough.

IV. SIMULATION

A computer program was written to evaluate n^*/n_6 for the directional coupler by having first determined n_{11}, n_{12}, n_{13}, n_{14}, n_{15}, n_{16}, n_2, n_3, n_4, and n_5 (Neumann, 1988). It may be reasonable to assume that n_2 to be approximately 1 since transmission attenuation can be as small as 0.2 dB/km (Karim, 1990), and likewise n_4 can be about 0.965 for an ideal fiber end-face, and n_5 is unity since the distance under consideration is very small. Fig. 2(a) shows the plot of n_{11} versus w_0/a where a is the core radius for different values of normalized frequency V, Fig. 2(b) shows the values of maximum n_{12} and optimum w_g/a versus V and Fig. 2(c) shows the plot of n_{13} versus w_0/w_g. On the other hand, Fig. 3 (a)-(c) respectively shows the three geometry related launching efficiencies n_{14}, n_{15}, and n_{16}. Fig. 4(a) shows the geometry of a directional coupler while Fig. 4(b) shows the values of n_3 versus wavelength for different values of coupler separation (Digonnet and Shaw, 1983). Fig. 5 shows n_5 for two of the most important modes LP_{01} and LP_{11} as functions of wavelength (Karim, 1990). Finally, Fig. 6 shows the plots of n^*/n_6 versus the three ratios z_w/z_r, $theta/theta_d$, and s/w_g for different values of w_0/a where z_w, theta, and s are the three offsets, z_r is Rayleigh distance and $theta_d$ is the divergence angle.

V. DISCUSSION AND RECOMMENDATIONS

Preliminary calculations made during this work is indicative of the fact that it is possible to use directional coupler in an heterodyne system as it may provide high SNRs provided proper geometrical and beam characteristics have been maintained.

Before the design characteristics of the directional coupler based heterodyne system is finally specified, a few more aspects need to be considered and investigated in the follow-up research. The follow-up mini-grant proposal is expected to address the following additional concerns:

(i)   consider the fact that the detectors are often nonuniform and it may contribute to detrimental consequences (Fink and Vodopia, 1976) that needs to be avoided;

(ii) account for the fact that the steered beams may have field profile that may result in a reduction of SNR (Fink, 1976; Paranto, 1988);

(iii) account for the fact that wavelength variation will play a role in determining system efficiencies (Digonnet and Shaw, 1983);

(iv) to account for the consequences of using an Y-coupler (Karim, 1990; Salzman, Sivan, Kapon, and Katzir, 1983) in the place of a directional coupler;

(v) to account for the fact that a directional coupler may not be strictly step-index type; and

(vi) to account for the realistic transmission characteristics of the available directional coupler.

An analysis of the sort described should be substantiated by experimental verification of the resulting heterodyne system.

VI. REFERENCES

Cohen, S. C., Heterodyne detection: phase front alignment, beam spot, and detector uniformity, Appl. Opt., Vol. 14, pp. 1953-

1959, 1975.

Digonnet, M., and Shaw, H. J., Wavelength multiplexing in single-mode fiber couplers, Appl. Opt., Vol. 22, pp. 484-491, 1983.

Fink, D., Coherent detection signal-to-noise ratio", Appl. Opt., Vol. 14, pp. 689-690, 1975.

Fink, D., and Vodopia, S. N., Coherent detection SNR of an array of detectors, Appl. Opt., Vol. 15, pp. 453-454, 1976.

Karim, M. A., Low voltage broadband beam steering devices using liquid crystals, Final Report to UES, AFSOR Contract No. F49620-87-R-0004, July 1988.

Karim, M. A., Analytical model of a unique electro-optic beam scanner, Final Report to UES, AFSOR Contract No. F49620-88-C-0053, September 1989.

Karim, M. A., Electro-Optical Devices and Systems, Boston, Massachusetts, PWS-KENT Pub. Co., January 1990.

Neumann, E. -G., Single-Mode Fibers Fundamentals, Springer-Verlag, Berlin, 1988.

Paranto, J. N., Investigation of heterodyne mixing efficiency variations due to optical misalignments and aberrations, MS Thesis, University of New Mexico, December 1988.

Salzman, J., Sivan, U., Kapon, E., and Katzir, A., Heterodyne detection using multimode waveguide Y-couplers, Appl. Opt., Vol. 22, pp. 3931-3934, 1983.
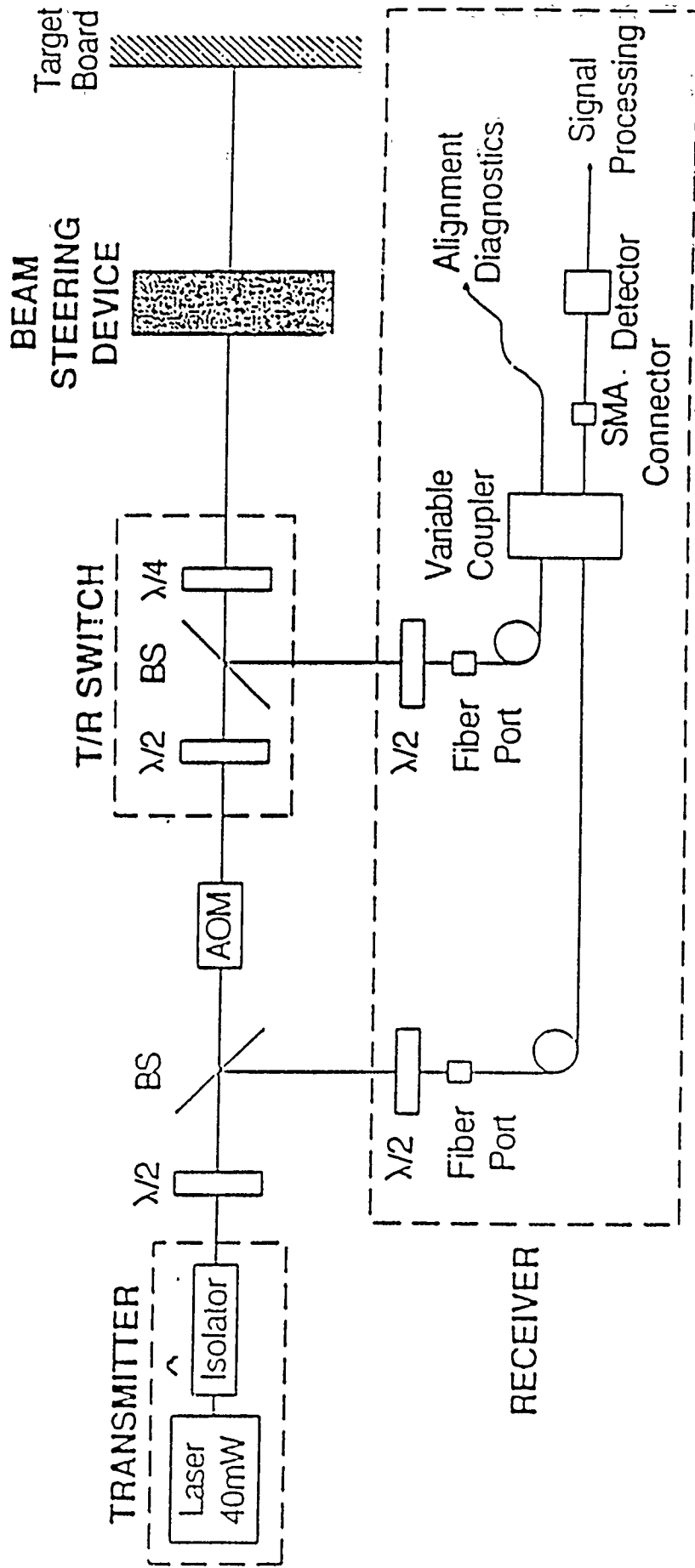
Fig. 1. The optical heterodyne system.
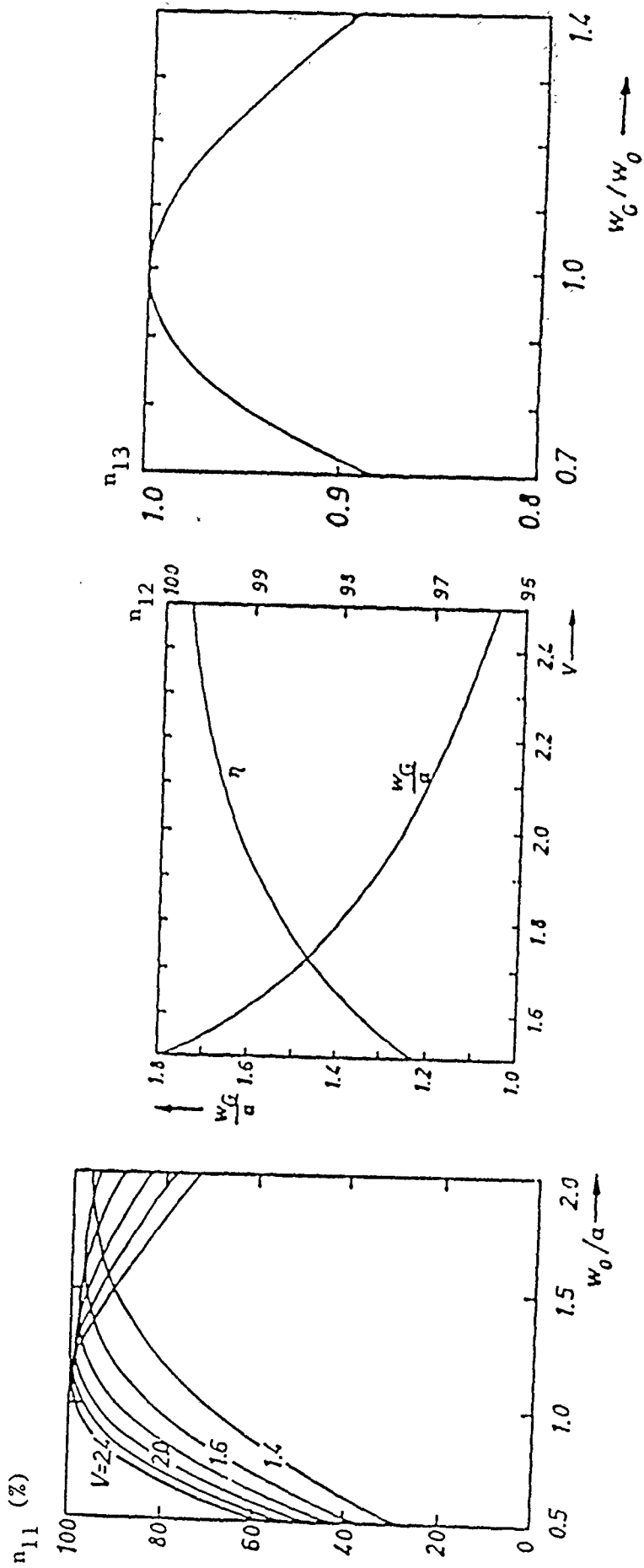
Fig. 2. (a) n_{11} versus w_o/a for different values of V; (b) n_{12} and optimum w_g/a versus V; and (c) n_{13} versus w_g/w_o.
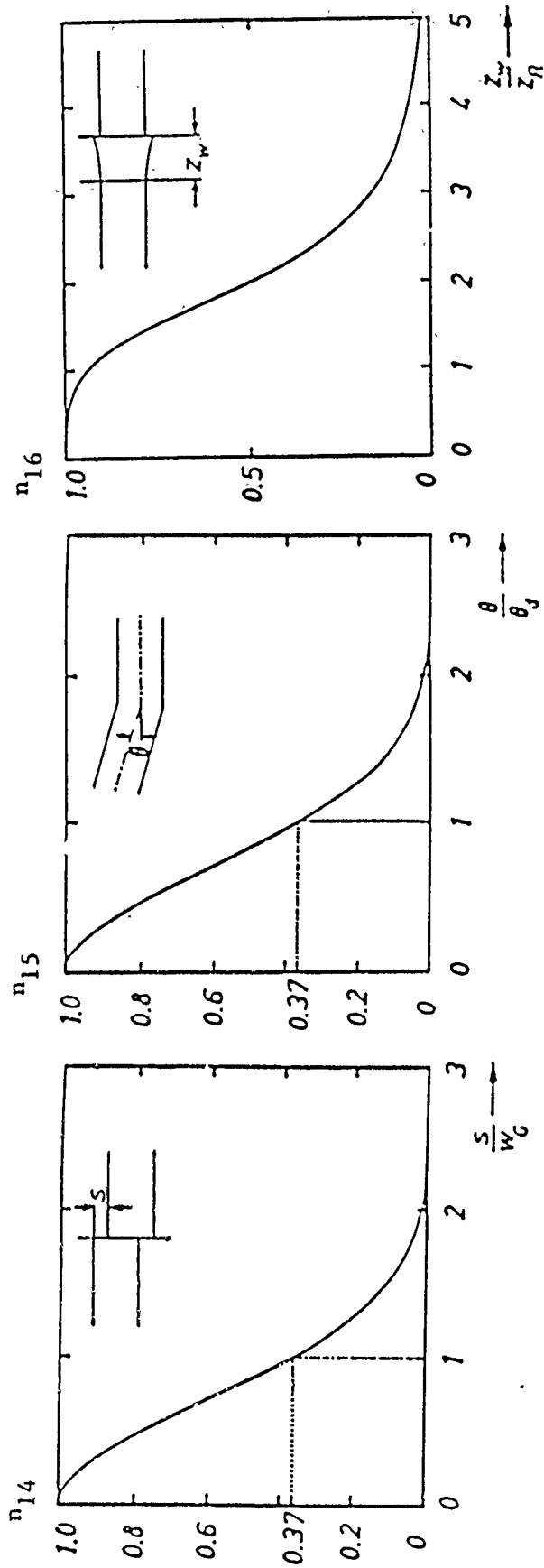
Fig. 3. (a) n_{14} versus s/w_g where s is the transverse offset; (b) n_{15} versus theta/theta$_d$ where theta and theta$_d$ are respectively the angular offset and divergence angle; and (c) n_{16} versus z_w/z_r where z_w and z_r are respectively the longitudinal offset and Rayleigh distance.
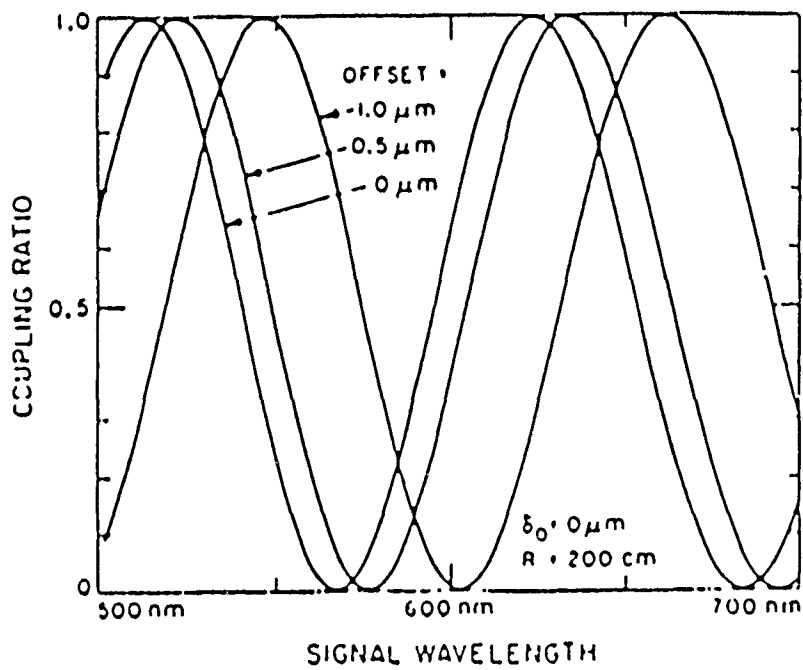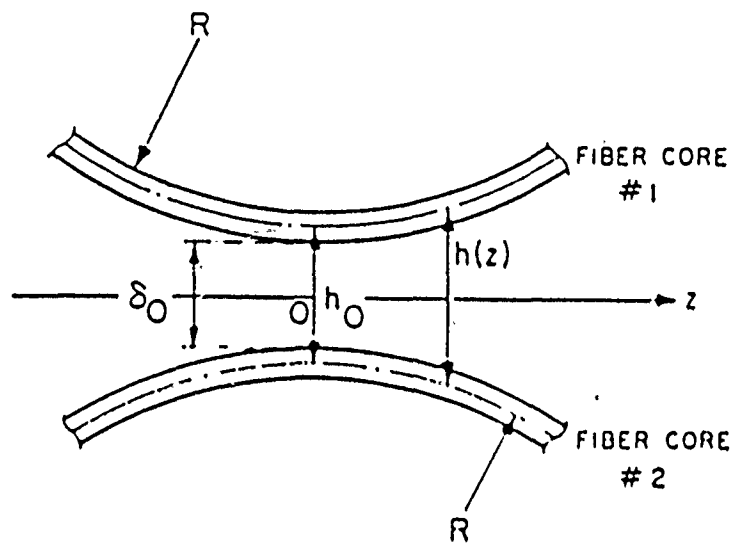
Fig. 4. (a) the directional coupler; and (b) n_3 versus wavelength for different coupler separations.
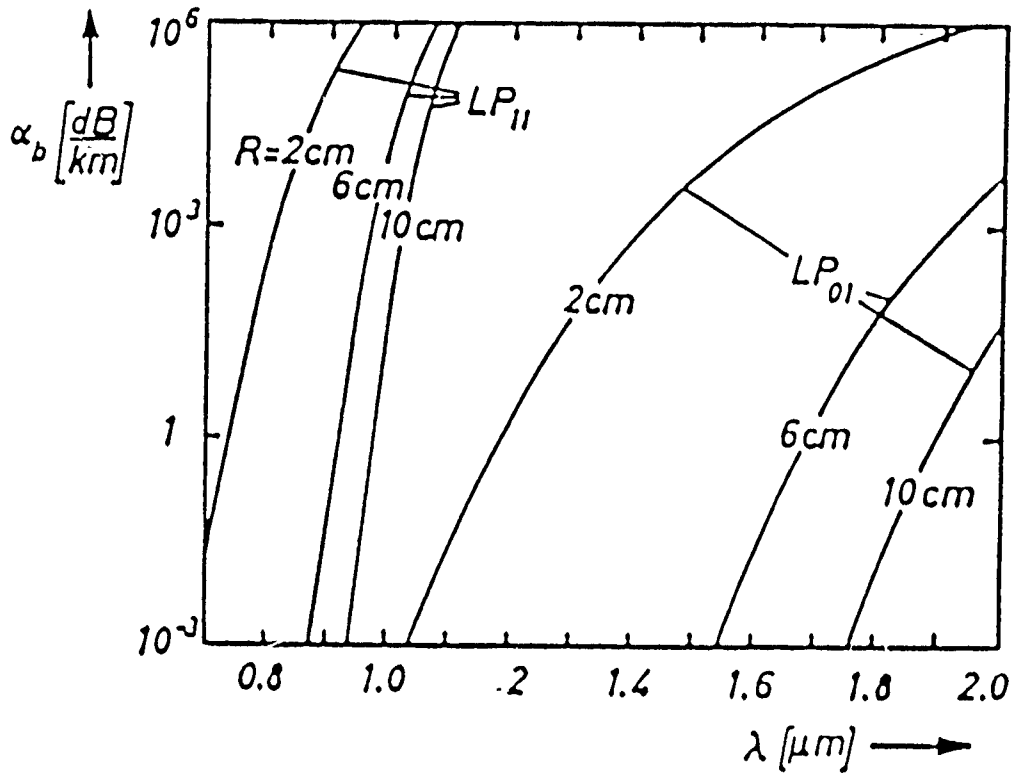
Fig. 5. n_5 versus wavelength for two modes LP_{01} and LP_{11} and different values of bend radius R.

Efficiency Factor
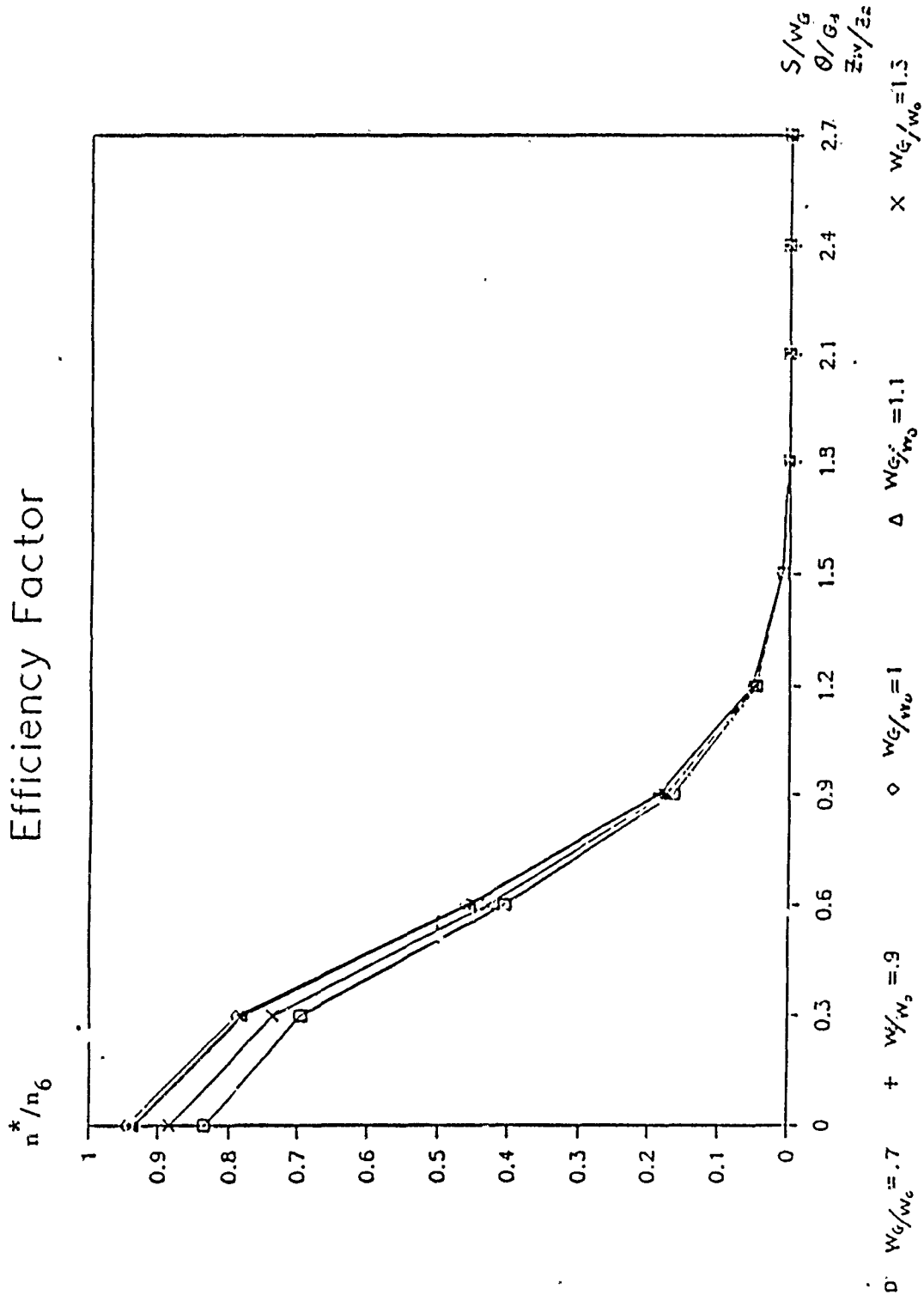
Fig. 6. The plot of the ratio n^*/n_6 versus z_w/z_r, theta/theta$_d$ and s/w_g for different values of the ratio w_g/w_o.

$\Box$ $w_G/w_o=.7$ + $w/w_o=.9$ $\diamond$ $w_G/w_o=1$ $\triangle$ $w_G/w_o=1.1$ $\times$ $w_G/w_o=1.3$

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM
GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Conducted by the
Universal Energy Systems, Inc.

FINAL REPORT

Context Dynamics in Neural Sequential Learning

Prepared by: Kevin G. Kirby, Ph.D.

Academic Rank: Assistant Professor

Department and   Computer Science and Engineering Department

 University: Wright State University

Research Location: AAAT-1
                     Wright Patterson AFB
                     Dayton, OH 45433

USAF Researcher: Louis Tamburino, Ph.D.

Date: September 30, 1990

Contract No: F49620-88-C-0053

Context Dynamics in Neural Sequential Learning

by

Kevin G. Kirby

ABSTRACT

A new neural architecture was developed for efficient learning of spatiotemporal dynamics. This architecture reduces the learning problem to two subproblems: (1) the formation of a "context" containing compressed input histories, and (2) the classification of context by an associational algorithm. The first subproblem was handled by introducing a nonlinear dynamical system into the neural network, which can be a low-connectivity random net or a continuous reaction-diffusion system. This enables the solution of the second subproblem to become simpler, requiring only a variant of the classical perceptron learning algorithm. A theoretical framework was developed in which the learning capabilities were analyzed in terms of finite automata theory. A computer simulation system was developed and used to show efficient learning of the sequential parity problem. Further simulations clarified the role of the context subsystem and demonstrated promising non-connectionist architectures for this problem.

Acknowledgements

I. INTRODUCTION

One goal of machine learning research is to develop a system which can make observations of a complex, unknown environment, and form its own internal model to make predictions of the environment's responses to various disturbances. This is an extension of what is known as the "system identification problem": if we watch an unknown system long enough, observing inputs and outputs, how can we reconstruct its internal states? This problem, in turn, is an extension of the learning problem that has been tackled by neural network researchers since the field's beginnings: how do we learn associations in space and time?

The Advanced Systems Research Group of the Avionics Laboratory at Wright Patterson Air Force Base has been particularly concerned with neural network learning for pattern recognition and control. As one example, the application to *damaged systems* is crucial. Can an automated system participating in the piloting of a severely damaged aircraft rapidly learn a model of the craft's dynamics to keep it functional? Neural network architectures truly give us special-purpose systems with autonomy and real-time capabilities. Exciting developments in the field have extended the technologies for spatial learning to temporal learning. However straightforward extensions of existing algorithms (such as backpropagation) to the temporal domain incur a large cost in terms of processing time and hardware resources.

Because of the cost problem, in my research I decided to approach the temporal learning problem in terms of a problem reduction: how can we fool a neural network into thinking it is solving a simple spatial problem when in fact it is handling a complex spatiotemporal one? Ironically, this line of research leads us into quite unorthodox neural network dynamics.

My research interests concern the role of dynamical systems in information processing. My earlier work on the biophysical simulations supporting a continuous neuron model, and its application to for robot navigation and path optimization, laid the foundation for my summer assignment. Aside from its benefits in developing a new technology, my work at the Avionics Laboratory helps create a foundation for a clearer understanding of the role of spatiotemporal dynamics in intelligence.

II. OBJECTIVES OF THE RESEARCH EFFORT

Currently, the problem of learning associations between temporal sequences of spatial patterns is one of the most challenging in the field of neural networks. Inspired by the facility with which real brains accomplish this learning, generalizations of successful connectionist gradient descent algorithms from the spatial to the spatiotemporal domain have attracted much interest in the past two years. Yet the advantages of traditional connectionist solutions have become less clear; these solutions are both time consuming and require large amounts of memory and inter-neuronal connections. This suggests it is fruitful to search for new ideas.

My summer fellowship allowed me to develop a key concept, called *context*, and elaborate it into a novel architecture employing both connectionist and non-connectionist elements. This architecture has successfully and efficiently solved simple sequential learning problems, and work is currently underway on more complex tasks. The idea of context allows us to decompose the problem of learning spatiotemporal sequences into two subproblems, solvable by a convergence of new and old technologies.

The overall problem may be stated in general terms as follows. We are given a set of multi-dimensional input signals $\{x(t)\}$ and we observe the corresponding output signals $\{y(t)\}$. Unlike a simple spatial mapping, the output at time t depends not merely on the input at that time, but on an undetermined history of earlier inputs. We wish to have our learning machine form a model that explains this $x \rightarrow y$ association. Our idea is first to generate a compressed representation of input histories, called context. We do this by sending the input signals to a dynamical system called a *context reverberation subsystem*, which serves as a repository for information relevant to the association. This system may be either a sparsely connected network of conventional linear-threshold neurons, or an unorthodox complex neuron with continuous internal states. Hardware requirements for these context reverberation subsystems are less stringent than for ordinary connectionist learning architectures. The remainder of the problem is to use a conventional learning algorithm to group the current inputs with the context, and associate these with output patterns. As a consequence, the introduction of the context concept has allowed us to fall back on well-studied learning algorithms such as the perceptron convergence method.

In the three technical sections that follow, I map out the ideas and summarize the results of our implementations. The accomplishments may be summarized as follows.

Accomplishments of the Research Effort 7/2/90 - 9/7/90

(1) The context reverberation architecture was shown to work efficiently on the sequential parity problem, which is a temporal generalization of the benchmark spatial parity problem. An interactive simulation environment written in C++ was developed for these experiments.

(2) The connectionist version of this architecture was shown to require only local connectivity between neuronal units. This permits a dramatic, order of magnitude reduction in connections. Our solution overcomes the high connectivity problem, the chief obstacle in hardware implementations of traditional neural networks.

(3) We have formulated necessary conditions for the context dynamics to produce efficient learning. The essential feature of the context reverberation subsystem is its ability to discriminate temporally different but spatially similar inputs. This capability stems from the highly aperiodic nature of the dynamics. We have found a striking correlation between period length and the efficacy of the subsystem. This allowed us to bring in results on the theory of phase transitions in random nets to suggest new context reverberation designs.

(4) We have investigated context reverberation subsystems that are modified by simple genetic algorithms. This allowed us to improve the amenability of the entire system to traditional learning algorithms, and results more rapid learning times.

(5) We have mathematically elucidated the nature of the context reverberation approach to sequential learning. When phrased in the language of finite automata theory, the learning system is modeling the observed system in a sense different from the textbook "simulation-as-automaton-homomorphism" definition. Instead, it is employing a mathematically dual notion we call "quotient simulation". These ideas have implications that transcend the neural domain that motivated them, and deepen the significance of this work.

(6) We have identified promising non-connectionist technologies for context reverberation systems. The discoveries of the sufficiency of local connectivity (item 2, above) and the characterization of the requisite dynamics (item 3) suggest that a dynamically richer neuron, the so-called *reaction-diffusion* neuron, can be employed in our architecture. A single such neuron would play the role of an entire local connectionist net of several dozen units.

The final section makes recommendations on the continuation of the research programme.

III. STRATEGIES FOR LEARNING AN UNKNOWN DYNAMICS

In this and the following two sections, we describe the details of our research programme. Since many of the ideas are new and unfamiliar, space limitations require us to be somewhat terse. Subsequent publications will give these ideas their more complete exposition.

The canonical problem in neural network research is the following: *Given a finite subset of the graph of a function* f, *guess what* f *is*. The subset of the graph we are given is called a training set. Elements of the domain are spatial patterns. The algorithm must process this set to come up with a representation for f. The representation can be used to compute f on elements of the domain not appearing in the training set, and this is called "generalization". Often we are interested in adding a temporal dimension. This could mean real time, or merely the information included in *sequences* of input patterns provided to the system. In other words, an output pattern is no longer determined only by an input pattern, but potentially by an unbounded sequence of previous input patterns. If we let X and Y be our input and output sets, respectively, we pass from the interpolation of a function $f: X \rightarrow Y$ to the simulation of a dynamical system $\delta: Q \times X \rightarrow Q$ where Q is the set of states of our observed system. The job of the learning system is to construct an internal model of this observed system. Training sets are now *sequences* of inputs paired with *sequences* of output. A good simulation of δ permits good generalization. Connectionist systems can be harnessed for such computation by using recurrent networks. Indeed, backpropagation formally generalizes fairly easily to the recurrent case, although such an extension seems to be fairly demanding of computational resources [19]. This is a very active area of connectionist research [3,16]. Of course, this is by no means the only sense of "temporal" learning. One can also phrase the problem as one of learning a sequence of actions in a network with scalar feedback; this is the work of Klopf [13].

Our architecture is sketched schematically in Figure 1. We assume a discrete time scale. Spatial input patterns arrive on input lines, and are sent to an output layer of conventional linear threshold neurons, and to an internal subsystem. This internal subsystem has explicit or implicit recurrent connections, and is used to store the state of the system. Unlike the state units in the work of Jordan [5], the state representation is arbitrary. We call this a *context reverberation (CR) subsystem*. In the work of Gallant [6], enhancing an earlier model of Rosenblatt [17], a totally connected net of linear threshold neurons was used for a similar function. The output of a sequential system depends on current input plus state. Our architecture captures this dependence in a very straightforward way; output units receive signals from the input units and the CR subsystem.
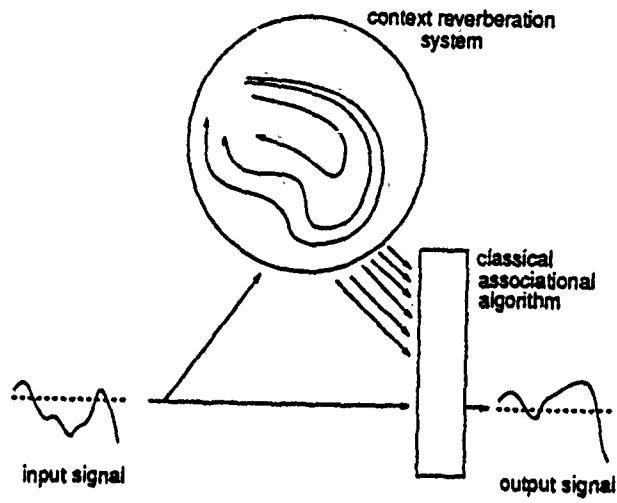
Figure 1. The context reverberation architecture.
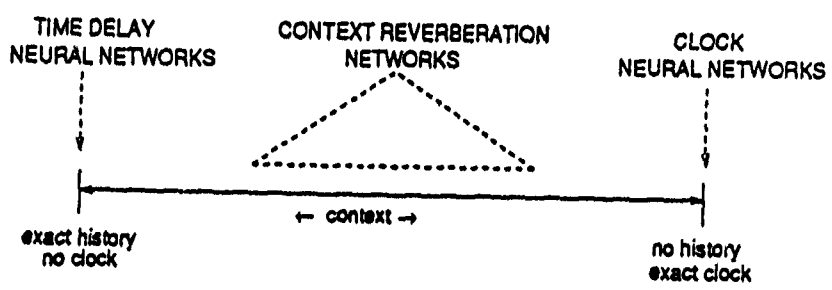
Figure 2. The spectrum of context usage.
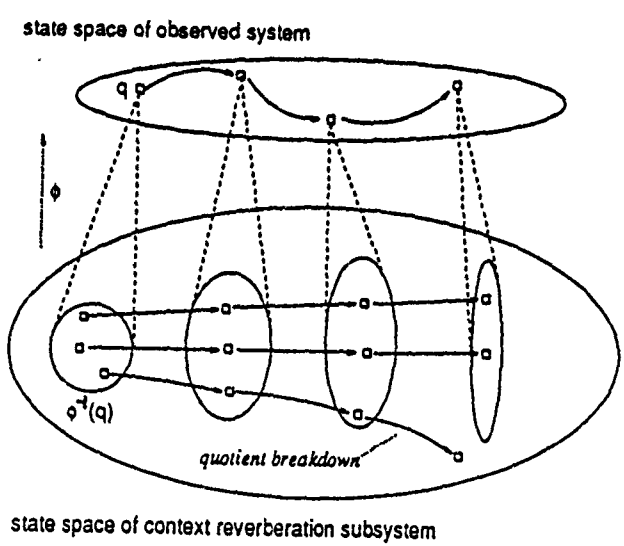
Figure 3. The over-representation of states by context.

The important research problem here is this: how can the CR-subsystem send an effective representation of input history to the output layer? Let us define *context* as the dynamical state of the CR subsystem. This concept can be clarified by considering a spectrum of approaches to connectionist architectures for sequential problems, shown in Figure 2. At one extreme, time-delay neural networks send exact delayed copies of the input signals to the output layer. A conventional learning algorithm learns to produce correct output from inputs $x(t), x(t-1), \cdots x(t-TMAX)$, where the input temporal window $[t-TMAX...t]$ is determined in advance. Hence context in this case exactly corresponds to history. On the other extreme, some sequential problems depend on time and not on history. A clock neural network calculates outputs from the input signal $x(t)$ coupled with an encoding of t.

In many problems, we need an unbounded temporal window, plus some access to encoded time, but do not want to maintain the overhead of a very large number of time-delayed inputs. The usual approach would be to designate some neurons as state units, and give them recurrent connections. These connections can be learned by "recurrent back-prop", for example. But our architecture differs at a deep level from that approach. In such recurrent adaptation algorithms, the idea is to create a homomorphism from the dynamics of the observed system to the dynamics of the neural net. If δ is the dynamics of the system that we want to model (a finite automaton that computes parity, for example), we would to set up the weights of our neural net so that its dynamics is given by δ', where there is a mapping h from external states Q to net states Q' that preserves the dynamics. This means that these two dynamics are coordinated by the relation $\delta'(h(q),x) = h(\delta(q,x))$ for all inputs x and states q of the observed system. This homomorphism of finite automata corresponds to the simulation relation: the learning system is supposed to *simulate* the observed system.

Our model constructs a representation of the observed system in a different, innovative sense. We have a mapping ϕ that collapses many states of our CR-subsystem onto each state q of the observed system. This is depicted in Figure 3. The equivalence classes of states, $\phi^{-1}(q)$, correspond to the ellipses in the bottom half of the figure. The set of these classes is called the "quotient space" under the mapping ϕ. It is this quotient space that, as learning proceeds, should come to represent the known system. As time proceeds, each class flows through the CR state space (bottom) tracking the transitions in the observed state space (above). In reality, since the complexity of the CR dynamics is so great, occasionally we will find that generalization off a training set of sequences is poor over long time periods. Referring to the figure, this happens when the dynamics of the CR system does not track the quotient structure. The trajectories

wander out of their proper equivalence classes, so the learned internal model would break down a few time steps past the end of the training data. A good CR architecture will make this a rare occurrence.

IV. LOW-CONNECTIVITY ARCHITECTURES FOR CONTEXT REVERBERATION

Having discussed the role of the CR subsystem, we now turn to its implementation. We first investigated a connectionist architecture, in which the CR subsystem contained a number of linear threshold units arranged in a grid, each connected to others within a limited neighborhood. (This contrasts with the total (i.e., $O(n^2)$) connectivity in the nets of Hopfield, Anderson, and others.) Synaptic weights are fixed and randomized. We had known from the work of Gallant and King [6] that a totally connected layer of random hidden units could learn sequential problems with some success. But dozens of units in a such a highly connected system will result in hundreds of wires. Our first step, undertaken with a graduate student at Wright State University, showed that for the so-called "robot plan task", low connectivity was more efficient [12]. During the summer, an "autopsy" of the networks (from large sequences of randomly generated trials) showed that the most successful ones were those that had highly irregular dynamical trajectories. We can see this if we plot the intensity of the firing states of the CR net versus time, for an arbitrary clamped input. (Intensity here means we add up the +1/-1 (firing/silent) values.) This is shown in the two plots at the bottom of Figure 4. The first plot shows a "good" CR net, which learned the task quickly, and the bottom a "bad" net that failed to learn. This shows that aperiodicity (or, strictly speaking, periods of length much longer than the time scales of interest) is an attribute correlated with good CR performance. One way to control this is to adjust the "gain" signal, which amplifies the input signals coming in to the hidden net. This is shown in the plot at the top of Figure 4. Low gain yields better results.

The robot plan task, however, merely involves memorizing a set of sequence pairs, and does not address the issue of generalization. To go beyond this, we had our network learn the sequential parity automaton (Figure 6.) The output at time t is the parity of the string of binary inputs from time 0 to time t. We used connectionist CR nets with various neighborhood sizes, and plotted the learning times in Figure 5. The perceptron algorithm was used to make the input/context associations. The y axis plots the number of perceptron epochs (passes through the training set), when trying to learn parity from five sequences of duration equal to 15 time units. We used 64
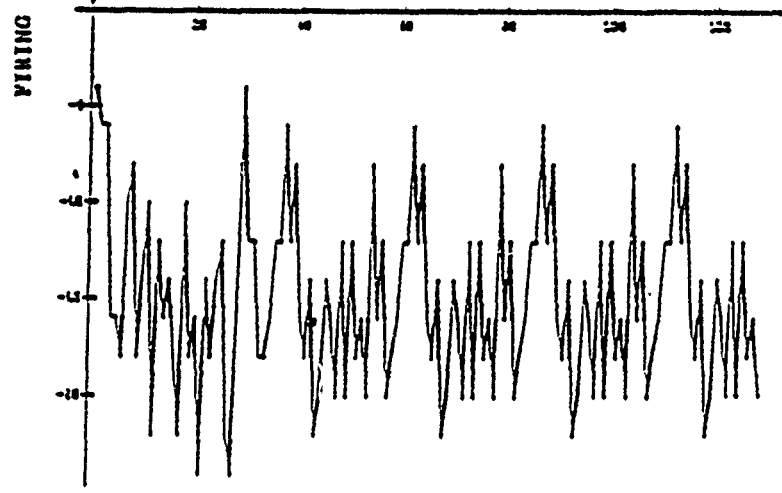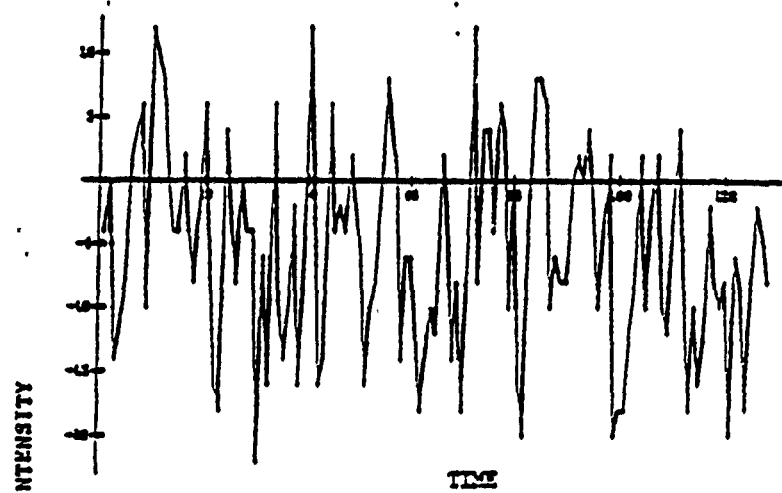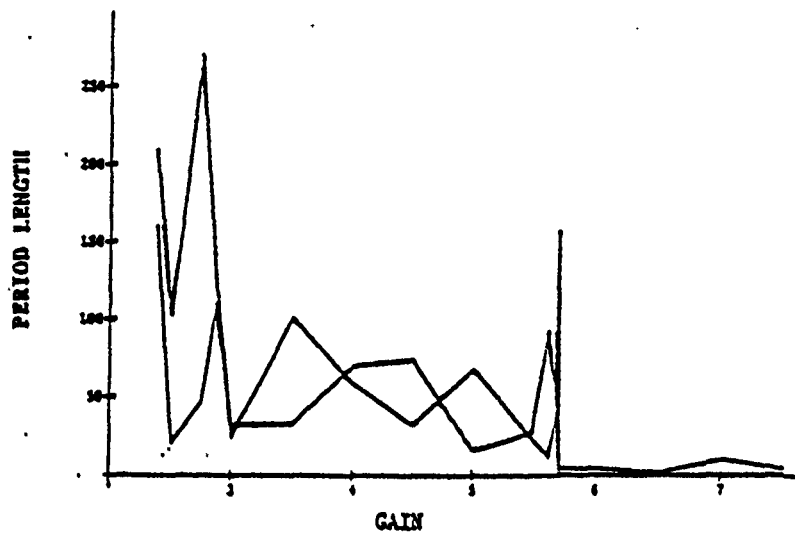
Figure 4. Periodicity and dependence of period on gain
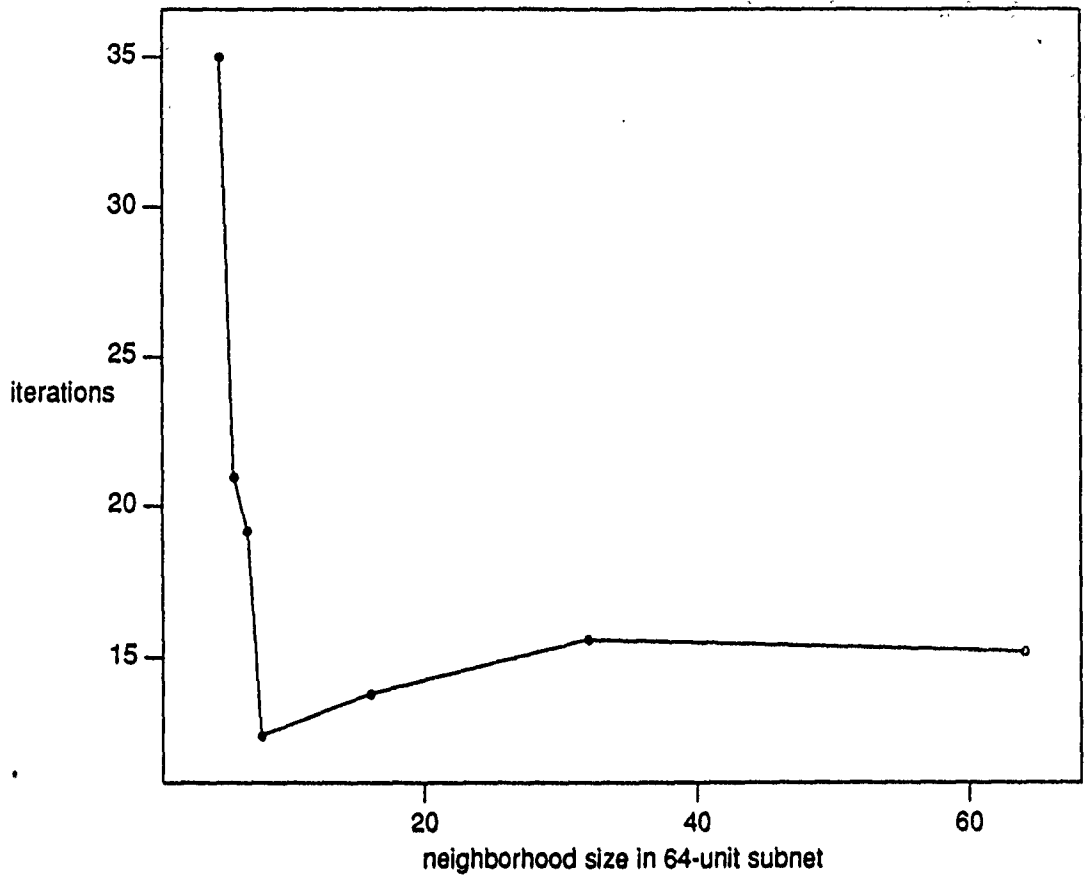
Figure 5. Learning times versus connectivity for the sequential parity problem.
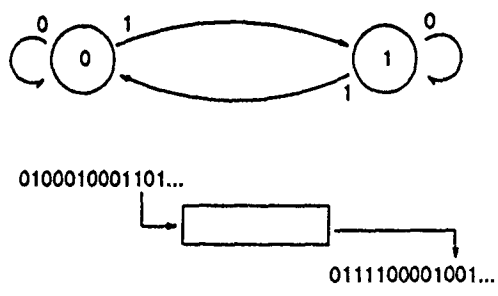
Figure 6. The state transition graph for the sequential parity problem.

CR units, and averaged over 20 different configurations. We only plot the result for cases when every instance learning the training data perfectly (zero error). Generalization on a test set of five random length 15 sequences resulted in an error under 6% for the 7-neighbor case. Preliminary results from a genetic modification scheme, which remove unchanging context units and swap weight values, showed modest improvements in learning rates with no significant change in generalization ability.

Figure 5 shows that not only is lower connectivity dramatically more efficient in terms of required connections, but it is even more efficient in absolute learning time. The curve's minimum occurs at a neighborhood size of 7, corresponding to 7×64 = 448 weights. This can be compared 64^2 = 4096 weights for the Gallant implementation, which took slightly longer to learn on average. As the neighborhood size shrinks to 6, 5 and 4, learning times increase, but for some applications weights may be traded off against learning time.

V. NON-CONNECTIONIST ENHANCEMENTS TO THE ARCHITECTURE

We can conclude from the experiments just described that the *locality* of the connections in the CR architecture is a feature to be exploited. This is an economic issue; fewer connections require less space and ease hardware implementations. But it also allows us to more rigorously investigate the dynamical properties of CR systems. In this section we discuss our work in relation to work in the dynamical systems disciplines. Insights from these disciplines are important, because they help us understand and extend the capabilities of the CR architecture.

Our connectionist CR-subsystem uses a random network. Kauffman [7] showed that networks of totally connected random boolean units exhibit an exponential growth in limit cycle length as the number of units increases. (A limit cycle is one period of the state trajectory.) In other words, a the state trajectories of even a small net are highly aperiodic. This is termed a "chaotic" phase, as opposed to the so-called "ordered phase" when cycle lengths increase polynomially with the number of units. Reducing from global to local connectivity slows this growth. With only 2-neighbor connectivity the periodicity grows as $\sqrt{N}$, too slowly for effective use as a context reverberation net. Kürten [8] studies threshold units with different local connectivities and shows that whereas low-connectivity systems in which neighbors are chosen randomly have exponential growth in cycle lengths, nearest-neighbor systems may show linear growth. The low-connectivity nets studied were the 3-neighbor "honeycomb" lattice, and the self+4-neighbor

square lattice. The zero-threshold honeycomb lattice shows linear growth in cycle length, and the zero-threshold square lattice with self-feedback shows exponential growth. Adding a unit threshold moves this system back into an ordered phase. Our research suggests that we should seek a chaotic phase in our context-reverberation systems. So if we are committed to using a local net of conventional neurons, the connectivity level (i.e., neighborhood size) should be at least 4. For connectivity less than this, learning should be impossible. We have experimentally confirmed this: the curve in Figure 5 goes to infinity on the left, when connectivity approaches 3.

Given the efficiency of low-connectivity threshold lattice automata as chaotic reverberation subsystems, we can investigate the potential of continuous local dynamical systems for setting up context. The general case is a reaction-diffusion equation, of the form

$$\frac{\partial u(\mu, t)}{\partial t} = \nabla \cdot D(\mu) \nabla u(\mu, t) + R[u(\mu, t)] \tag{1}$$

Here $u(\mu, t)$ is an excitation signal diffusing across a space with coordinates μ. D is the diffusion coefficient, which may vary across the space. R is the reaction term, a function of the excitation level. In the two terms we have the two ingredients necessary for effective context reverberation: local communication (via the diffusion term), and local computation (via the reaction term). In one dimension, in analogy to the discrete set of threshold units we used for learning the parity problem in the previous section), we can compartmentalize the system to create a ring of compartments. The diffusion term with discretized compartments becomes $\sum\limits_{j=k\pm 1} d_{jk}(u_j - u_k)$. This system was introduced by Alan Turing [18] to study the destabilizing effect of diffusion in morphogenesis. (Turing used two diffusing signals.) Othmer and Scriven [15] examined how the dynamical properties of this reaction-diffusion system depended on the topology, studying rings and lattices in what was a continuous analog to the studies of Kürten [8] on lattice automata cited above.

Can we expect such continuous dynamics to improve our CR-subsystems? We claim that it should be possible by using a neuronal model based on the Turing morphogenesis equations. This model is called the *reaction-diffusion neuron* [9,10], a continuous extension of a discrete linear unit used in a Darwinian brain model. These neurons take input signals and map them into continuous gradients, which evolve by reaction-diffusion equations. Gradients are read by spatially fixed sensors, whose response induces the firing of the neuron. With suitable dynamics one reaction-diffusion neuron can play the role of an entire CR-subnetwork. It is interesting to note that one of the features that make such neurons amenable to learning algorithms is their

gradualism property; slight changes in the internal sensor distribution produce slight changes in the dynamics. Indeed, this is like the case of totally connected random boolean nets, which have low structural *and* Lyapunov stability [7]. On the other hand, we may trade off gradualism for long cycle lengths, which, as we have seen in the previous section, are strongly correlated with good learning performance in CR systems. This suggests that a highly aperiodic reaction-diffusion neuron would be advantageous. We have the extreme case of this in the chaotic Lorenz-Turing neuron [11], in which trajectories proceed along a strange attractor. Figure 7 shows a plot a a two-compartment LT-neuron, showing the diffusion-caused interference of the familiar Lorenz butterfly attractor.

Once we allow continuous dynamics to enhance the CR system, we can consider adaptation of this dynamics. Recall that in the experiments discussed so far, the CR-subsystem was random and fixed, except possibly for sporadic localized genetic modification. This allows the context dynamics to evolve concurrently with the representation. We can view this idea of context as analogous to some phenomena observed in other research. In the continuous case we can have formation of topological features such as those that arise in Turing-type morphogenesis systems, e.g., the stripes of Meinhardt and Gierer [14]. A well-known neural analog is in the work of Amari [1] on pattern formation in neural fields. A neural field changes the representation of neurons from a set of finite units to a manifold of mathematical points. Firing signals do not propagate along connections, but spread out along the manifold, governed by equations of the form:

$$\tau \frac{\partial u(\mu,t)}{\partial t} = -u(\mu,t) + \int w(\mu,\mu') f[u(\mu',t)] \, d\mu' + \text{input terms.} \tag{2}$$

The weighting "matrix" $w(\mu,\mu')$ internal to the neural field is fixed in advance; only the input weights change (according to a Hebbian algorithm). Efferents from the neural field in a sense use the acquired patterns as state information, since the topographic arrangement on the field is also a kind of repository for input history. Let us call the kind of input context in these morphogenesis and neural field models *topographic context*. This contrasts with the concept of *scrambled context* used by our CR systems to represent history.

Scrambled and topographic context promise to be important notions in the theory of sequential learning. We have shown that the requirements for a good context-reverberation subsystem do not include the high connectivity required for connectionist solutions to other problems. This may encourage molecular electronic hardware implementations [4]. Low-connectivity nets of linear threshold functions with nearest-neighbor topology provide long cycle lengths and are an effective means for providing scrambled context information to a single-layer learning

LT-neuron ($n = 2$)
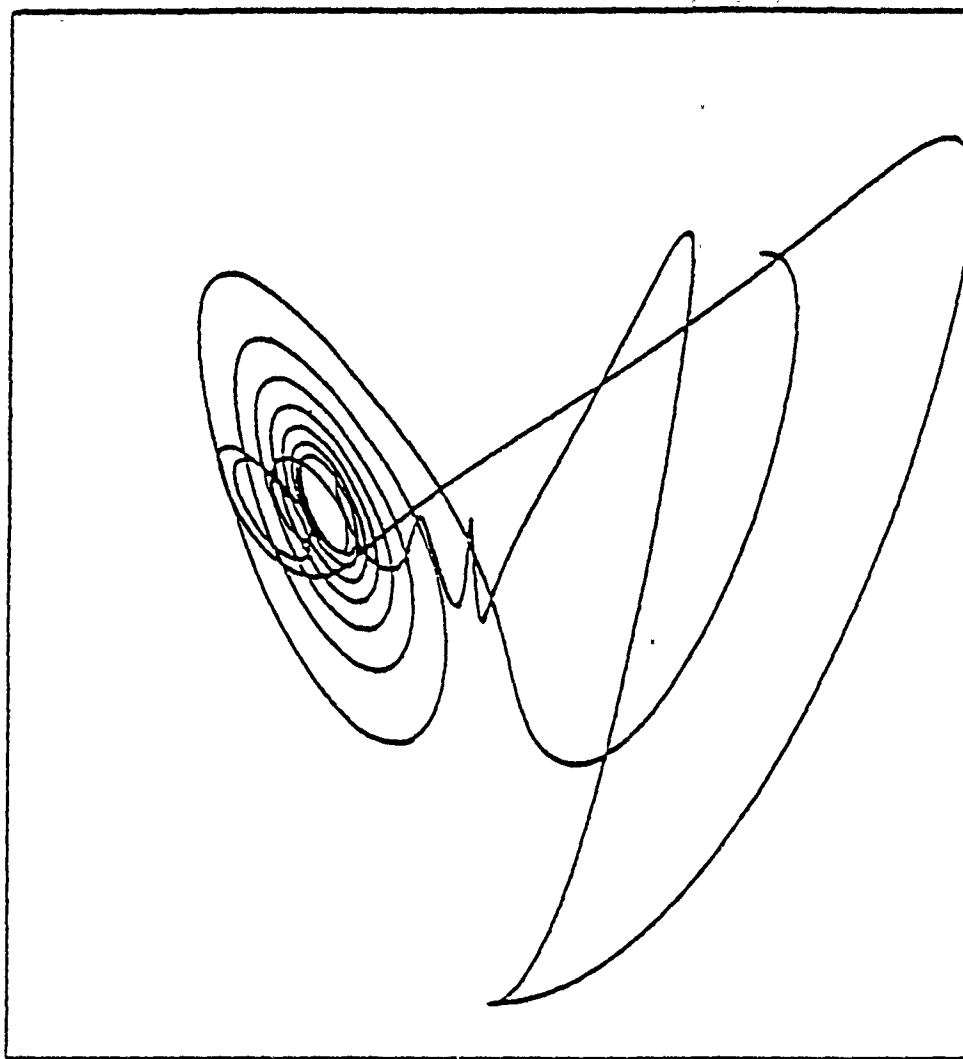
$\rho = 28.00$ $\beta = 2.667$ $d = 10.00$ (unit)

seed= 801

Figure 7. Trajectory of a 2-compartment Lorenz-Turing neuron

algorithm for learning the parity dynamics. We believe that such a result is encouraging for the study of non-symbolic non-connectionist continuous systems for solving "real" artificial intelligence problems. Far from merely providing a new technology for machine learning, the CR system has produced fertile ideas enabling a more profound understanding of the learning problem itself.

VI. RECOMMENDATIONS

In conclusion, we have constructed a new neural architecture, called the *context reverberation architecture*, that (1) learns the sequential dynamics of an observed system; (2) learns rapidly and with accurate generalization; (3) does so with dramatically low connectivity requirements; (4) uses aperiodic dynamics to represent context; (5) can be sped up by simple genetic algorithms; (6) suggests a new paradigm for the sequential learning problem; and (7) may be enhanced with a continuous time non-connectionist neural model.

The continuation of this project requires a working simulation of a CR system solving a highly complex sequential problem. It also demands a full theoretical characterization of the computational power of the architecture. Toward this end, the following recommendations are made.

(1) The learning system should attempt to learn a finite state system with a dynamics two orders of magnitude more complicated than the parity problem. This means moving from two states in the observed system, to over a hundred states. With conventional architectures, this would be intractable. Experimentation with this more complicated system will require an enhancement of the simulation environment. Figure 8 shows a screen from the current environment.

(2) The connectionist version of the context reverberation subsystem should be coupled with a version of the Doya-Yoshizawa algorithm [2] to get the most computational work out of the context units. This algorithm is a continuous time extension of recurrent backpropagation. A detailed study of the tradeoff between this algorithm's cost and the efficiency gain is in order. This would smooth the transition to handling continuous input and output signals.

(3) The coupling of the Lorenz-Turing neuron to the learning system should be perfected. Mathematical analysis done during the summer suggests the Doya-Yoshizawa algorithm can be modified to act between compartments in a reaction-diffusion system like the Lorenz-Turing neuron. This ties in to fundamental research on the computational potential of chaos.

An SFRP follow-on grant from the AFOSR Research Initiation Program will be sought to pursue these recommendations. We expect contributions in both technological and theoretical areas.
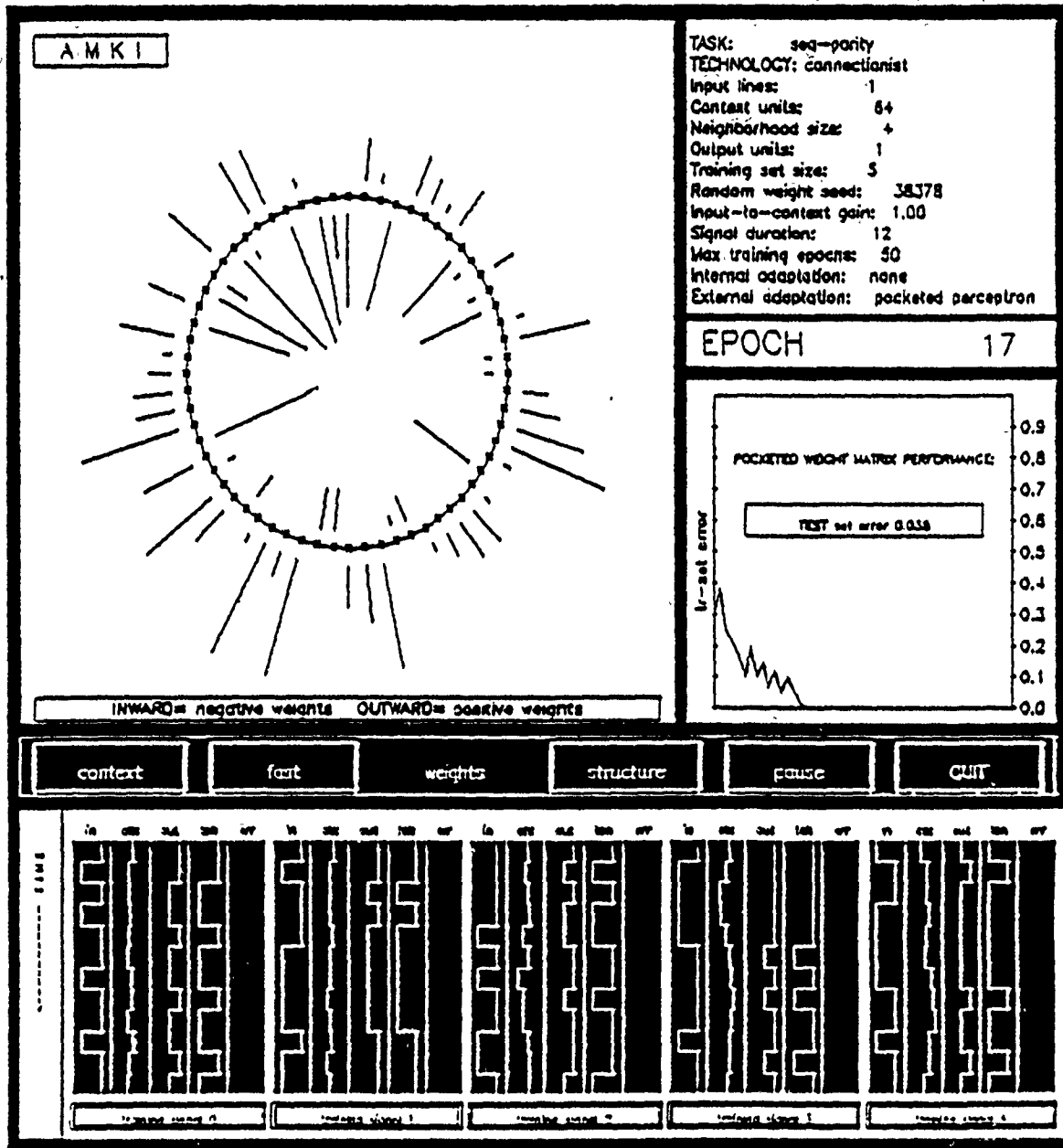
Figure 8. A screen from the AMKI interactive simulation program for CR systems

REFERENCES

1. Amari, S., "Dynamics of Pattern Formation in Lateral-Inhibition Type Neural Fields," Biological Cybernetics. 1977, Vol. 27, pp. 77-87.

2. Doya, K. and S. Yoshizawa, "Memorizing Oscillatory Patterns in The Analog Neuron Network," Proc. IEEE/INNS Conference on Neural Networks, January 1990, pp. 127-132.

3. Giles, C.L., G.Z. Sun, H.H. Chen, Y.C. Lee, and D. Chen, "Higher-Order Recurrent Networks and Grammatical Inference," Neural Information Processing Systems 2, D. Touretzky, Ed., San Mateo, California, Morgan-Kaufmann, 1990, pp. 380-387.

4. Hong, F.T. "The Bacteriorhodopsin Model Membrane System as a Prototype Molecular Computing Element," Biosystems. 1986, Vol. 19, pp. 223-236.

5. Jordan, M.I., "Serial Order: A Parallel, Distributed Processing Approach," Institute for Cognitive Science Report 8604, University of California, San Diego, May 1986.

6. Gallant, S.I., and D.J. King, "Experiments with Sequential Associative Memories," Cognitive Science Society Conference, Montreal, August 1988.

7. Kauffman, S.A., "Principles of Adaptation in Complex Systems," In Lectures in the Sciences of Complexity, D. Stein, Ed., Addison-Wesley, 1989.

8. Kürten, K.E., "Dynamical Properties of Threshold Automata with Nearest-Neighbor Interactions on a Regular Lattice," Proc. IEEE International Conference on Neural Networks, August 1988, pp. 137-143.

9. Kirby, K.G., and M. Conrad, "The Enzymatic Neuron as a Reaction-Diffusion Network of Cyclic Nucleotides," Bulletin of Mathematical Biology. 1984, Vol. 46, pp. 765-783.

10. Kirby, K.G., M. Conrad, and R. Kampfner, "Evolutionary Learning in Reaction-Diffusion Neurons," Applied Mathematics and Computation, To appear, 1991.

11. Kirby, K.G., "Information Processing in the Lorenz-Turing Neuron," Proc. IEEE Engineering in Medicine and Biology Conference, Molecular Electronics Track, November 1989, pp. 1358-1359.

12. Kirby, K.G., and N. Day. "The Neurodynamics of Context-Reverberation Learning," Proc. IEEE Engineering in Medicine and Biology Conference, Molecular Electronics Track, November 1990, to appear.

13. Klopf, A.H., "A Neuronal Model of Classical Conditioning," Psychobiology. 1988, Vol. 16, pp. 85-125.

14. Meinhardt, H., and A. Gierer, "Generation and Regeneration of Sequences of Structures During Morphogenesis," **Journal of Theoretical Biology.** 1980, Vol. 85, pp. 429-450.

15. Othmer, H.G., and L.E. Scriven, "Instability and Dynamic Pattern in Cellular Networks," **Journal of Theoretical Biology.** 1971, Vol. 32, pp. 507-537.

16. Pineda, F.J., "Recurrent Backpropagation and the Dynamical Approach to Adaptive Neural Computation," **Neural Computation.** 1989, Vol. 1, pp. 161-172.

17. Rosenblatt, F., **Principles of Neurodynamics,** Washington, DC, Spartan Press, 1961.

18. Turing, A.M., "The Chemical Basis of Morphogenesis," **Phil. Trans. Royal Soc. B** 1965, Vol. 237, pp. 37-72.

19. Williams, R.J., and D. Zipser, "A Learning Algorithm for Fully Running Recurrent Neural Networks," **Neural Computation.** 1989, Vol. 1, pp. 270-280.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc

FINAL REPORT

Fiber Laser Preamplifier for Laser Radar Detectors

Prepared by:	Richard E. Miers
Academic Rank:	Associate Professor
Department and	Physics
University:	Indiana University/Purdue University at Fort Wayne
Research Location:	WRDC/AARI-2 Wright Patterson AFB Dayton, OH 45433
USAF Researcher:	Paul F. McManamon
Date:	10 Aug 90
Contract No:	F49620-88-C-0053

Fiber Laser Preamplifier for Laser Radar Detectors

by

Richard E. Miers

ABSTRACT

A study was made of the feasability of using a fiber laser preamplifier as a means of improving the detectability of laser radar signals. Although fiber laser amplifiers at the wavelength of interest, 1.064 μm, have not been developed, a study of the development of Er-doped fiber laser amplifiers for 1.55 μm indicates the usefulness of such amplifiers. Also the properties of Nd-doped fibers indicates that such fibers should amplify 1.06 μm wavelength signals as well as or better than the Er-doped amplifiers. Recommendations for development and testing of such an amplifier are given.

Acknowledgements

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. The help in all administrative and directional aspects of this program by Universal Energy Systems is acknowledged.

The friendliness and helpfulness of the personnel of AARI-2 are greatly appreciated. Paul F. McManamon provided enthusiastic support for my efforts. His advise and help were invaluable. I would like to thank Captain Larry Myers for providing the initial idea for this project and for his assistance in gathering valuable information. Carol Heben and Lt. Deanna Won provided information and help that was quite useful for development of this project.

I. INTRODUCTION

The Electro-optics Division of the Avionics Laboratory at Wright- Patterson
Air Force Base is involved in the development of laser radar systems. One
possible method of increasing the detectability of a returning laser radar
signal might be to use a fiber optical laser preamplifier immediately before
the photo detector.

In the competition between signal and noise, immediate amplification of the
signal eliminates detector noise sources as areas of concern (due to the higher
signal level after amplification). Fiber laser amplifiers may allow immediate
signal preamplification.

My research interests in recent years have involved the use of various lasers
in atomic physics research. As part of this research I have constructed lasers
and have used a number of different types of commercial lasers. My
knowledge and experience involving lasers contributed to my assignment to
the Electrooptics Division of the Avionics Laboratory.

II. OBJECTIVES OF THE RESEARCH EFFORT:

Light amplification by stimulated emission was first accomplished with the advent of Maiman's Ruby Laser in 1960. [1] Light amplification of a traveling light wave in an active fiber was reported by Koester and Snitzer in 1964. [2] They measured the amplification of a 1.06 μm wavelength signal in a one meter long Nd-doped fiber that was pumped with a flash lamp situated at the center of the spirally wound fiber. In 1969 Holst and Snitzer detected 1.06 μm signals with a fiber laser preamplifier. [3] They suggested use in a laser radar system.

Subsequent to the initial efforts noted above, very little was reported on traveling wave fiber laser amplifiers until around 1985. However, doped fibers have been used in fiber lasers. In 1973 Stone and Burrus observed lasing in two types of Nd-doped silica fibers. [4] One fiber consisted of an active core of fused silica, aluminum oxide, and Nd-oxide surrounded by a fused-silica cladding. The second was made up of an active core of fused silica and Nd oxide only. In both cases the Nd content was less than 1% by weight. They were end-pumped with a chopped argon-ion laser beam and with a pulsed dye laser operating at 590 nm. The first fiber oscillated at 1.06 μm and the second at 1.08 μm. In a later paper they reported cw operation of Nd-doped fiber lasers at room temperature. [5] In 1979 they reported a self-contained LED-pumped single-crystal Nd:YAG fiber laser. [6] Since 1985, due to the interest of the communications industry, a great deal of development and testing has been done on Er-doped fibers to amplify light signals at wavelengths at or near 1.55 μm. However, as of now there seems

to have been no effort to develop traveling wave fiber laser amplifiers at 1.06 μm which is the wavelength of immediate interest for laser radar at the Avionics Laboratory.

My assignment as a participant in the 1990 Summer Faculty Research Program (SFRP) was to determine the potential for increased laser radar receiver sensitivity based upon immediate optical signal preamplification through use of a Nd-doped fiber laser amplifier. Results of this investigation show that this type of laser preamplifier may be useful. Therefore the development of an experimental program to build and study such an amplifier is a secondary goal. It will be proposed that this investigation be continued with funding from the Research Initiation Program.

III.

Fiber laser amplifiers seem to have a number of advantages over other optical amplifiers. Fibers have been developed to transmit a single mode signal. These fibers are polarization independant. Laser diodes with wavelength 807 nm, which is within a very efficient absorption pump band for Nd^{3+}, are readily available for end pumping . Theoretically, it appears possible to develop a Nd-doped fiber amplifier with ideal i.e. quantum limited signal-to noise ratio.

Kingston derives the signal-to-noise ratio for a laser amplifier operating in single mode applied to both coherent and direct detection systems. [7] His treatment is similar to that of Yariv. [8] They show that the signal-to-noise ratio for an ideal 4-level laser amplifier can be given by

$$(S/N)_P = P_s/(h\nu\Delta\nu) \tag{1}$$

where $(S/N)_P$ is the signal-to-noise ratio for a single mode of amplified spontaneous emission, P_s is the power of the detected signal, $h\nu$ is the photon energy at the center of the stimulated emission band, $\Delta\nu$ is the detector bandwidth, h is Planck's constant. This is the ideal quantum limited receiver equivalent to a heteradyne detector with quantum efficiency of 1.

Nd-doped fibers are four-level laser systems while Er-doped fibers are three-level systems. In a four-level laser amplifier, population inversion should exist at very low absorbed pump powers since the lower lasing level is initially unpopulated. In the three-level amplifier the lower lasing level is at or near the ground state, therefore the pumping rate must be high enough to move at least half of the electrons from the lower level to the upper level before amplification can occur.

Many of the techniques and experiments leading to the development of Er-doped fiber laser amplifiers may be used in the development of Nd- doped amplifiers at 1.06 μm. Also a future goal at the Avionics Laboratory may be to develop laser radar systems using the eyesafe wavelength of 1.55 μm. In that event an Er-doped laser preamplifier may be used.

Yamada et al. have obtained a signal gain as high as 37.8 dB in an Er-doped fiber amplifier pumped by 0.98 μm laser diodes. [9] They used a 30 m long

Er^{3+} doped fiber at 20 mW launched pump power. The pumping source was an InGaAs strained quantum well LD that could deliver up to about 85 mW output. They found by comparison that 0.98 μm pumping was more efficient than 1.48 μm pumping. Giles et al. measured the noise performance of an Er-doped fiber amplifier pumped at 1.49 μm. [10] They achieved a gain of 37.4 dB at a wavelength of 1.53 μm and a noise figure of 5.3 dB. The sensitivity of the optical receiver without preamplifier was −21 dBm. The receiver sensitivity with preamplifier was −41 dBm. Olshansky defines the noise figure F as 2γ, where γ is the population inversion factor, given by

$$\gamma = g_s/(g_s - a_s) \tag{1}$$

where g_s is the rate of stimulated emission and a_s is the rate of stimulated absorption. [11] For complete population inversion, $\gamma = 1$, and the noise figure F is 3 dB.

End pumping has been accomplished using various pumping lasers and wavelengths. In particular pumping with diode lasers has been achieved. To date laser diode pumping has been reported at 0.807 μm, 1.48 μm, and at 0.98 μm with the pumping at 0.98 μm reported to be the most efficient. Yamada, et al. tested the noise characteristics of Er^{3+} doped fiber amplifiers pumped by 0.98 μm and 1.48 μm laser diodes. [12] The spontaneous-spontaneous beat noise and the spontaneous shot noise for the 1.48 μm pumping were higher than those for the 0.98 μm pumping. The noise figures estimated from the signal-spontaneous beat noise were 3.2 dB for 0.98 μm pumping and 4.1 dB for 1.48 μm pumping. Bour, et al. describe a 980 nm diode laser suitable for pumping Er^{3+} fiber amplifiers. [13] These lasers are capable of supplying 125 mW of CW power. Takada et al. measured picosecond laser diode pulse

amplification up to 12 W using laser diode pumping at 1.48 μm. [14] In another test of an optical receiver with an Er-doped fiber preamplifier, an 8.3 dB improvement of the signal-to-noise over the receiver without the preamplifier was recorded. [15] The Er-doped fiber had a 7.2 μm core, 17 m length, and 300 ppm Er^{3+} concentration in pure silica fiber. The signal was detected with an APD/FET receiver.

An important issue related to optical amplifiers is the phenomenon of amplification of spontaneous emission (ASE). This occurs concurrently with signal amplification and degrades the signal-to-noise ratio (SNR). Some of the features that affect this are dopant concentration distribution across the fiber, the pump and signal mode spatial overlap, the signal absorption due to the lower level population, the amplification of spontaneous emission, the detailed structure of the fluorescence and the absorption spectra, and the gain homogeneous and inhomogeneous broadening. Excited state absorption which is known to occur at some pump wavelengths may also have an effect. These features have been investigated both theoretically and experimentally for Er-doped single-mode fibers. [16] For long fibers it has been shown that end-pumping codirectionally with the signal provides lower ASE at the detector than does pumping contradirectionally. [11][17]

Optical amplification characteristics as a function of length and doping concentrations are important. These effects have been investigated for Er-doped silica single-mode fibers. [18] It was found that the increase of the Er concentration causes deterioration of the amplification characteristics in the Er-doped silica glass system, even when the Er concentration was less than

1000 ppm. This relationship is very important in the optimum design of the Er-doped fiber amplifiers.

A number of detectors are available at 1.06 μm. Silicon detectors have the advantage of low noise at room temperature, however 1.06 μm is at the edge of their useful detection range which lowers the quantum efficiency. Germanium PIN avalanche photodiodes are probably the most sensitive detectors at this wavelength, however the noise at room temperature may be considerably higher. InGaAs detectors lie somewhere between these two in both sensitivity and room temperature noise. The quantum efficiency of silicon detectors at 1.06 μm is quite low, around 10% or less. The quantum efficiency of Ge or InGaAs detectors can be quite high, however the dark current at room temperature is much higher. Since a four-level laser amplifier such as Nd^{3+} doped fiber can be theoretically equal to the ideal quantum limited detector with quantum efficiency of 1, use of such an amplifier could make a direct detector as sensitive as the ideal heteradyne detector if the filtered bandwidth can be kept as low as that for the heteradyne detector.

Many of the papers of Desurvire and others examine the theory of Er- doped fiber amplifiers much of which seems to be verified by experiment. [19][20]

IV. RECOMMENDATIONS

Noise generated by a laser fiber amplifier is generally due to quantum or shot noise generated by amplified stimulated emission. Researchers appear to have come close to bringing it near the quantum limit for Er-doped fiber laser

amplifiers. It seems reasonable to assume that similar techniques applied to the development of Nd-doped fiber amplifiers would produce efficient noise-free gain at the 1.06 μm wavelength.

Digonnet and Gaeta have theoretically analyzed an optically pumped fiber laser amplifier. [21] They show that for an ideal 4-level amplifier the unsaturated single pass gain factor can be given as

$$\gamma = \sigma\tau_f P_{abs}/(h\nu_p A_p^*) \qquad (2)$$

where the gain is given by

$$G = I_{out}/I_{in} = e^\gamma \qquad (3)$$

I_{out}/I_{in} is the ratio of amplified signal to the input signal. σ is the stimulated emission cross section for the amplified wave. τ_f is the flourescent lifetime of the upper lasing level, $h\nu_p$ is the photon energy of the pump light, P_{abs}/A^* is the effective intensity of the absorbed pump light in the fiber.

Applying this theory to a Nd-doped silica fiber assuming $\sigma = 3 \times 10^{-20}$ cm^2, $\tau_f = 4.5 \times 10^{-4}$ s, $h\nu_p(800 \text{ nm}) = 2.48 \times 10^{-19}$ J, $A_p^* \sim A_f(r_f = 3 \times 10^{-4}$ cm$) = 3 \times 10^{-7}$ cm^2 gives a slope efficiency of

$$\gamma/P_{abs} = 0.18/\text{mW or } G(dB)/P_{abs} = 0.78 \text{ dB/mW}. \qquad (4)$$

A_f is the cross-sectional area of the pumped fiber of radius r.

Po, et al. reported a slope efficiency of 0.47 dB/mW in a Nd-doped silica fiber when pumped at 800 nm. [22]

Since doped fibers will probably have to be purchased from some manufacturer, the absorption and emission spectra may already have been investigated and published as technical specifications. If special fibers have to be manufactured for test purposes, it may be necessary to make measurements of these properties to evaluate the wavelengths at which they should be pumped and the probable gain profile of the particular fibers. One company, York VSOP, (Chandler's Ford, Hants, UK) is manufacturing single-mode fibers in which 300 ppm of Nd ions have been incorporated. The Gain bandwidth centers at 1088 nm. These fibers probably would not be useful as amplifiers at 1060 nm. Fibers have been made that have bandwidth centered at 1060 nm. [4][22][23][24][25] These are co-doped with Al_2O_3. Researchers at Rutgers University and at Brown University have the capability of manufacturing custom doped fibers that have amplification centered at 1060 nm and have indicated the willingness to provide fibers for test purposes. A number of manufacturers are making erbium doped fibers and may have the facilities to custom dope other fibers. These options are being investigated at this time.

The gain of a fiber amplifier will be a function of pump wavelength and power, fiber length, doping, efficiency of pump coupling into the fiber, and spacial overlap of pump and signal modes. In a four-level laser medium the population inversion and therefore the gain factor is proportional to the pump power absorbed. On the other hand, for a three-level system such as Er-doped glass, the gain factor is only positive when the pump intensity exceeds the threshold intensity for which the upper level population exceeds that of the lower or ground level. Therefore for any given doping concentration and

power level there is an optimum length for maximum gain in an Er-doped fiber. The same is not true for the four-level Nd-doped system. In such a system the fiber can be arbitrarily long in order to absorb as much pump power as possible. Nominal unpumped lengths should have an insignificant effect on signal absorption. These are properties that will need to be studied in the development of a Nd-doped fiber amplifier.

To ascertain the usefulness of a fiber amplifier, the most important aspect will be to study the sensitivity of the amplified detector compared with the sensitivity of an unamplified detector Parameters that have to be studied are: gain, signal, reflections, methods of launching pump light into the fiber, methods of coupling the input of the signal to be amplified into the fiber, methods of coupling the amplified output into the detector, and methods of reducing the detected bandwidth. All of these factors could affect the signal-to-noise ratio of the amplifier. These are all aspects that investigators of Er-doped fiber amplifiers have studied.

To build a Nd-doped fiber laser amplifier a single mode fiber doped to amplify signals at 1.06 μm wavelength will be needed. A suitable diode laser at 800 nm for pumping the fiber is needed. Also needed are either a pigtail coupler or a dichroic mirror with reflecting and transmitting properties such that both the pump light and signal light can be coupled simultaneously into the fiber. Positioners for precise coupling and coupling eyepieces would be needed. A photodetector for detecting 1.06 μm light will be needed along with a narrow band filter centered at 1.06 μm to reduce the spontaneous emission noise and to eliminate any unabsorbed pump light.

Equipment needed for studying the properties of the amplifier are a power meter for measuring input and output signal and pump light, A signal laser at 1.06 μm, and attenuators for varying the signal.

One method of ascertaining the center of signal gain band of the fiber is to place appropriate mirrors at either end of the fiber and measure the wavelength at which it lases or oscillates. Also this may provide an appropriate signal laser for testing the gain of the fiber laser amplifier.

Other miscellaneous test equipment will include an appropriate optical table or breadboard for mounting, a monochrometer for measuring wavelengths of signal and pump light, and mounting equipment for mirrors, lenses etc.

REFERENCES

[1]   T. H. Maiman, Nature 187, 493 (1960); Brit. Commun. and Electr. 7, 674 (1960)

[2]   C. J. Koester and E. Snitzer, "Amplification in a Fiber Laser," Applied Optics, vol. 3, pp 1182-1186, 1964.

[3]   C. G. Holst and E. Snitzer, "Detection with a Fiber Laser Preamplifier at 1.06 μ," IEEE J. of Quantum Elect., vol. QE-5, pp 319- 320, 1969.

[4]   J. Stone and C. A. Burrus, "Neodymium-doped silica lasers in end-pumped fiber geometry," Appl. Phys. Lett., vol. 23, pp 388-389, 1973.

[5]   J. Stone and C. A. Burrus, "Neodymium-Doped Fiber Laser: Room Temperature cw Operation with an Injection Laser Pump," Appl. Optics, vol. 13, pp 1256-1258, 1974.

[6]   J. Stone and C. A. Burrus, "Self-Contained LED-Pumped Single-Crystal Nd:YAG Fiber Laser," Fiber and Integrated Optics, vol. 2, pp 19-46, 1979.

[7]   R. H. Kingston, Detection of Optical and Infrared Radiation," Springer-Verlag, (Berlin), 1978, Ch 8.

[8]   A. Yariv, Optical Electronics, Third Ed., Holt, Rinehart and Winston, (New York), 1985, Appendix D.

[9]   M. Yamada, M. Shimizu, T. Takeshita, M. Okayasu, M. Horiguchi, S. Uehara, and E. Sugita, "Er^{3+} –Doped Fiber Amplifier Pumped by 0.98 μm Laser Diodes," IEEE Photon. Technol. Lett., vol 1, pp 422-424, 1989.

[10] C. R. Giles, E. Desurvire, J. L. Zyskind, and J. R. Simpson, "Noise Performance of Erbium-doped Fiber Amplifier Pumped at 1.49 μm, and Application to Signal Preamplification at 1.8 Gbits/s," IEEE Photon. Technol. Lett., vol 1, pp 367-369, 1989.

[11] R. Olshansky, "Noise Figure For Erbium-Doped Optical Fibre Amplifiers," Electron. Lett., vol. 24, pp 1363-1365, 1988.

[12] M. Yamada, M. Shimizu, M. Okayasu, T. Takeshita, M. Horiguchi, Y. Tachikawa, and E. Sugita, "Noise Characteristics of Er^{3+} doped Fiber Amplifiers Pumped by 0.98 and 1.48 m Laser Diodes," IEEE Photon. Technol. Lett., Vol 2, pp 205-207, 1990.

[13] D. P. Bour, N. A. Dinkel, D. B. Gilbert, K. B. Fabian, and M. G. Harvey, "980 nm Diode Laser for Pumping Er^{3+} Doped Fiber Amplifiers," IEEE Photon. Technol. Lett., Vol 2, pp 153-155, 1990.

[14] A. Takada, K. Iwatsuki, and M. Saruwatari, "Picosecond Laser Diode Pulse Amplification up to 12 W by Laser Diode Pumped Erbium-Doped Fiber," IEEE Photon. Technol. Lett., vol 2, pp 122- 124, 1990.

[15] N. Henmi, Y. Aoki, S. Fujita, Y. Sunohara, and M. Shikada, "Rayleigh Scattering Influence on Performance of 10 Gb/s Optical Receiver with Er-Doped Optical Fiber Preamplifier," IEEE Photon. Technol. Lett., vol 2, pp 277-278, 1990.

[16] E. Desurvire and J. R. Simpson, "Amplification of Spontaneous Emission in Erbium-Doped Single-Mode Fibers," IEEE J. Lightwave Tech., vol 7, pp 835-845, 1989.

[17] E. Desurvire, "Spectral Noise Figure of Er^{3+} – Doped Fiber Amplifiers," IEEE Photon. Technol. Lett., vol 2, pp 208-210, 1990.

[18]   M. Shimizu, M. Yamada, M. Horiguchi, and E. Sugita, "Concentation Effect on Optical Amplification Characteristics of Er-Doped Silica Single-Mode Fibers," IEEE Photon. Technol. Lett., vol 2, pp 43-45, 1990.

[19]   E. Desurvire, "Analysis of Erbium-Doped Fiber Amplifiers Pumped in the $^4I_{15/2} - \, ^4I_{13/2}$ Band," IEEE Photon. Technol. Lett., vol 1, pp 293-296, 1989.

[20]   E. Desurvire, J. R. Simpson, and P. C. Becker, "High-gain Erbium-doped Traveling-wave Fiber Amplifier," Optics Lett., vol 12, pp 888-890, 1987.

[21]   M. J. F. Digonnet and C. J. Gaeta, "Theoretical Analysis of Optical Fiber Laser Amplifiers and Oscillators," Appl. Optics, vol 24, pp 333-342, 1985.

[22]   H. Po, F. Hakimi, R. J. Mansfield, B. C. McCollum, R. P. Tumminelli, E. Snitzer, "Neodymium Fiber Laser at 0.905, 1.06 and 1.4 μm, "Abstracts of Annual Meeting of Optical Soc. of America, Seattle, WA (1986) Paper FD4, p 103.

[23]   P. R. Morkel, M. C. Farries and S. B. Poole, "Spectral Variation of Excited State Absorption in Neodymium doped Fibre Lasers," Optics Comm., vol 67, pp 349-352, 1988.

[24] E. Snitzer, H. Po, F. Hakimi, R. P. Tumminelli, B. C. McCollum, "Double Clad, Offset Core Nd Fiber Laser," Optical Fiber Sensor Conf. Postdeadline Paper, New Orleans, LA (1988) Paper PD5.

[25] H. Po, E. Snitzer, R. P. Tumminelli, L. Zenteno, F. Hakimi, N. M. Cho, T. Haw, "Double Clad High Brightness Nd Fiber Laser Pumped by GaAlAs Phased Array," Optical Fiber Communication Conf. Postdeadlline Paper, Houston TX (1989) Paper PD2.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
(AFOSR)

conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Reusable Ada Software -

Evaluating the Common Ada Missile Packages (CAMP-3)

Prepared by: Brian J. Shelburne Ph.D.

Academic Rank: Associate Professor

Department and   Mathematics and Computer Science

 University   Wittenberg University

Research Location: WRDC/AAAF-3
                          Wright-Patterson AFB
                          Dayton, Ohio 45433

USAF Researcher: Marc J. Pitarys

Date: August 17, 1990

Contract No   F49620-88-C-0053

Reusable Ada Software –

Evaluating the Common Ada Missile Packages (CAMP-3)

by

Brian J. Shelburne

ABSTRACT

One of the largest and earliest projects involving reusable Ada software was the United States Air Force sponsored CAMP effort with McDonnell-Douglas Corporation. This summer's AFOSR project evaluated CAMP for its usefulness and suitability for avionics applications.

During the process of evaluation, errors were discovered in some of the CAMP software parts. The tight dependencies among the various CAMP parts caused by "withing" and the poor internal documentation made tracking down these errors extremely difficult.

CAMP is overly complex, poorly documented, and contains errors. The final conclusion arrived at is that CAMP software is not suitable for avionics applications.

Acknowledgements

I wish to thank the Air Force Systems Command (Avionics Laboratory) and the AFOSR for sponsorship of this research. Universal Energy Systems must be mentioned for their concern and help to me in all administrative and directional aspects of this program.

I wish to especially recognize Marc Pitarys for his guidance and assistance with this project and to thank him for the many and fruitful discussions we had on the problems and pitfalls related to software reusability. I also want to thank Kenneth Littlejohn and James Williamson both for their assistance with my project and for the many fruitful discussions we had.

I.   INTRODUCTION:

A concise explanation of CAMP found on page 195 of Appendix I of

Developing And Using ADA Parts in Real-Time Embedded Applications

states

> "The main goal of the CAMP program has been to
> establish the feasibility and value of reusable Ada
> software within the mission critical real-time domains.
> This has required a careful evaluation of a particular
> domain, the development of reusable components, the
> development of automated support for software reuse in
> the software development lifecycle, and the application
> of both reusable components and the automated tools to a
> realistic application."

There are three phases to the CAMP project. Again, to quote from

the above document

> "[T]he CAMP program began in 1984, with a 12-month
> feasibility study. There were two major objectives: (1)
> to determine if sufficient commonality existed within
> the missile operational flight software to warrant the
> development of reusable software parts; and (2) to
> determine the aspects of parts engineering that could be
> fully or partially automated, and to develop the
> requirements and top-level design for a parts
> composition system to support reuse".

> "While CAMP-1 concentrated on feasibility analysis,
> Phase 2 of the CAMP program (CAMP-2) was a 32-month
> technology demonstration phase that began in September
> 1985. The goal of CAMP-2 was to demonstrate the
> technical feasibility and value of reusable Ada missile
> parts and a PCS [Parts Composition System] by building
> and using them on a realistic application."

It was during the CAMP-2 phase that 454 software parts were

identified and cataloged from 10 different missile systems. A

prototype parts composition system (called AMPEE or Ada Missile

Parts Engineering Expert system) was constructed to manage these

parts for later reuse. The CAMP parts and the AMPEE system were

then used to construct the navigation and guidance systems of an 11th missile application which was tested in a MIL-STD 1750A hardware-in-the-loop simulation. At the same time a suite of "armonics" benchmarks were developed from the CAMP parts to measure the effectiveness and efficiency of Ada compilers for armonics applications.

The third phase of CAMP (CAMP-3) was technology transfer. The CAMP parts were extended to over 500 in number and the AMPEE parts composition system was re-engineered, rewritten in Ada, and renamed the Parts Engineering System (PES). A user's manual, Parts Engineering System Catalog User's Guide - Version 1.1 was also written to support it.

II. OBJECTIVES:

The objectives of my research were threefold: learn the CAMP Parts Engineering System, use it to generate a sample avionics application, and evaluate the usefulness of CAMP.

III. THE PARTS ENGINEERING SYSTEM (PES)

The CAMP Parts Engineering System (PES) is used to access the CAMP parts database. It has two basic functions : submission of new parts into the database and user examination of parts currently in the database. A part is an Ada package, subroutine or task, often a generic. The exact definition is unclear but one criterion for a part is that it must be able to stand alone.

Parts can and frequently do "with" other parts.

PES is menu driven and easy to use. It has some built-in help facilities. The written documentation for PES is well written.

PES accesses the parts database by letting the user generate a search list of parts to examine. There are 24 different search criteria (e.g. part ID, part number, part name, keyword, etc.) with a limited capacity to generate compound search queries using Boolean AND's and OR's. Once a list of parts is selected, the user has the option of examining any of the 22 attributes of a part in the current search list or examining the source code of the part (specification or body). Examples of attributes are abstract, keywords, classification, projects used by, sample usage etc. Final selection of a part for use generates a small text file containing the specification and body file names plus compilation instructions.

III.1. Critique of PES :

The limited screen capabilities of the VT220 terminal made PES slow and awkward to use particularly with the way it generated new screens. A window-like environment would be better.

Of the 24 different search criteria, only two, search by keyword and search by classification, were useful. Of the 22 attributes attached to a part, only three, the abstract, compilation instructions and sample usage, were useful. In many cases the abstract was too sketchy to be of much use.

For a given part, one could either examine the specification file or the body file. Unfortunately CAMP grouped many different parts into one specification file so a search of the specification file had to be made to find the part specification. Fortunately PES used a VAX editor to access CAMP part files so one could use the "find" command to locate the specification. Unfortunately, many body files use the Ada "is separate" feature to shift the part code into a second file making the code inaccessible from PES. The only way to find the code was to exit PES, do an operating system level search (such as the directory command) for a file name that was similar to the body file name, and then access that file outside of PES. Fortunately CAMP's file naming scheme made such a search easy.

Each CAMP part source code file was highly structured in term of comments and a standard format for comments was followed for all files. Many of the comments were not particularly useful (e.g. the entire revision history) while other comments that would have been useful were missing (e.g. the mathematics underlying the algorithm used to implement the part). There was no documentation on what conditions must exist before a subroutine was called or on what conditions were true after a subroutine completes. In any case, the large quantity of comments tended to get in the way and long searches were often necessary to find code. There were also errors in the comments, generally minor annoying but they could cause problems if one was not familiar with the parts database or with that part's function. PES also referenced parts whose files did not exist; sometimes PES simply

had the wrong file name.

One final and very annoying feature of the specification and body files was that they were often more than 80 columns wide so they would not fit on an 80 column screen.

The most telling fact about PES was the way I eventually got around having to use it. I made a hardcopy of the "taxonomy" of the parts data base which gave me a classification of the CAMP parts and a general overview of all parts. (This taxonomy was used by PES as the classification search criterion.) I then made a hardcopy listing of the directory of specification files. The standardized file naming conventions for CAMP parts made this easy to do. Then as needed for each specification file, I made a hardcopy listing of the body files for that specification. Fortunately CAMP file names were meaningful. In this way I could examine the CAMP parts with an editor without using PES. Since the part attributes provided by PES were in the source code as comments, I still had access to them.

It is significant that a simple way around PES was found that allowed easier and better access to the parts database.

IV. BUILDING APPLICATIONS WITH CAMP PARTS

The main goal was to study the suitability of CAMP parts for avionics software. To accomplish this goal and to get a feel for using CAMP parts, three small programs of increasing complexity and a final avionics application were written.

The three small programs were

1. A square root program using CAMP's math routines
2. A program to calculate actual north and east velocity given nominal north and east velocity and the wander angle
3. A program to calculate great circle distance between two points given their latitude and longitude.

The final avionics application was a very simple "waypoint navigation" program. It accepted a sequence of coordinates (latitude and longitude) for the starting point, waypoints, and terminating point of the route of an aircraft then computed the great circle distance for each segment of the route and the turning angle at each waypoint.

The idea for the last two programs and the avionics application were suggested by various CAMP parts.

IV.1. Square Root Program

The program was to read a real number, use the CAMP part to calculate the square root, and then print it. Since the specification of the CAMP square root part stated that the exception "Negative_Input' would be raised if the argument to the square root function was negative, an exception handler was written.

Unfortunately or negative input, the "Negative_Input" exception was not raised as it should have been. Instead a "Constraint_Error" exception was raised which crashed the program.

This program revealed the first serious problem with CAMP; that is, the complexity of using the CAMP parts.

CAMP parts are, for the most part, generic packages and procedures. Many of them "with" other parts and the resulting web of interconnections and dependencies can be complex and confusing. Tracing down the origin of the square root constraint error required going back through seven files only to discover that CAMP used the standard VAX Ada square root function. It did not reveal why the "Negative_Input" exception was not raised.

The second of the seven files, General_Purpose_Math, contained the line "package body Square_Root is separate". The third of the seven contained the line "separate General_Purpose_Math" thus bridging the link from the second to the third file. Unfortunately, there happened to be an (unknown) eighth file which also contained the line "separate General_Purpose_Math" and this was the one that was used. There was no indication that this other file existed and indeed the file was found by accident only after I was forced to cast around for some other outside "source" of the error. The file contained a different method to implement a square root function and its documentation stated 'The exception "Negative_Input": is raised if "Input" is negative'. This was false; the code failed to do this.

The unknown eighth file was used instead of the third file because it was apparently compiled later. This brings up a second problem with CAMP parts aside from the heavy use of "withing" and the difficulty of back-tracking problems; that is

the use of Ada's "separate" clauses that lead to configuration management abuses like having multiple "separate" files. The third problem is simply a lack of clear concise part dependency documentation. Fourth, there is no warning that if one uses a part, it won't generate a tangled web of dependencies. Finally, using hidden code in implementing a part increases its complexity and the likelihood of an error. The part user has no knowledge of and no control over this.

IV.2. NE Velocity Program

The main work for this program was done by two CAMP generic procedures contained in the package Wander_Azimuth_Navigation_Parts. Both generic procedures had three type parameters and two subprogram parameters.

In designing parts, CAMP uses an approach known as the "semi-abstract data type" method, a method based on generics and use of overloaded operators. It means that many CAMP parts are generics that are tailorable to user defined data types but with CAMP providing a set of default types and operators. (An excellent description of the Parts Design Alternatives is found in Appendix III of the CAMP manual Developing And Using Ada Parts in Real Time Embedded Applications.)

Instantiating the generic package required the user to declare three data types and two subroutines as generic parameters. This insures that strong typing, one of Ada's best features, is maintained but it requires more work to instantiate the generic. However, CAMP provides a part called Basic_Data_Types which

provides many of the necessary data types and operators that operate on them. To assist the user, each CAMP part documents a "sample usage" on how to use the default data types for that part.

The NE Velocity program was written using the sample usage examples given in the source code since the easiest way to write the program was to follow CAMP's suggested guidelines. However, the program would not compile due to an "inconsistency detected during overload resolution" error when trying to find an actual parameter corresponding to one of the generic formal procedure parameters.

The problem is somewhat complicated to explain but it has to do with instantiation of generics. According to the Ada Language Reference Manual, if a generic package is instantiated and a new type is derived from the instantiation then subroutines from the instantiation ought to be visible with parameters of the derived type. This does not seem to be the case.

The problem once understood was easy to fix but either the CAMP documented sample usage is wrong or there is an Ada compiler problem. (This example was run on two different compilers and the same error occurred. The CAMP manuals did mention problems with immature compilers instantiating generics.) Whether the error was actually a compiler problem or the result of incorrect code was never determined.

IV.3. Great Circle Distance Navigation Program

This program implemented the CAMP package Great_Circle_Arc_Length and again the documented sample usage was used to guide the implementation. Using the lesson learned from the NE Velocity program, the "overload resolution inconsistency" problem was avoided.

However, when run, the program sometimes crashed with a constraint_error raised by a "Modified_Newton_Raphson" routine for square root. Use of a debugger eventually traced down the error to a calculation of 1.0000000000000000000000001295525 for a square root (obviously using quadruple precision) where the output value was constrained to the range -1.0 .. 1.0. which generated the constraint error.

Again because of the tangle of "withed" parts and part dependencies, the error was very difficult to trace. Because a lot of parts were hidden, the occurrence of this particular error came as a complete surprise.

The error did reveal a fundamental design flaw : a square root routine whose design guaranteed only a certain amount of precision (which was not documented) was "misused" by permitting higher precision arithmetic internally.

V. WAYPOINT NAVIGATION APPLICATION

The Waypoint Navigation Problem : Given the earth coordinates (latitude and longitude) of three points A, B, and C, find the great circle distances AB, BC, and the turning angle at point B.

This particular application was chosen because CAMP had the parts that would make the implementation fairly easy. The one drawback was that the mathematics behind the particular method CAMP used to compute great circle distances and turning angles was not documented. The mathematics had to be extracted out of the code which again was made difficult by the complex web of CAMP part dependencies. Understanding the math behind the CAMP parts was important because the exact meanings of the part parameters was not well documented.

Unfortunately the CAMP package subroutines for Waypoint Navigation contained a fundamental error in the method used to calculate turn angles. Thus the application never worked.

VI. EVALUATION

CAMP is overly complex, poorly documented, and contains errors.

The complexity of the CAMP package derives from heavy usage of "withing" which generates a complex web of part dependencies. This is complicated by use of Ada's "is separate" which means that a part might be scattered over three or more files.

The complex web of part dependencies has three drawbacks. First, tracking down an error where there is no indication of what file the part is in or even what files the part depends on is difficult, time consuming, and frustrating. Second, the ramifications of the use of any part are unknown. The user has no feel for how or why the part works since too much is invisible

to him or her. Third, it makes for more complicated and slower running code if a part is implemented using calls to other parts. Complicated code is more susceptible to error.

Another aspect of the complexity of CAMP parts deals with the "semi-abstract data type" method of parts design. Many CAMP parts were generic units which were tailorable to user-defined types. Using a part required defining a number of data, constant, and subroutine parameters to be used in the instantiation of the generic. While use of CAMP's default types took some of this burden off the user, it was not clear why all were needed or even why certain distinctions are made. Default types always added to the complex web of dependencies between parts. In all fairness to the CAMP authors, they took the best way out of a difficult situation. While generics aid software usability, Ada's strong typing requirements are hostile to it.

Complexity of part dependencies and loss of visibility into the inner workings of a part decreases user confidence in the part. While one may argue that a specification is all that is needed to use a part, any responsible programmer would at least want to check out the code to be assured that it works.

The specification of a part should explain clearly, concisely, and exactly what the part does. It should state the constraints required on any parameters and it should guarantee the precision of any output. Only with this will a user have confidence in it. The CAMP parts fail to do this.

In general the comment documentation was well laid out and

followed a standard format but there was too much useless documentation. The mass of documentation got in the way when one was trying to find a particular code segment.

Finally there were errors in the CAMP parts. Each of the four programs written uncovered problems ranging from errors in the parts documentation to serious mathematical design errors.

VII. CAMP AND REUSABLE SOFTWARE

My experience with CAMP was not one of success. I found too many errors that in many cases were difficult to track down. I lost confidence in CAMP.

For software reusability to succeed, a user must be able to use a reusable part having the assurance that it will take less time to implement and that it will be correct. The parts and all its inner workings must be totally visible to the user so that he or she knows the exact behavior of the part and so that he or she can be assured that the part works correctly. The part must be a white box, not a black box, in order that the user may have confidence in it. CAMP parts, by and large, are black boxes.

Regarding software reusability, there is one fundamental issue in that must be addressed. When a software engineering project is undertaken, choices are made at all levels of the design process from the highest level of abstraction to the lowest level of implementation. Two designers given the same specification will come up with two equally good designs yet at some point the

designs will diverge and from that point become increasingly incompatible so that it is impossible to reconcile one design with the other.

This raises the issue of whether it is even possible to take the software parts from one design domain (armonics) and use them in a second design domain (avionics).

As I became more and more acquainted with CAMP parts I began to perceive the underlying design choices that were made. When I wrote my programs I had to reconcile my own design tendencies with the design of CAMP parts. In each instance I had to ask the question, will my design take less time and be more assured of being correct or can I use a CAMP part and achieve the same degree of software quality in less time?

VIII. RECOMMENDATIONS

My recommendations address five issues : Parts Engineering Systems, documentation of software parts, quality of software parts, Ada reusable software, and software engineering issues.

VIII.1. PES Issues :

A Parts Engineering System supporting a reusable software library should use a windows-like interface. Hypertext might be a good approach.

A selection process should be used that allows the user to eventually narrow his or her search to the examination of a few

parts. Selection criteria used to search for parts should be useful and meaningful.

A PES should display the salient information about a part so that the user can decide quickly whether to investigate it further or reject it.

VIII.2. Documentation

Documentation for a software part should state the exact constraints required on parameters and it should guarantee the precision of output. The algorithm or implementation used by the part should be explicitly stated or a reference given for it. Documentation of sample usage should be included. Dependencies on other parts should be explicitly stated. The user should not be overwhelmed by useless documentation. Quality documentation is a necessary precondition for a Parts Engineering System.

VIII.3. Quality

Reusable software parts should be totally visible to the user so that the user may understand its function, proper application, and correctness. Reusable software parts should be white boxes, not black boxes.

Reusable software parts should not be overly dependent on other parts. In addition, any dependency should not go back more than one level. All dependencies should be clearly stated in the documentation.

VIII.4. Ada and Reusable Software Parts

The use of Ada's "is separate" feature should be avoided.
The grouping of many software parts into a single file should be avoided.

Reusable software is supported by Ada generics but inhibited by Ada's strong typing. CAMP's "semi-abstract data type" based on the generic and overloaded methods (see page 212-213 of Developing and Using Ada Parts in Real-Time Embedded Applications) is a reasonable compromise provided software part dependencies are minimized and well documented.

VIII.5. Software Engineering

Software reusability across different application domains might not be possible. If different designers create diverging designs from the same specification, divergence will be much greater with different specifications.

Users of reusable software from a particular applications domain should be proficient in that domain in order to allow the rapid evaluation of the usefulness and correctness of a part. Human judgment should be part of the loop to insure the part performs the task correctly.

REFERENCES

Biggerstaff, T.J and Perlis, A.J. ed, <u>Software Reusability - Vol.</u>
<u>I Concepts and Models</u>; ACM Press, New York, 1989.

<u>Developing And Using Ada Parts in Real-Time Embedded Applications</u>
<u>(CAMP-3)</u>; McDonnell Douglas Missile System Company, 1990.

<u>Parts Engineering System Catalog User's Guide Version 1.1</u>;
McDonnell Douglas Missile System Company, 1990.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

COMPUTER SIMULATION OF NMOS INTEGRATED CIRCUIT CHIP

PERFORMANCE INDICATORS

Prepared by:	Ashok K. Goel, Ph.D.
Academic Rank:	Assistant Professor
Department and	Department of Electrical Engineering
University:	Michigan Technological University
Research Location:	Electronics Technology Laboratory
	WRDC/ELE
	Wright-Patterson Air Force Base
	Dayton, OH 45433-6543
USAF Researcher:	Dr. Fritz Schuermeyer
Date:	July 20, 1990
Contract No:	F49620-88-C-0053

COMPUTER SIMULATION OF NMOS INTEGRATED CIRCUIT CHIP
PERFORMANCE INDICATORS

by

Ashok K. Goel

ABSTRACT

For an integrated circuit chip based on the silicon NMOS technology, a computer-efficient model of the various chip performance indicators has been developed and a user-friendly computer program called "NCHIPSIM" suitable for the simulation of the chip performance indicators for an NMOS microprocessor or a gate-array chip has been developed. In addition to predicting the various chip performance indicators such as its maximum clock frequency, power consumption, computational capacity, power efficiency, fabrication yield, functional throughput rate and the size of an NMOS chip with the given technology parameters, the program NCHIPSIM has also been used to simulate the dependence of the various chip performance indicators on the technology feature size in the range 0.1-2.5 microns and the chip integration level in the range 100-100,000,000 transistors on the chip. The results have been compared with and found in excellent agreement with those known for several single-chip microprocessors based on the silicon NMOS technology.

ACKNOWLEDGEMENTS

First, I like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Electronics Technology Laboratory at the Wright-Patterson Air Force Base for sponsoring this research. I also wish to thank the Universal Energy Systems, Inc. for helping me with all the administrative and directional aspects of this program.

I also like to thank several individuals who helped in making this experience truly rewarding and enriching for me. First, I am grateful to Dr. Fritz Schuermeyer for providing me with support, encouragement and a truly enjoyable working atmosphere. Second, I am thankful to Mr. Don Peacock for his personal attention at several occasions during this work. The concern of Dr. Ben Murphy was greatly appreciated. The help of Mr. Darrell Barker, Mr. Ben Carroll and Mr. Tim Seiter was invaluable in overcoming several technical roadblocks.

I. INTRODUCTION:

It is extremely important to be able to simulate the performance of an integrated circuit chip before its fabrication is undertaken. This can be accomplished by executing the following steps:

a) Development of a computer-efficient model of the various chip performance indicators;

b) Development of a user-friendly computer program suitable for the chip performance simulation; and

c) Application of the computer simulator for the determination of the performance indicators for a chip with known values of the various technology parameters.

Continuous advances in the integrated circuit technology have resulted in more complex chips integrating millions of devices and interconnections. In the recent years, it has become necessary to use interconnections in two or more levels to achieve higher packing densities, shorter propagation delays and smaller chips. Further, because of the much higher mobility of electrons in Gallium Arsenide (GaAs), it has emerged as a preferred substrate for the development of the high speed circuits.

The Device Technology Branch of the Electronics Technology Laboratory at the Wright Patterson Air Force Base is interested in the development of very high speed integrated circuits based on the GaAs FET and other technologies. In

particular, they are concerned about the parasitic effects that adversely affect the performance of a high-speed high-density integrated circuit. My previous research work has been focused on the contribution of the parasitic effects on the crosstalk and the propagation delays in the GaAs MESFETs and the high-density interconnections on the GaAs-based very high speed integrated circuits and this contributed to my present summer assignment.

II. OBJECTIVES OF THE RESEARCH EFFORT:

The long range objective of this research is to develop the computer-efficient algorithms and the related user-friendly computer software modules suitable for the simulation of the performance indicators for the small-geometry high-speed high-density integrated circuit chips based on the Silicon NMOS, Silicon CMOS, Silicon BJT and the Gallium Arsenide FET, HBT and other high-speed technologies.

It was decided that, first, a computer-efficient model and a user-oriented computer program would be developed to predict the various performance indicators of a silicon NMOS chip such as its maximum clock frequency, power consumption, computational capacity, power efficiency, fabrication yield, functional throughput rate and its size. In addition to simulating the performance of a single-chip microprocessor or a gate-array with the given technology parameters, the program

should enable the user to analyze the effects of scaling the feature size or changing the integration level of the chip on its performance indicators. In order to validate the algorithm and the simulator, the simulation results will be compared with those for several existing single-chip microprocessors based on the silicon NMOS technology.

III. ALGORITHM DEVELOPMENT:

To date, an NMOS chip performance simulator called "NCHIPSIM" has been developed. In the algorithm, first, the average interconnection length on the chip in units of the average logic gate dimension is determined by the given number of transistors or gates on the chip and the Rent's constant by using the equations derived by Donath [1]. Then, for a logic intensive interconnection-capacity limited VLSI chip, the average logic gate dimension on the chip is obtained by setting the interconnection available per gate equal to that required per gate. Then the value of the average interconnect length on the chip is evaluated and the values of the chip size and the chip area were determined. Next, the average delay time in a gate on the chip is obtained by adding the various component delays such as the delay due to the output resistance of the driving gate and the interconnect capacitance, the delay due to the input capacitance of the gate at the next stage, the distributed-RC delay of the

interconnections and the delay due to the resistance of the interconnections and the input capacitance of the gates. Next, the total delay suffered by an input signal on the chip is determined by adding the delay through the logic gates, the contribution of the global interconnection delay, and the contribution of the speed-of-light limit depending on the propagation speed of the electromagnetic waves on the chip. The maximum clock frequency of the chip is then determined. Next, the power consumption of the chip is calculated by adding the power consumption in the logic gates and the dynamic power consumption at the I/O buffers. This depends on the power supply voltage, fraction of the on-chip gates that switch during a clock period, the total capacitance at an output pin and the number of pins per chip as determined by using the Rent's rule [2] or provided by the designer. Next, the computational capacity of the chip is determined by using the number of gates on the chip and its maximum clock frequency. Then the power efficiency of the chip is obtained by using the values of the computational capacity and the power consumption. Next, the functional throughput rate of the chip is obtained by using the values of the computational capacity and the chip area. Finally, for a known value of the density of defects on the chip, the fabrication yield of the chip is obtained by using the Price law [3]. The flowchart of the program "NCHIPSIM" is shown on the following two pages.

```
                        ┌─────────────┐
                        │    START    │
                        └──────┬──────┘
                               │
                               ▼
              ┌────────────────────────────────┐
              │ Choose Chip type:              │
              │ Microprocessor  or  Gate-Array │
              └────────────────┬───────────────┘
                               │
                               ▼
              ┌────────────────────────────────┐
              │ Choose Interconnection and     │
              │ Dielectric Materials           │
              └────────────────┬───────────────┘
                               │
                               ▼
              ┌────────────────────────────────┐
              │ Define Chip, Transistor and    │◄──────────┐
              │ Interconnection Parameters     │           │
              └────────────────┬───────────────┘           │
                               │                            │
                               ▼                            │
         ┌─────────────────────────────────────────┐       │
         │ Calculate Interconnection Resistance,   │       │
         │ Interconnect Capacitance, Gate Resistance,│     │
         │ Number of Gates and Pins on the Chip    │       │
         └─────────────────────┬───────────────────┘       │
                               │                            │
                               ▼                            │
         ┌─────────────────────────────────────────┐       │
         │ Calculate Average Interconnect Length,  │       │
         │ Gate Dimension, Chip Size and Chip Area │       │
         └─────────────────────┬───────────────────┘       │
                               │                            │
                               ▼                            │
         ┌─────────────────────────────────────────┐       │
         │ Calculate Typical Gate Delay and        │       │
         │ Maximum Clock Frequency                 │       │
         └─────────────────────┬───────────────────┘       │
                               │                            │
                               ▼                            │
      ┌────────────────────────────────────────────┐       │
      │ Calculate Power Consumption, Power Efficiency,│     │
      │ Throughput Rate and Fabrication Yield      │       │
      └────────────────────────┬───────────────────┘       │
                               │                            │
                               ▼                            │
              ┌────────────────────────────────┐            │
              │ Write Results on Screen        │            │
              │ and NCHIPSIM.OUT               │            │
              └────────────────┬───────────────┘            │
                               │                            │
                               ▼                            │
                          ╱─────────╲        YES            │
                         ╱ Redefine  ╲──────────────────────┘
                         ╲ Reference ╱
                          ╲  Chip?  ╱
                           ╲───┬───╱
                               │ NO
                               ▼
```

(Continued on Next Page)

FLOW CHART FOR "NCHIPSIM"

IV. THE PROGRAM "NCHIPSIM":

Using the flowchart and the steps outlined in the above
section, an extremely flexible and user-friendly program
called "NCHIPSIM" suitable for the simulation of the various
performance indicators for an NMOS integrated circuit chip
with known technology parameters was written in FORTRAN-77 and
run successfully on the VAX mainframe computer. First, the
program allows the user to choose the chip type, a
microprocessor or a gate array. Depending on the chip type,
the program lists the constants, found empirically, required
for determining the average interconnect length and the number
of pins on the chip. Next, the program lists, but allows the
user to change interactively, the values of several chip
parameters such as the number of transistors, fan-out of a
typical gate, capacitance at the output pin, density of
fabrication defects and the probability of an on-chip gate to
switch during a clock period. Next, the program lists the
typical values of the transistor related parameters such as
its feature size, input capacitance, output gate resistance,
power supply voltage and the ratio of optimum-size to feature-
size transistors but permits the user to change any of these
values for the chip being simulated. Next, the user can choose
a dielectric material out of silicon dioxide, polyimide,
alumina and epoxy glass or select one of his own and define
its dielectric constant interactively. Next, the user can

choose an interconnection material out of aluminum, copper,
silver, tungsten and molybdenum or select any other material
and define its electrical resistivity interactively. Next, the
program lists the typical values of the other interconnection
parameters [4] for the technology feature size defined earlier
but allows the user to modify any of these values
interactively. These parameters include width, pitch and
thickness of on-chip interconnects, thickness of the
dielectric material, number of interconnection layers and the
utilization coefficient of the interconnections. Then, for the
chip defined above, the program calculates and displays the
values of its performance indicators such as its size, the
maximum clock frequency, power consumption, computational
capacity, power efficiency, functional throughput rate and the
fabrication yield. Next, the program allows the user to scale
the reference chip defined above by a certain scaling factor
and determines the performance indicators for the scaled chip.
Finally, the user can change the number of transistors on the
reference chip and study the dependence of the various chip
performance indicators on its integration level. In addition
to displaying the simulation results on the screen, the
program also writes the various chip parameters and the
corresponding results for the reference chip, scaled chips and
the chips with different integration levels on an output file
called "NCHIPSIM.OUT".

V. SIMULATION RESULTS:

The program "NCHIPSIM" has been used to compare the simulation results for several actual NMOS microprocessor chips to the available data [5]. For example, such a comparison for the 1.5-micron NMOS microprocessor chip called HP Focus (1982) is shown in Table 1 and that for the 3-micron NMOS microprocessor chip called Stanford MIPS (1984) is shown in Table 2. The program has also been used to study the dependence of the chip performance on its minimum feature size as well its integration level i.e. the number of transistors or gates on the chip. For example, Figures 1 and 2 show the dependences of the maximum clock frequency and the power consumption of an NMOS microprocessor chip on its minimum feature size in the range 0.1-2.5 microns while Figures 3 and 4 show the maximum clock frequency and the power efficiency for a 1-micron NMOS microprocessor chip on the number of transistors on it in the range 100-100,000,000 respectively.

VI. RECOMMENDATIONS:

At present, the program "NCHIPSIM" is expected to simulate NMOS integrated circuit chips of feature size 1 micron or greater quite accurately. However, for simulating the submicron chips correctly, the algorithm used in NCHIPSIM must be modified to include the high-density effects.

Thereafter, this work must be extended and the appropriate mathematical algorithms and the related software modules must be developed for the integrated circuit chips based on the silicon CMOS, silicon BJT, GaAs FET, GaAs HBT and the other state-of-the-art VHSIC technologies. It is important that, for the small-geometry high-density chips, the capacitance calculation routines include the contributions of the fringing fields, coupling capacitances with the neighbors and the effects of shielding by the neighboring conductors. For the high-frequency GaAs chips, the interconnection resistance calculations should consider the skin effect, the gate delay calculations should include the longitudinal as well as the transverse delays in the field effect transistors and the interconnection delay calculation. should include the effects of the distributed capacitances and inductances in the on-chip interconnections.

It was decided that the author will continue with the summer research effort and carry out the research work recommended above at the author's institution with funding from the Research Initiation Program of the Air Force Office Of Scientific Research.

REFERENCES

1. Donath, W.E., "Placements and Average Interconnection Lengths of Computer Logic," IEEE Trans. Circuits and Systems, Vol. CAS-26, April 1979, pp. 272-277.

2. Landman, B.S. and R.L. Russo, "On a Pin versus Block Relationship for Partitions of Logic Graphs," IEEE Trans. Computers, Vol. C-20, Dec. 1971, pp. 1469-1479.

3. Price, J.E., "A New Look at Yield of Integrated Circuits," Proceedings of IEEE, Vol. 58, Aug. 1970, pp. 1290-1291

4. Bakoglu, H.B. and J.D. Meindl, "Optimal Interconnection Circuits for VLSI," IEEE Trans. Electron Devices, Vol. ED-32, No. 5, May 1985, pp. 903-909.

5. Toong, H.D. and A. Gupta, "An Architectural Comparison of Contemporary 16-Bit Microprocessors," IEEE Micro, Feb. 1983, pp. 26-38.
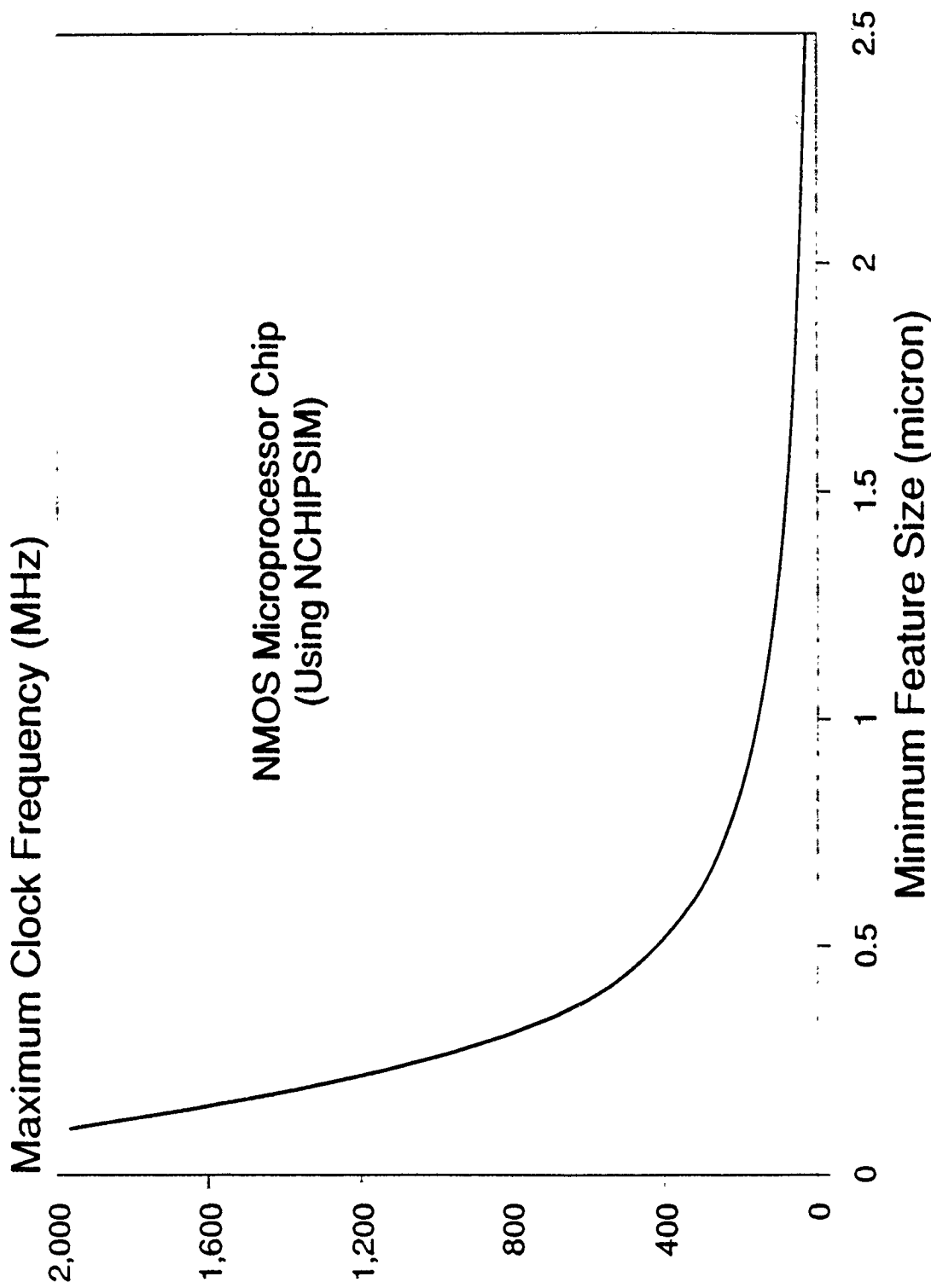
Figure 1: NCHIPSIM prediction of the dependence of the maximum clock frequency of a 200,000-transistor NMOS microprocessor chip on its minimum feature size in the range 0.1-2.5 microns.
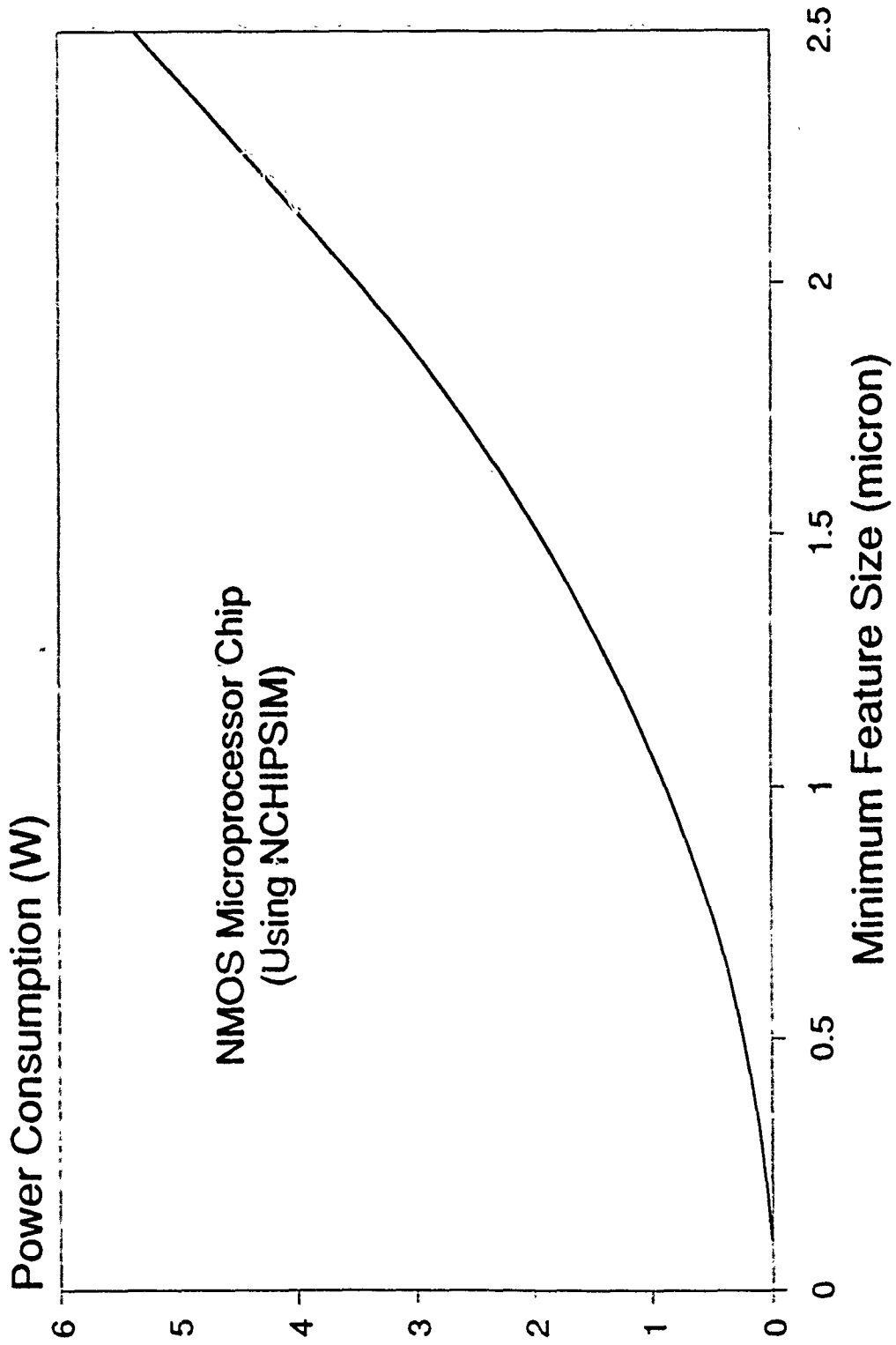
Figure 2: NCHIPSIM prediction of the dependence of the power consumption of a 200,000-transistor NMOS microprocessor chip on its minimum feature size in the range 0.1-2.5 microns.
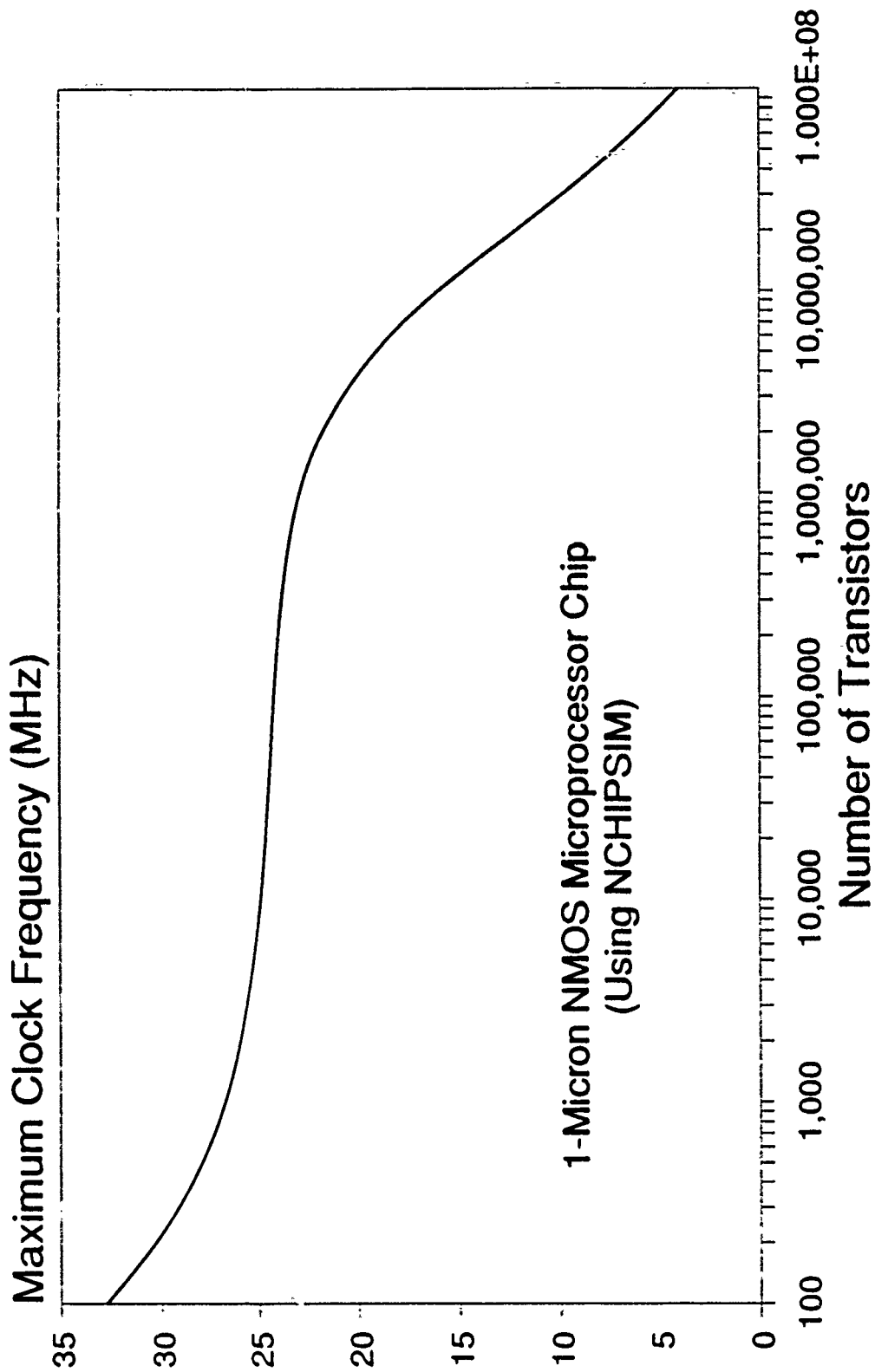
Figure 3: NCHIPSIM prediction of the dependence of the maximum clock frequency of a 1-micron NMOS microprocessor chip on the number of transistors on it in the range 100-100,000,000.
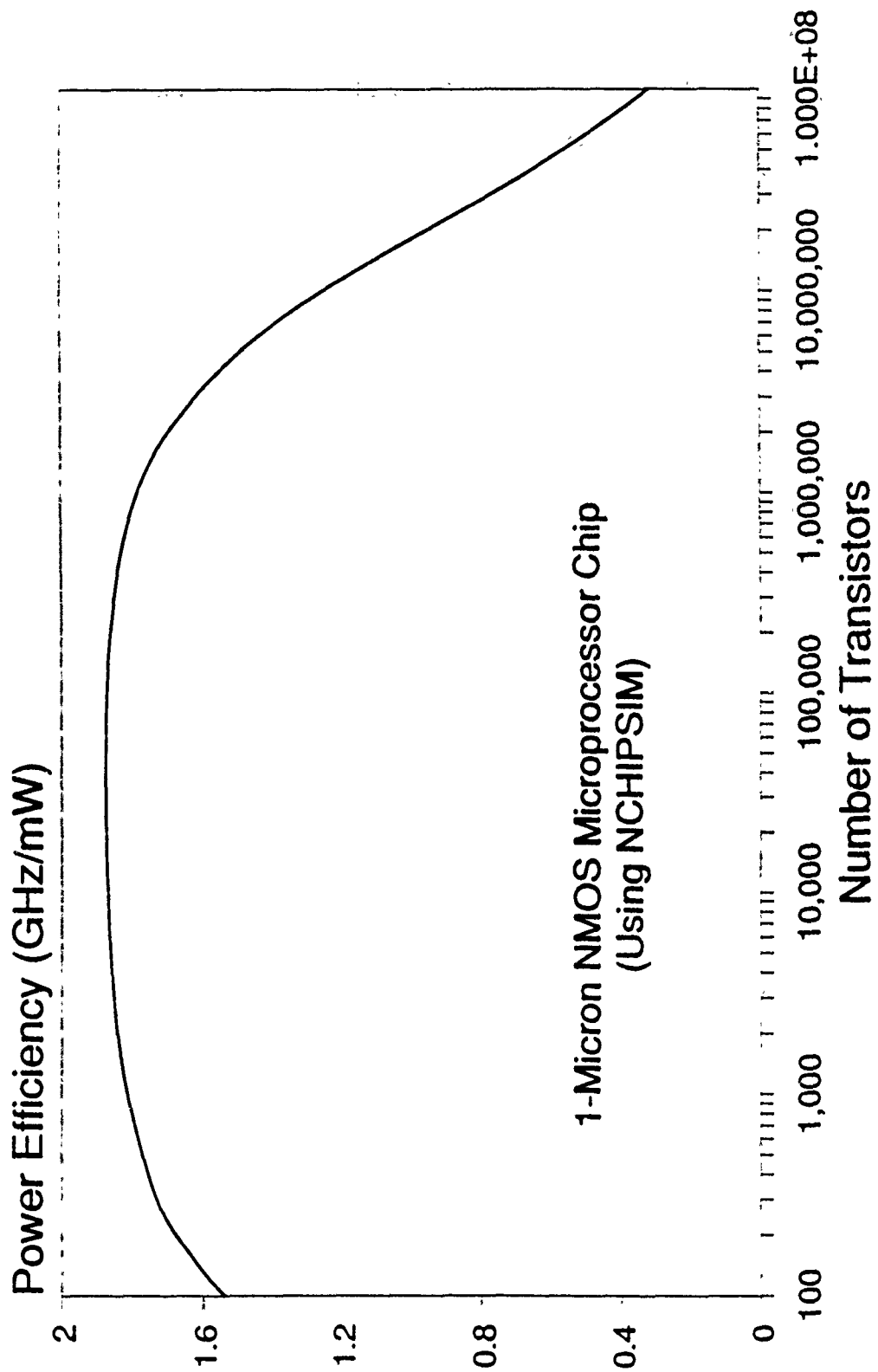
Figure 4: NCHIPSIM prediction of the dependence of the power efficiency of a 1-micron NMOS microprocessor chip on the number of transistors on it in the range 100-100,000,000.

TABLE 1: Actual Data and Simulation Results for NMOS Microprocessor Chip HP FOCUS (1982)

Performance Indicator	Actual Value	Using NCHIPSIM	
Chip Size	5.6	5.43	(mm)
Maximum Clock Freq.	18	22.68	(MHz)
Power Consumption	4	4.61	(Watt)
Computational Cap.	*	2551.8	(GHz)
Power Efficiency	*	0.55	(GHz/mW)
Functional Throughput	*	86.63	(GHz/mm^2)
Fabrication Yield	*	40.44	(%)

Chip Parameters

Minimum Feature Size = 1.5 microns
Number of Transistors = 450,000
Interconnect Layers = 3
Number of Pins = 83
Interconnection Pitch = 2.5 microns

* Data is not available

TABLE 2: Actual Data and Simulation Results for NMOS Microprocessor Chip STANFORD MIPS (1984)

Performance Indicator	Actual Value	Using NCHIPSIM	
Chip Size	5.8	6.7	(mm)
Maximum Clock Freq.	4	4.63	(MHz)
Power Consumption	2	2.07	(Watt)
Computational Cap.	*	27.76	(GHz)
Power Efficiency	*	13.43	(MHz/mW)
Functional Throughput	*	619.1	(MHz/mm^2)
Fabrication Yield	*	30.85	(%)

Chip Parameters

Minimum Feature Size	= 3.0 microns
Number of Transistors	= 24,000
Interconnect Layers	= 2
Number of Pins	= 84
Interconnection Pitch	= 9.0 microns

* Data is not available

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal energy systems, Inc.

FINAL REPORT

Application of Photoreflectance to Novel Materials

Prepared by: Muhammad Z. Numan

Academic rank: Assistant Professor

Department and   Physics Department

University: Indiana University of

                          Pennsylvania

Research location: USAF/WRDC

                          Wright Patterson A.F.B.

                          Dayton, OH 45433

USAF Researcher: M. Omar Manasreh

Date: Sept 15, 1990

Contract No: F49620-88-C-0053

Application of Photoreflectance to Novel Materials

by

Muhammad Z. Numan

ABSTRACT

Photoreflectance spectroscopy was applied to the InGaAs/GaAs single quantum well structures of different well thickness and to the low temperature molecular beam epitaxy grown GaAs cap layers on both n- and p- type GaAs substrates. The PR spectra at both room temperature and 77K have been studied. The GaAs study clearly indicates a lowering of surface potential associated with the unpinning of the Fermi levels reported for these systems. Both 200°C and 400°C caps demonstrate a disappearence of the Franz-Keldysh oscillation. Suggestions for future experiments are made.

Acknowledgements

I. INTRODUCTION:

Photoreflectance (PR) spectroscopy is becoming an important tool for the characterization of bulk semiconductors and semiconductor microstructures, such as quantum wells and superlattices[1]. Its popularity stems from the nondestructive, contactless, and room temperature nature of the technique, involving no special sample preparation or mounting requirements. PR found notable application in.measuring direct gap of compound semiconductors, intersubband transitions in multiple quantum wells (GaAs-AlGaAs, GaAs-InGaAs), and change in surface Fermi level of GaAs caused by photowashing[2].

At the WRDC/ELRA laboratory there exists excellent facilities for both room temperature and low temperature PR. In the past, summer fellows have used these facilities in fruitful collaboration with resident scientists[3]. Two of the areas of interest to the laboratories, namely, GaAs-InGaAs single quantum wells and low temperature MBE grown GaAs passivation layers, lend themselves to PR analysis with prospects of unique and important results. Current research have been undertaken to explore these possibilities.

My research interest in the past has spanned various experimental aspects of semiconductor materials and devices including structural (Rutherford backscattering and channeling) and electrical characterization. Recently, at IUP Physics department, we developed an optical characterization laboratory with funding from Research Corporation and local grants. A vacuum far infra-red FTIR system (Bio-rad/ Digilab FTS-40V) has been added to the lab. The summer fellowship will strengthen our program at IUP and establish fruitful collaboration between the labs in the areas of materials characterization and analysis.

II. OBJECTIVES OF THE RESEARCH EFFORT:

This summer project addresses two different materials systems of technological interest. The objectives in each area are delineated below. Firstly, we consider strained quantum wells of $In_xGa_{1-x}As$ in GaAs for room temperature PR study of conduction to valence band transitions of quantized energy levels. Strain dependent calculation of the transition energies as a function of well width can be compared with the experimental peaks in PR spectra to extract band offset for this system. We decided to grow MBE samples of $In_xGa_{1-x}As$/GaAs single quantum wells (x=.10 - .20) of well width ranging from 90A - 150A for PR measurements at room temperature and 70K. The results will be fitted with existing and new theoretical models to extract band offset parameter

Secondly, PR will be used to understand the reduced surface pinning of low temperature MBE grown GaAs layers. According to recent reports, the well known surface pinning effect of the Fermi level in GaAs is reduced by 75% when a low temperature MBE grown layer is used to cap the doped GaAs[4]. Photoreflectance is a sensitive tool in determining surface electric field in bulk materials[5]. For large fields, Franz-Keldysh oscillations appear in the PR spectra in contrast to the low field case[6]. If the surface field in the capped GaAs is significantly lowered a shift in the period of the FK oscillation will permit us to measure the change from room temperature PR spectra. Samples of both p-type and n-type MBE GaAs will be capped with 200A layers of low temperature material. PR spectra will be compared with control samples to see if a change in FK oscillation can be observed.

III. PHOTOREFLECTANCE DATA ANALYSIS:

In Photoreflectance spectroscopy, a chopped laser beam is used to modulate the built in electric field at the surface or an interface of the semiconductor (Fig. 1). The relative change in reflectance is then recorded

using lock in detecton method.   Photoreflectance spectra is customarily
analyzed by the third derivative functional form due to Aspenes[7],
according to:

$$\Delta R/R = Re[C\, e^{i\theta}(E - E_g + i\Gamma)^{-n}]\qquad\qquad(1)$$

where E is the energy of the probe beam, G is a broadening parameter for
the critical point energy E, and n refers to the type of critical point. Eqn (1)
applies to the case of low built-in electric field typified by the absence of
Franz-Keldysh oscillation.   In the presence of FK oscillation, $\Delta R/R$ takes an
asymptotic form:

$$\Delta R/R = \cos\{\tfrac{4}{3}[(E - E_g)/\hbar\Theta]^{3/2} + \phi\}\qquad\qquad(2)$$

where $\hbar\Theta$ is a characteristic energy,

$$(\hbar\Theta)^3 = e^2\hbar^2 F^2/2\mu.\qquad\qquad(3)$$

Photon energy at the m-th extrema, E_m, satisfies:

$$m p = \frac{4}{3}[(E_m - E_g)/\hbar\Theta]^{3/2} + \phi\qquad\qquad(4)$$

A straight line fit of $\frac{4}{3\pi}[(E_m - E_g)]^{3/2}$ vs. m gives the characteristic energy,
$\hbar\Theta$.   Surface electric field, F_s, can then be found from Eqn. (3).

IV. RESULTS:

Fig. (2)-(3) shows the room temperature and 80K photo-reflectance data
from the single quantum wells of well width 90A and 140A. We have not
been successful in analyzing this data using TDFF method.   A more
comprehensive fitting procedure is being developed.

Fig. (4)-(5) shows the results for the low temperature capped layers overlayed with the uncapped samples. The presence of FK oscillation is evident in case of the uncapped (control) samples, which disappears for the capped layers grown both at 200°C and 400°C. This result is a direct evidence of the reduction in band bending for the latter cases. In Fig. (5) G2-994 is the uncapped Be doped sample and G2-994 is the capped sample. G2-992 is the n-type uncapped sample of Fig. (4) for comparison of FK oscillation period in the two impurity types. In the p-type GaAs an additional low energy feature was noticed, which may be attributed to the shallow impurity levels.

V RECOMMENDATIONS:

The photoreflectance results for the low temperature MBE cap layers are very interesting and could resolve many unanswered questions about this technologically important material. Low temperature measurements should be performed on these samples to see if FK oscillation can be observed in the capped material. This will allow one to determine the build in field and the surface potential quantitatively. The low energy feature should be studied more carefully to ascertain its origin.

As to the quantum wells, the signals are not well resolved in the room temperature spectra. The MBE chamber was known to have some problem when these samples were grown. New samples should be grown carefully and the measurements repeated. To improve the apparatus, a new grating should be used in the monochromator of the room temperature system along with a Ge detector to sufficiantly extend the wavelength range to allow one to study higher mole fraction of Indium in the well. A variable neutral density filter can be placed in front of the monochromator as shown in Fig. (6) to have uniform intensity of the probe beam. A follow-up mini grant proposal will be submitted to achieve some of these recommendations.

REFERENCES

1. N. Botka, D.K. Gaskill, R. S. Sillmon, R. Henry and R. Glosser, "Modulation Spectroscopy as a Tool for Electronic Materials Characterization", J. Electr Materials, 17 (161),1988.

2. F. H. Pollak, and H. Shen, "Photoreflectance Characterization of Semiconductors and Semiconductor Heterostructures", J. Electr Materials, 19 (399),1990.

3. Michael Sydor, J. R. Engholm, M. O. Manasreh, C. E. Stutz, L. Liou, and K. R. Evans, "Photoreflectance and the electric fields in GaAs depletion region" Appl Phys. Lett. 56(1769),1990.

4. D. C. Look, C. E. Stutz, and K. R. Evans, "Unpinning of GaAs Fermi Level by 200°C Molecular Beam Epitaxial Layer", Unpublished.

5. R. N. Bhattacharya, H. Sen, P. Parayanthal, and Fred H. Pollak, "Electroreflectance and Photo-reflectance study of the space charge region in semiconductors: (In-Sn-O)/InP as a model system", Phys. Rev. B37(4044)1988

6. Fred H. Pollak and O. J. Glembocki, "Modulation Spectroscopy of Semiconductor microstrucrures: an Overview", SPIE Proceedings, Vol. 946, P. 2, 1988.

7. D. E. Aspenes, "Third Derivative Modulation Spectroscopy with low-field electroreflentance", Surface Science 37418(1973)

(a) LASER OFF

CONDUCTION BAND

SURFACE STATES

IONIZED DONORS

VALENCE BAND

(b) LASER ON

CB

VB

Figure 1

Fig 2
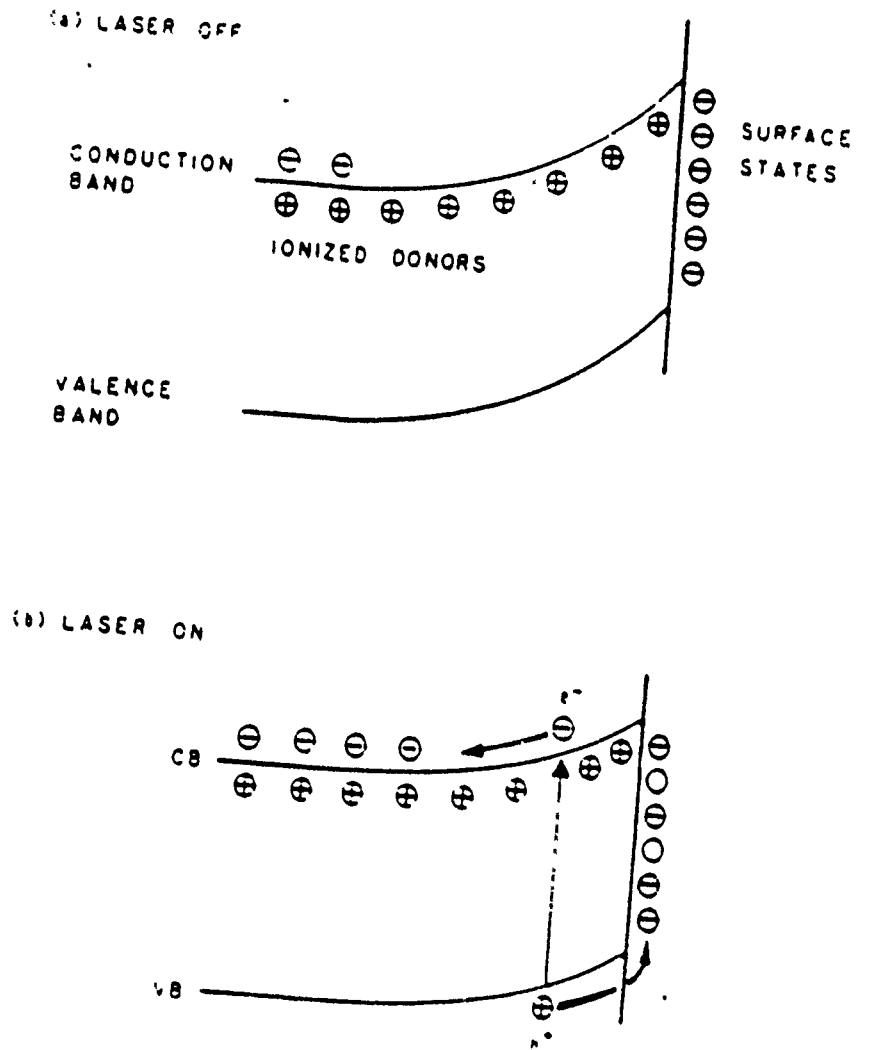
ARB UNITS

G2~1008

InGaAs 90Åx1 well 70A Si in well

nanometers

——— 80K
····· 300K
(×5)

G2—1011

InGaAs (140A×1) 120A Si in Well

ARB UNITS

nanometers

——— 300K   ———— 80K
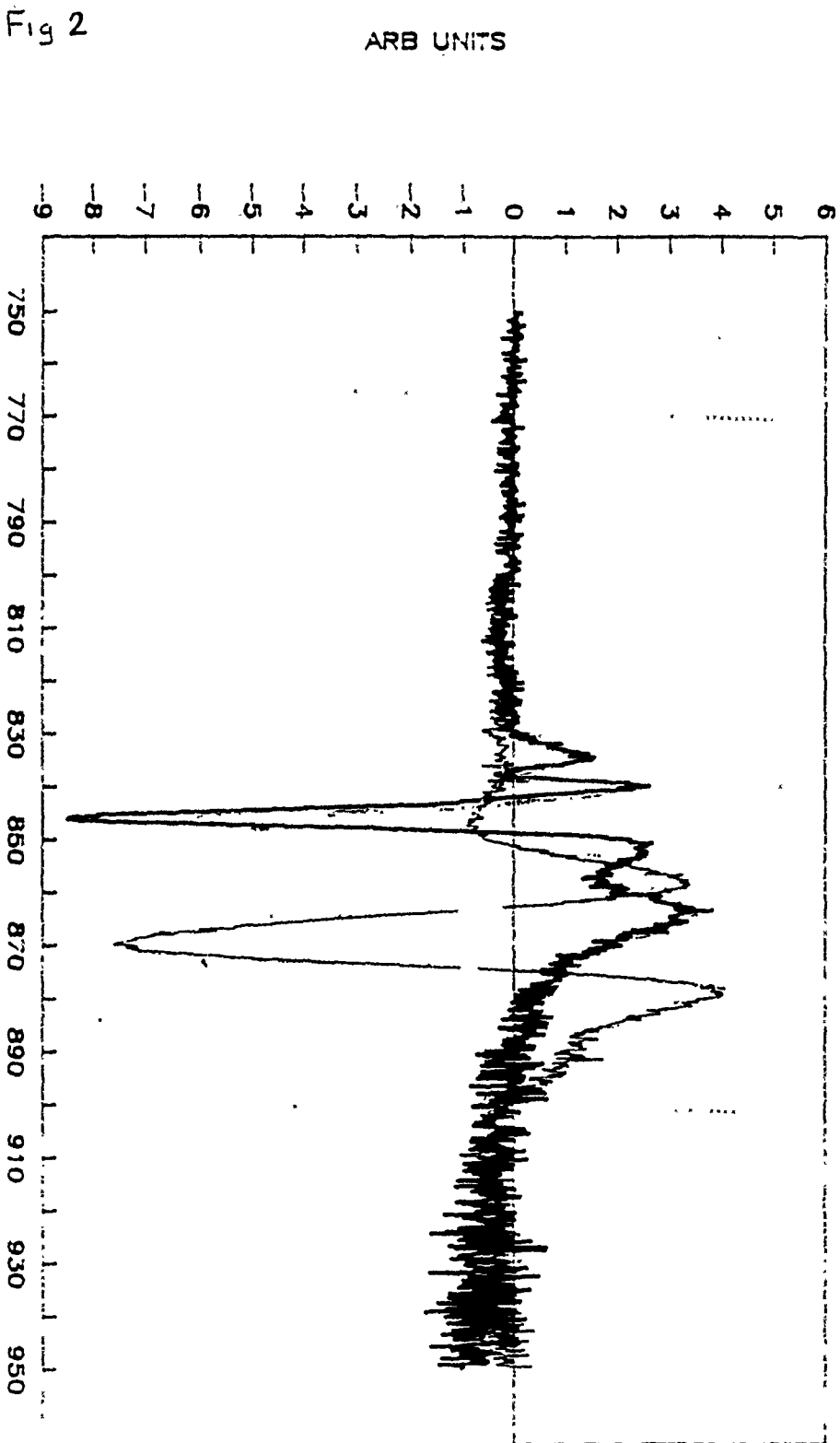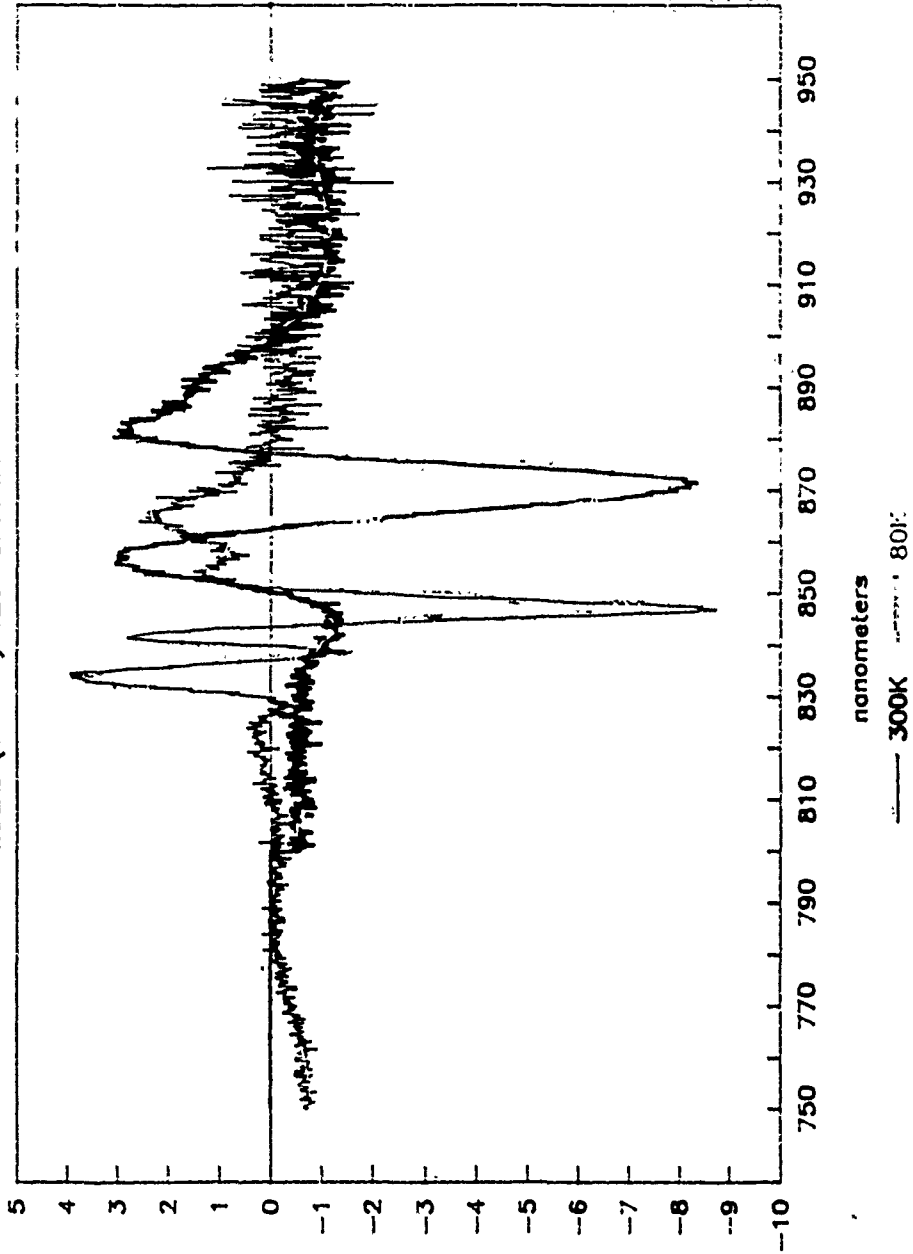
Fig 3

LOW TEMP MBE

P type material
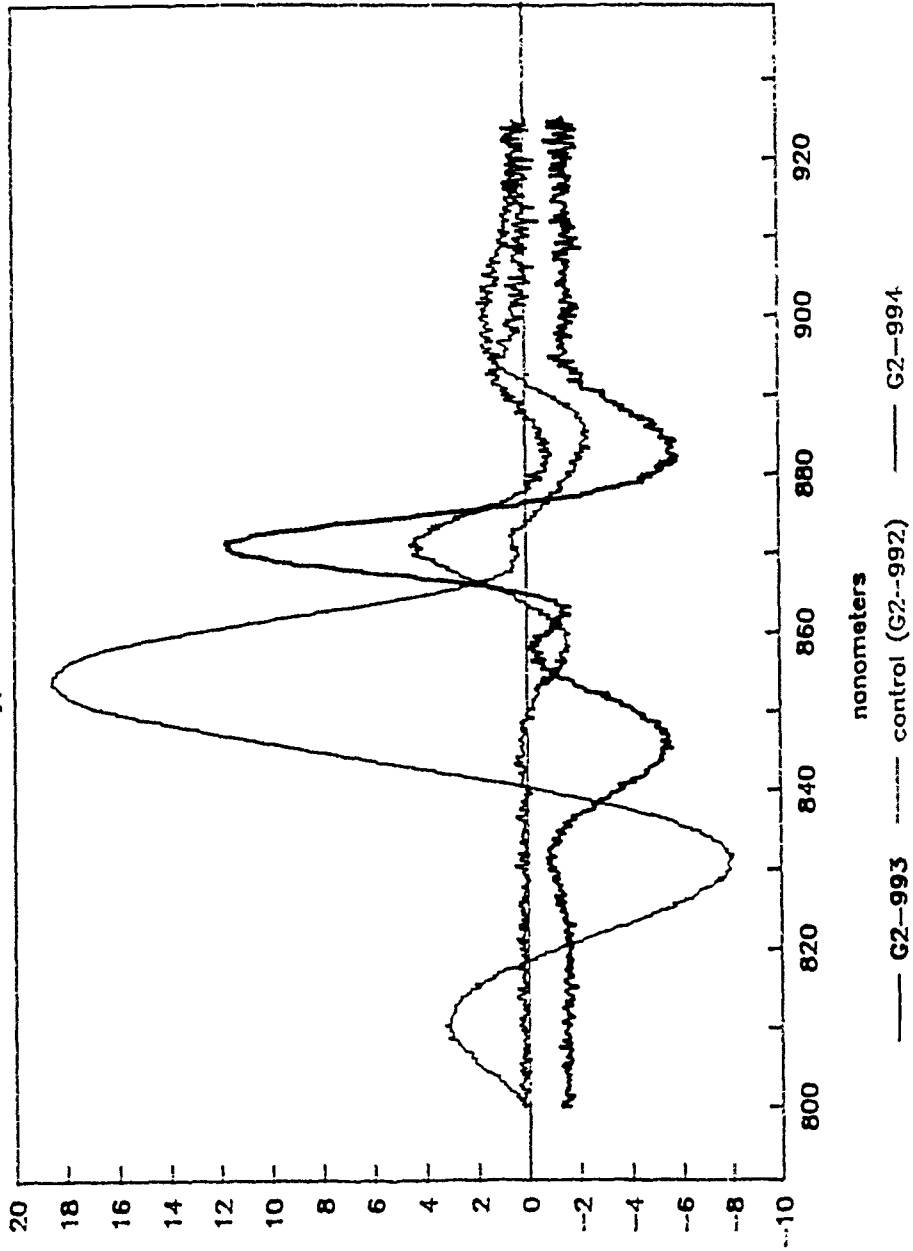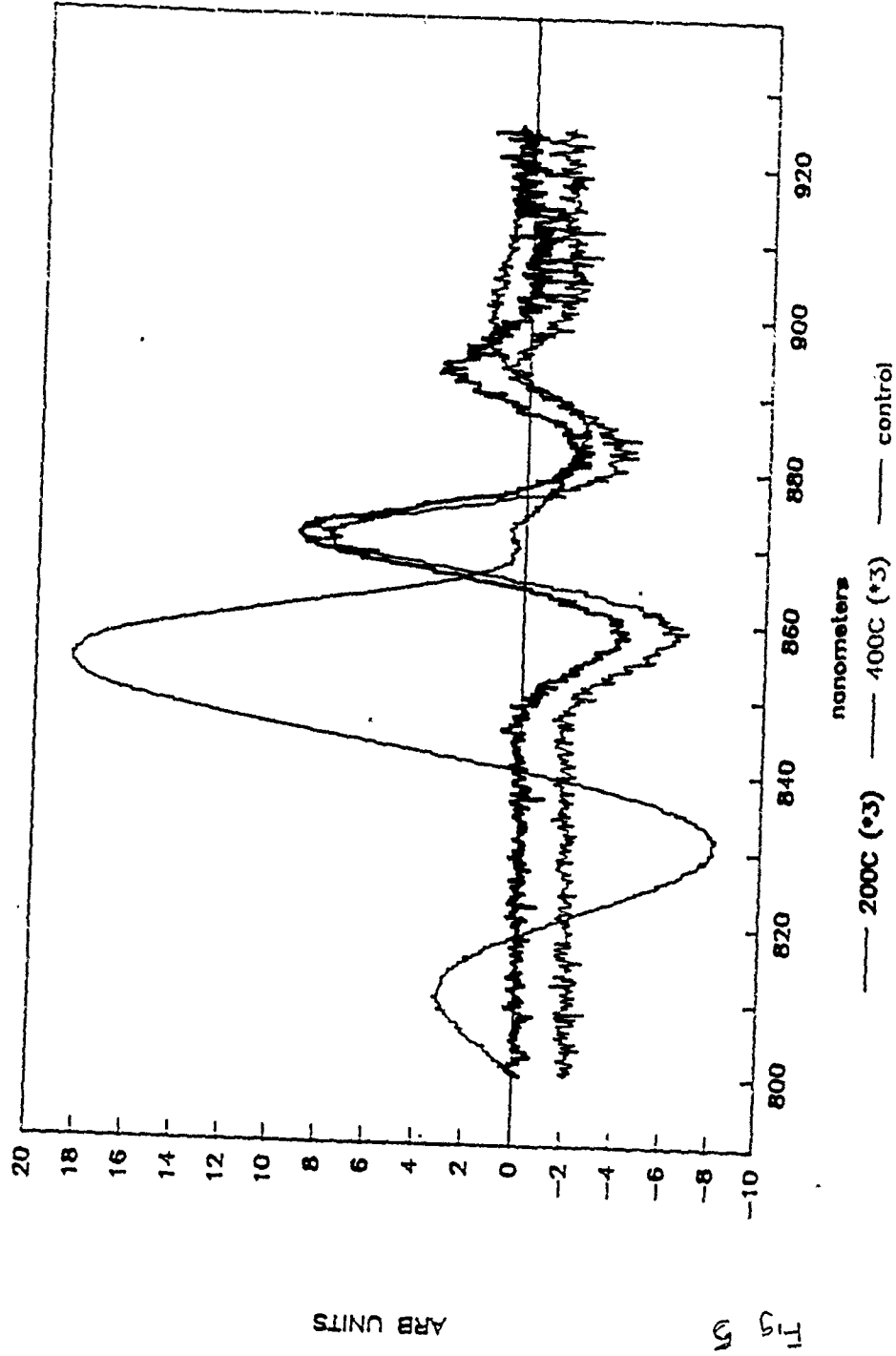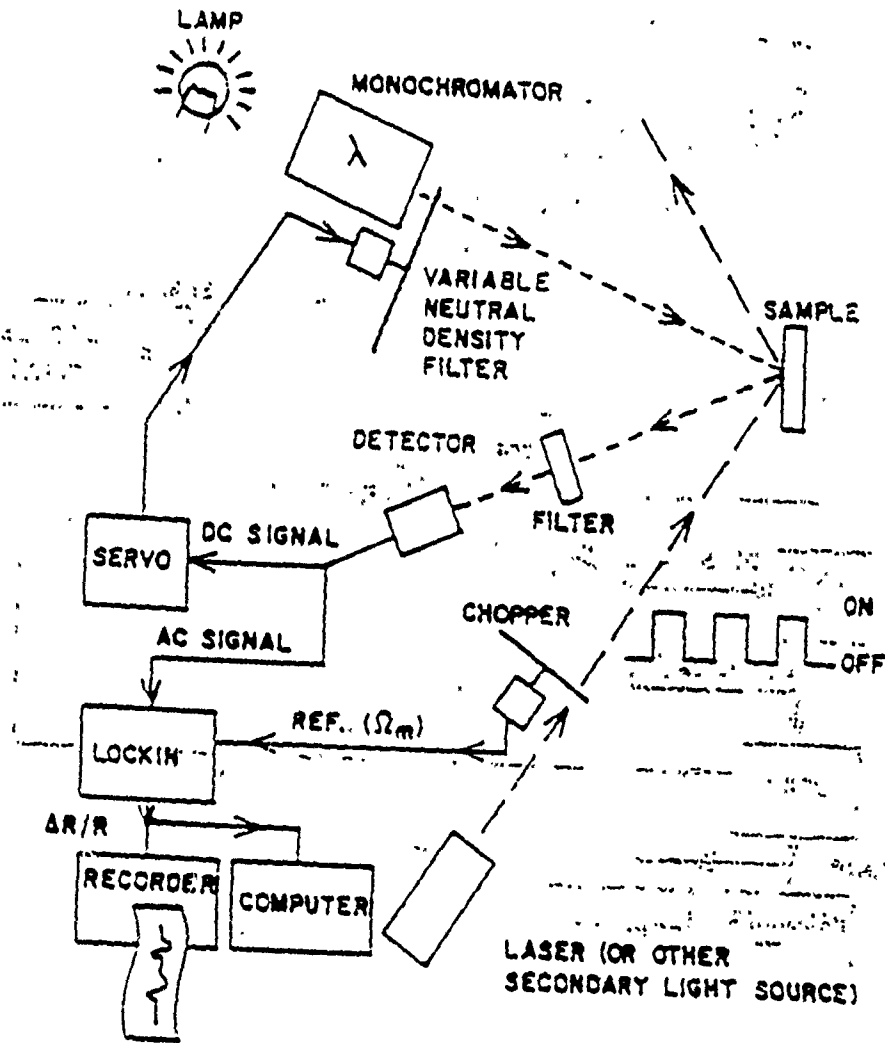
ARB UNITS

nanometers

800 820 840 860 880 900 920

20 18 16 14 12 10 8 6 4 2 0 -2 -4 -6 -8 -10

——— G2-993   ------ control (G2-992)   ——— G2-994

Fig 4

Low temp MBE

ARB UNITS

nanometers

200C (*3) —— 400C (*3) —— control

Fig 5

Fig 6

LAMP

MONOCHROMATOR

λ

VARIABLE
NEUTRAL
DENSITY
FILTER

SAMPLE

DETECTOR

FILTER

SERVO

DC SIGNAL

AC SIGNAL

CHOPPER

ON

OFF

LOCKIN

REF. (Ω_m)

ΔR/R

RECORDER

COMPUTER

LASER (OR OTHER
SECONDARY LIGHT SOURCE)

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Electronic structure and deep impurity levels in GaAs related compound semiconductors and superlattices

Prepared by: Devki N. Talwar and Alan Coleman

Academic Rank: Assistant Professor and Graduate Student

Department and   Department of Physics

University   Indiana University of Pennsylvania

Research Location: Electronic Technology Laboratory

                          ELRA/WRDC, Wright Patterson Air

                          Force Base, Ohio, 45433

USAF Researcher   Dr. M. O. Manasreh

Date: Sept. 27, 1990

Contract No: F49620-88-C-0053

Electronic structure and deep impurity levels in GaAs related compound semiconductors and superlattices

by

Devki N. Talwar
Alan Coleman

ABSTRACT

The band structure of periodic, ultra-thin, lattice matched $(Al_xGa_{1-x}As)_m$ /$(GaAs)_n$, and strained layer $(GaAs)_m$/$(In_xGa_{1-x}As)_n$ superlattices (SL's) grown along the three main crystallographic orientations (001), (110), and (111) is studied by using a second-neighbor tight binding theory. The SL wave functions are described as a linear combinations of bulk Bloch functions (*sixteen*, if spin is included) for each of the *two* constituent materials while the alloy $Al_x Ga_{1-x}As$ (or $In_x Ga_{1-x}As$) is treated in the virtual crystal approximation. To incorporate the effects of strain in strained layer SL's, a new method is developed, based on Harrison's scaling scheme which properly includes the variation of bond lengths and bond angles. While studying the band structure, we found that the band gap in SL's depends not only on the layer thicknesses through quantum mechanical effects but also through the strains in the constituent layers. Our calculated results for the band structure of $(GaAs)_1$/$(AlAs)_1$ and $(GaAs)_2$/$(AlAs)_2$ SL's are found in excellent agreement when compared with the existing sophisticated self-consistent pseudopotential data. Using the above information of band structure, a Green's function theory of impurity levels is being developed. This theory will allow us to predict the role of deep levels in SL's and QW's and may prescribe a method to overcome the effects of deep traps which we believe are responsible for limiting the performance of Si-doped HEMT's, and other technologically important devices.

Acknowledgements

The author wishes to express his sincere thanks to the Air Force System Command (AFSC), and the Air Force Office of Scientific Research (AFOSR) for the award of a "1990 UES Summer Faculty Research Fellowship" that enabled him to work at the Electronic Technology Laboratory, Wright Patterson Air Force, Base, Ohio. Thanks are due to Dr. Rodney C. Darrah, of UES and to Col. Claude Cavender for arranging an excellent slide show which provided the participants an indepth knowledge of the research intersts of the AFOSR.

The ten week stay at WPAFB was short but very exciting, enjoyable, and rewarding. Col. Ken Soda and Dr. Omar Manasreh of the Electronic technology laboratory (ELR/ ' WRDC) provided the author with support, encouragement, and a truly friendly working atmosphere. Fruitful discussions on the subject matter with Drs. Dave Look, Omar Manasreh, Phil Won Yu, D. Reynolds, Keith Evans, Ed Stutz, Dave Fischer, Dennis Whitson, and Cole Litton were invaluable in learning the new and exciting area of semiconducting superlattices.

I. INTRODUCTION

Modern crystal-growth technique such as molecular-beam epitaxy (MBE) has made possible the laboratory synthesis of a variety of ultra-thin layer superstructures. Among others, there is a special group of systems which spontaneously construct the superstructure under certain experimental conditions. One of them is the monolayer based strained GaP/InP structure.[1-6] Other systems include lattice matched GaAs/AlGaAs, and strained GaAs/InGaAs quantum wells (QW's), and superlattices (SL's). The strained GaP/InP structure exhibits an atomic ordering which has been revealed recently by the transmission electron microscopy to be a (111) SL. Strained layer materials exhibit special importance as they permit study of the relation between band structure and strain. Again they retain all the advantages of lattice-matched systems and exhibit additional features, such as the potential for the highest valence states to be light-hole-like, which is of benefits for many tailorable engineering applications.

A major advance for potential high performance devices was made when Strömer, et al.[7] introduced n-type doped modulation doped samples. The underlying idea is that, at equilibrium, charge transfer occurs across a heterojunction to equalize the chemical potential (i.e., the Fermi level) on both sides. If the wide band gap side of AlGaAs/GaAs heterojunction is doped, the electrons will be transfered to the GaAs layer until an equilibrium is reached. This occurs because electron transfer raises the Fermi energy on the GaAs side due to level filling and also raises the electrostatic potential of the interface region because of the more numerous ionized donors in the AlGaAs side. The charge tansfer effect makes possible an old dream of semiconductor technologists, i.e. getting conducting electrons in a high-purity, high-mobility semiconductor without

having to introduce mobility-limiting donor impurities. Since then, modulation doping has been applied to a number of situations involving various semiconductor pairs and also to hole modulation doping. In recent years, the impressive development of the subject has resulted in several high-performance electronic devices including faster complementary logic circuits,[8] efficient and stable long-wavelength lasers,[9] and high electron mobility transistors (HEMT's), etc.[10-12]

II. Need for a reliable theory of band structure in SL's

a. The Present Situation

In device design using heterostructures, it is imperative to know the electronic band structure and the behavior of defects in SL's and QW's from a reliable theoretical model. In bulk III-V semiconductors, the point defects have already played a leading role: they compensate and scatter the free carriers, shorten their life time, and introduce spurious effects. This is the case, for instance, of the famous EL2 defect in GaAs and DX center in $Al_xGa_{1-x}As$.[13] In heterostructures, in addition to point defects located in the layers, there may also be the possibility of the existance of interface defects between layers.[14] Thus, an understanding of the role of defects in heterostructures is atleast as important as in bulk or ternary compound semiconductors. However, very few studies have been performed in order to characterize defects and to predict their behavior in SL's.

b. Objectives

Earlier we developed[15-16] an empirical second-neighbor tight-binding theory and applied it successfully to treat the energy-band structure (See Figs. 1 a-c)and

deep impurity levels (See Figs. 2 a-b) in elemental and III-V compound semiconductors. The strength of our approach was in its chemical trends of the matrix elements of the TB Hamiltonians, since these are the local functions of the bulk semiconductor. The observed trends[15-16] in the short-range interactions are particularly suited for the study of the interface disorder and for the inclusion of strain effects in strained-layer SL's. To incorporate strain in strained-layer SL's we proposed here a new method, based on Harrison's scaling[17] scheme, which includes the variation of the bond lengths and bond angles. With these ideas properly considered, a systematic transition of the TB method from bulk[15-16] to SL's is possible for studying both the band structure and the impurity levels. Such a theory is needed for predicting the conditions under which a specific impurity produces a shallow donor level and "dopes" the semiconductor, versus the possibilities of deep-level formation and the trapping of charge carriers. The present work done under *1990 UES Summer Faculty Research Program (SFRP)* had two major objectives: (i) to develop a second-neighbor Tight-binding theory for the Band structure of SL's, and (ii) to formulate the TB-Green's function theory for deep-levels in SL's

III. Theory

a. Band structure

The method used for the calculations of the band structure of SL is similar to that of Ihm[18] with *two* exceptions: First, formulae for the TB matrix elements have been corrected to yield the same output for the bulk X-point energies as the input data. Second, the interactions up to and including second nearest neighbors

have been introduced to allow an exact description of L-point energies in bulk semiconductors.

Consider an SL of two different alternating zinc-blende crystals labelled 'ab' and 'AB' with (001) interface, where a(A) and b(B) are cation and anion atoms, respectively. The SL $(ab)_n/(AB)_m$ contains $2(n+m)$ atoms in a unit cell at R_i with four (sp^3) orbitals each ; $|\alpha j>$ where α denotes the s, x $(=p_x)$, y $(=p_y)$ and z $(=p_z)$ orbitals, and j represents the site index in a unit cell which runs from 1 through $2(n+m)$. If the Bloch sums defined by

$$| \chi_j^\alpha(\mathbf{k})> = N^{-1/2}\sum_i e^{i\mathbf{k}(R_i+\tau_j)} |\alpha j> \tag{1}$$

are used as basis, the SL Hamiltonian H^0 can be written as

1	2	3	...	...	2n	1	2	3	...	...	2m
a	ab	aa								Aa	Ba
	b	ba									Bb
		a									
			...	...	bb						
					ab	aA					
					b	bA	bB				
						A	AB	AA			
						B	BA				
							A				
									...	...	BB
	h.c.									...	AB
											B

$H^0 =$

$\tag{2}$

where $k = \frac{2\pi}{a_o}(k_1, k_2, k_3)$ (a_o is the lattice constant) is the wave vector, τ_j is the atom position in a unit cell and N is the number of Braivais lattice points. Each element in Eq. (2) represents a [4x4 (or 8x8, if spin is included)] matrix and, for instance, ab (ba) and aa (bb) correspond to nearest neighbor interactions $\{<\chi_j^\alpha|H^o|\chi_{j+1}^\beta>\}$ and second-nearest neighbor interactions $\{<\chi_j^\alpha|H^o|\chi_{j+2}^\beta>\}$ having j of site 'a' ('b'). The diagonal elements as a or b symbolize $\{<\chi_j^\alpha|H^o|\chi_j^\beta>\}$ with j of site 'a' or 'b' and contain both intrasite energies and second neighbor interactions. Elements bA, Ba, aA and bB correspond to intermaterial interactions at interface. Additionally, the Hamiltonian for the bulk crystal ab is expressed as follows:

$$Ho\ (n=1, m=0) = \begin{pmatrix} a + aa + aa^\dagger & ab + ba^\dagger \\ \\ ba + ab^\dagger & c + cc + cc^\dagger \end{pmatrix} \tag{3}$$

which is equivalent to that reported by Talwar and Ting.[15] The TB-parameters obtained earlier[15-16] for the bulk III-V compounds are assumed not to change at the interior layers of the corresponding compound when the constituent materials are combined to form a SL. In addition to the bulk constituents, the interface region must be described properly. For interface region, we will use averages of the bulk parameters for both compounds or the atoms are assigned to either bulk regions. A properly specified valence-band offset, combined with the bulk electronic structure, will determine the rest of the band line-ups.

b. Inclusion of Strain Effects

A new method is suggested here based on Harrison's scaling scheme[17], to incorporate the effects of strain in the band structure calculations of SL's. Unlike

other methods this method does not require experimental data for the shifts or the splittings of energy bands and has the advantage of being applied to any strained system. Two extreme cases are considered here. One of them is the so-called free standing model, which describes the strain distribution when the thickness of the SL exceeds a certain critical value or the SL is lattice matched to the substrate. The other case considered here is the coherent growth model which has been suggested to be true both experimentally and theoretically when the thickness of the SL is under the critical value.

c. TB-Green's Function Theory of Deep-Levels in SL's

Quite recently in random alloys such as $Al_xGa_{1-x}As$, we have proposed[19] the possibility of "shallow-deep" transition for isolated donor defects (DX-center), when x reaches a critical value. It is our belief that the deep levels are much less sensitive to changes of alloy composition x than the band edges. Slightly overstating the point, we may also say that the deep levels are almost constant in absolute energy, and that by varying x one can cause the conduction or valence band edge to move through the deep levels. In $Al_xGa_{1-x}As$, for example, by changing x the band gap varies and it opens up a window of observability of the deep levels (c. f. Fig. 2). Thus by a proper selection of the alloy composition x in mixed semiconductors, a given impurity can produce a shallow level rather than a deep one, is a form of "band-gap engineering". We believe that similar ideas of "band-gap engineering" can be achieved by altering the atomic ordering in the properly selected semiconductors which make the hosts either nearly lattice matched or strained-layer SL's, rather than a random alloy.

IV. Preliminary Results

The tight-binding method outlined above (c.f. Sec. II) offers a useful frame-work for the subband energy calculations in both lattice matched and strained-layer SL's. Here, we present via our preliminary calculations, a picture of the successful transition of the bulk TB-theory to the band structure calculation of GaAs/Al$_x$Ga$_{1-x}$As (001)SL's. By comparison with the existing sophisticated theoretical[20-21] (c.f. Fig. 3) and reliable optical (PL or PLE) data,[22-23] we have obtained an extensive knowledge which supported our TB approach. While the methodology adopted and the group theory used are of general applicability, the choice of one example presented here will leave out many important issues that we are planning to undertake. These issues are:

a. Calculation of Energy Band Gap variation with period of the SL's

This type of study may help us attain new band gap regimes not otherwise achievable with the bulk III-V compound semiconductors. For instance, InAsSb strained layer detectors are believed to operate at wavelengths up to about ~12μm and may possibly be more stable than HgCdTe detectors.

b. Cross over from Type-I to Type-II (GaAs)$_m$/(AlAs)$_m$ SL's

An equally important and still unresolved issue in short-period (GaAs)$_m$/(AlAs)$_m$ SL's is to understand the cross over (c.f. Fig. 4) from type-I to type-II superstructure. Here, m represents the number of monolayers in one SL period (a monolayer of GaAs or AlAs is ~2.83 Å thick). In Fig. (4a), we have displayed both Γ-like quantized states (which are mainly confined in GaAs layers)

and X-like states (which are mainly confined in AlAs layers) in $(GaAs)_m/(AlAs)_n$ SL's. When m is relatively large, the Γ-like energy levels are lower than the X-like levels and the resulting superlattices are of type-I in which both electrons and holes are confined in the same constituent layers (see Fig. 4a). However, when m is small enough, the X-like states will become the lowest states in the conduction band. The superlattices become type-II in which the electrons and holes are confined in the adjacent constituent layers, respectively, (as shown in Fig. (4b)). Obviously, for certain critical values of m the Γ-like levels are nearly equal to the X-like ones $i.e.$ a type-I to type-II transition occurs as shown in Fig. (4c). Experimental evidence for such a crossover is likely to be important[24] for the vertical transport (tunnelling) in SL structures. Although, some PL data is known for the critical thickness (m) for which the transition from type-I to type-II occurs at ambient pressure. However, the existing results are rather sparse primarily because in type-I SLs the X-like states are difficult to detect either by PL or by PLE measurements.[22-23]

c. Relative Position of the X-valleys in Type-II $(GaAs)_m/(AlAs)_m$ SL's

Another important and unsolved problem, from the basic physics standpoint, is to know which X-valley states (in the conduction band), the X_{xy} and or the X_z in type-II short-period $(GaAs)_m/(AlAs)_m$ SL's attain the minimum energy with respect to the valence band maximum. Here, X_z is the X-valley with wave vector (**k**) parallel to the growth axis, while X_x and X_y have **k** in the plane of the layer. From simple Group-theoretic analyses and by considering appropriate Brillouin zones of the SL's, it can be shown that for the (001) SL's, there is a possibility of mixing of Γ-X_z states. On the other hand, for the (011) SL, it is the X_z - X_x states that mix

together and plays a dominant role, while for the (111) SL, it is the mixing of L-valley with the Γ-valley state which is important. In SL's the mixing and splitting of the energy states is manifestly an interface effect which cannot be treated within the Krönig-Penney or **k.p** (EMA) models. This issue, we believe can be resolved only if realistic theories are considered and if the band structure of $(GaAs)_n/(AlAs)_m$ (011) short-period SL's is calculated.

d. Problems Related to the Modulation Doping in AlGaAs/GaAs SL's

Modulation doping in GaAs/AlGaAs SL's, by which donor (say, Si) impurities are inserted into the large gap AlGaAs layers of a SL but donate their electrons to the small band-gap GaAs layers, has already played a role in the development of high speed III-V compound semiconductor devices. However, practical devices based on $Al_xGa_{1-x}As$ are often limited to alloy compositions x < 0.3. This is due to the formation of Si-related defects that are deep traps. Some devices, such as HEMT's operate using QW structures at or near GaAs/AlGaAs interfaces and the performance of these devices depends on the donor-doping, the alloy composition and the SL's structure. Clearly, the tight-binding Green's function theory in SL's and heterostructures will help us learn the electronic characteristics of point defects giving rise to deep-levels.

References

1. Smith, D. L., and C. Mailhiot, *Rev. Mod. Phys.* 1990, Vol. 62, p. 1.

2. O'Reilly, E. P. *Semiconductor Sci. Tech.* 1989, Vol. 4, p. 121.

3. Hayakawa, T. T., K. Suyama, M. Takahashi, M. Kondo, S. Yamamoto and T. Hijikata, *Appl. Phys. Lett.* 1988, Vol. 52, p. 339.

4. Dingle, R., Editor *'Applications of Multi-quantum wells, Selective Doping, and Superlattices'* 1987, (Academic, New York).

5. Osbourn, G. C., R. M. Biefield, and P. L. Gourley, *Appl. Phys. Lett.* 1982, Vol. 41, p. 172.

6. Fritz, I. J., L. R. Dawson, and T. E. Zipperian, *Appl. Phys. Lett.* 1983, Vol. 43, p. 846.

7. Störmer, H. L., R. Dingle, A. C. Gosard, W. Wiegmann, and R. A. Logan, *Conf. Ser. - Inst. Phys.* 1979, Vol. 43, p. 557.

8. Lepore, A. N., M. Levy, H. Lee, and E. Kohn, *Electron Lett.* 1988, Vol. 24, p. 364.

9. Henderson, T. et al., *IEEE Electron Dev. Lett.* 1986, Vol. EDL7, p. 649.

10. Mizuta, H. et al., *IEEE Trans. Electron Dev.* 1989, Vol. 36, p. 2307.

11. Watanabe, Y. et. al., *IEEE GaAs IC Symp. Dig. Tech. Papers* 1988, p. 86.

12. Notomi, S. et. al., *IEEE GaAs IC Symp. Dig. Tech. Papers* 1987, p. 177.

13. Henning, J. C. M., and J. P. M. Ansems, *Semicond. Sci. Tech.* 1987, Vol. 2, p. 1 .

14. Bourgoin, J. C., and H. J. von Bardeleben (private communication)

15. Talwar, D. N., and C. S. Ting, *Phys. Rev.* 1982, Vol. B25, p. 2660.

16. Talwar, D. N., K. S. Suh, and C. S. Ting, *Phil. Mag.* 1986, Vol. B54, p. 93 : K. S. Suh, *Ph.D. Thesis* (University of Houston, 1988), (unpublished).

17. Harrison, W. A. in *Electronic Structure and the Properties of Solids,*
1980 (Freeman, San Francisco).

18. Ihm, J., *Appl. Phys. Lett.,* 1987, Vol. 50, p. 1068.

19. Talwar, D. N., M. O. Manasreh, K. S. Suh and B. C. Covington, *Appl. Phys. Lett.* 1987, Vol. 51, p. 1358.

20. Gilbert, T. G., and S. J. Gurman, *Superlattices and Microstructures,*1987, Vol.3, p. 17.

21. Srivastava, G. P. in *Gallium Arsenide and Related Compounds 1987* (Inst. Phys. Conf. Ser. 91) p. 529.

22. Jiang, D. S., K. Kelting, T. Isu, H. J. Queisser and K. Ploog, *J. Appl. Phys.* 1988, Vol. 63, p. 845 .

23. Nakazawa, T., H. Fujimoto, K. Imanishi, K. Taniguchi, C. Hamaguchi, S. Hiyamizu, and S. Sasa, *J. Phys. Soc. Jpn.* 1989, Vol. 58, p. 2192.

24. C. Minot, H. Le Person, J. F. Palmier and R. Planel, *Superlattices and Microstructures* 1989, Vol. 6, p. 309.

Figure captions

Fig. 1a. Calculated band structure of GaAs based on the second-neighbor tight-binding parameters of Talwar and Ting (Ref. [29]).

Fig. 1b. Same key as of 1a but for AlAs.

Fig 1c. Calculated band structure of $Al_{0.43}Ga_{0.57}As$ in the virtual-crystal approximation.

Fig. 2a. Schematic illustration of the dependence of deep impurity levels, caused by donor defects (DX-centers) occupying cation sites, on x in AlxGa1-xAs (after Ref. [19]). The zero of energy is the valence band maximum of the alloy.

Fig. 2b. Same key as of 2a but for donor defects on anion sites.

Fig. 3. Calculated electronic band structure of (001) oriented $(GaAs)_1/(AlAs)_1$ superlattice based on our second-neighbor tight-binding theory described in the text. The results are found in excellent agreement with the sophisticated self consistent pseudopotential calculation.

Fig. 4 Schematic illustration of the energy band structure for the $(GaAs)_m/(AlAs)_m$ SL's of the E^Γ transition (from the Γ-like state to the heavy hole state) and the E^X transition (from the X-like state to the heavy hole state). The dashed lines represent the X-valley energy of GaAs and AlAs bulk materials. The bold lines represent the lowest quantized energy states in the wells. (a) The type I SL. (b) The type II SL. (c) The SL at the transition point from the type-I, to type- II SL. (a') Case (a) under hydrostatic pressure, having been the type II SL.

(1a)

(1b)

(1c)

Fig. 1

Fig. 2

Fig. 3

GaAs/AlAs SLS (001)

ENERGY (EV)

WAVE VECTOR
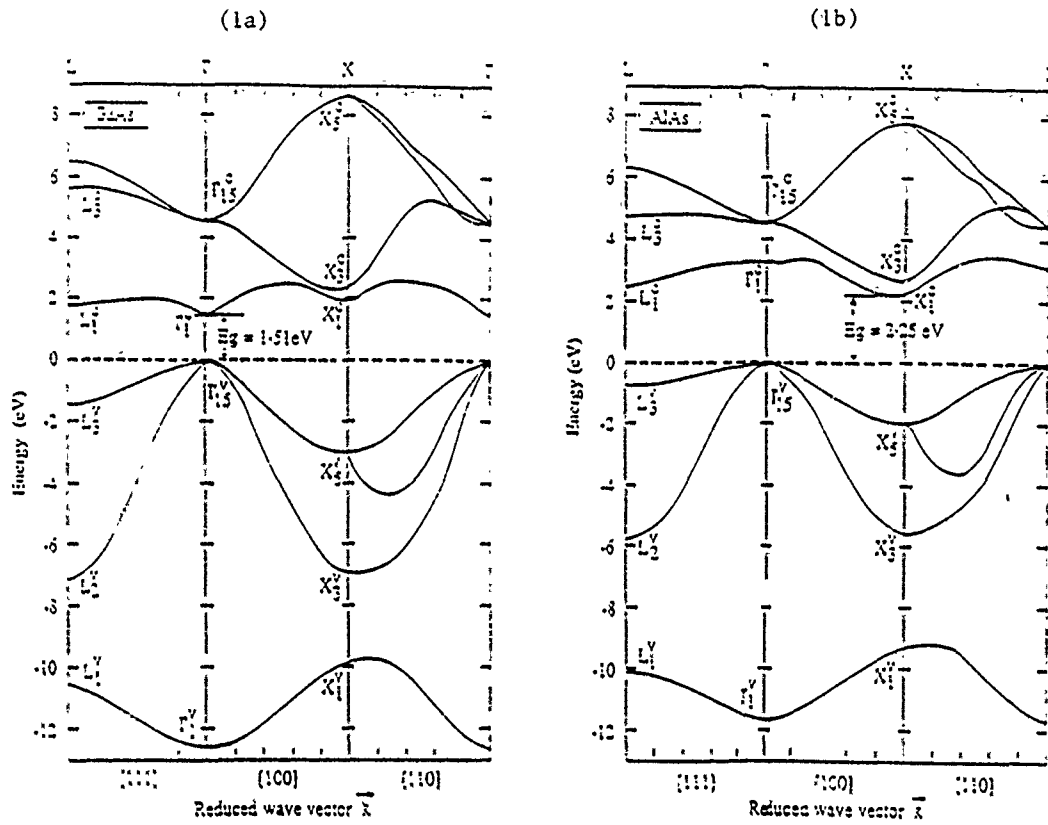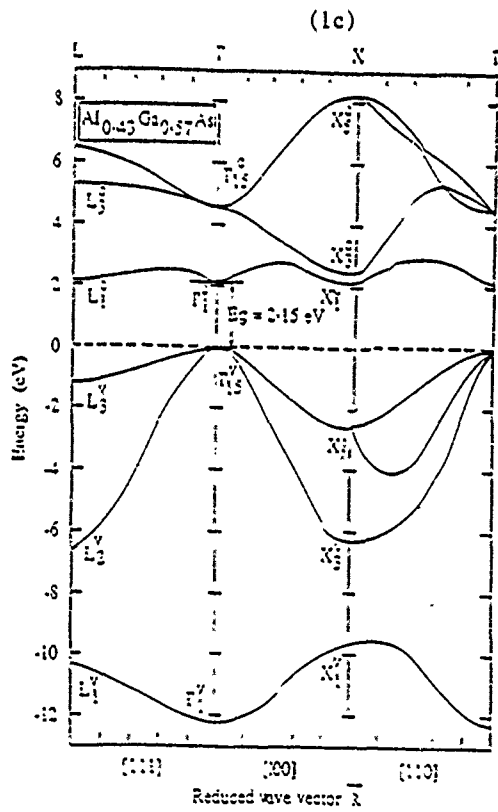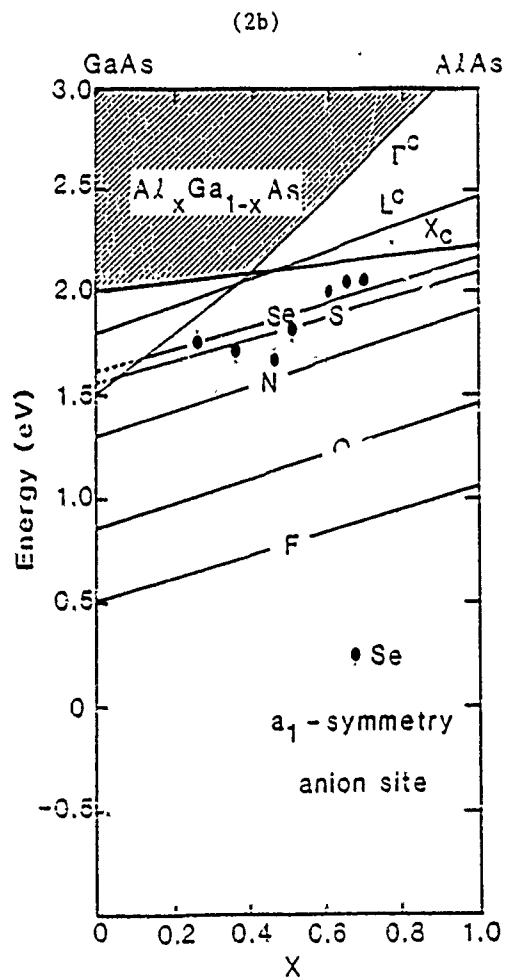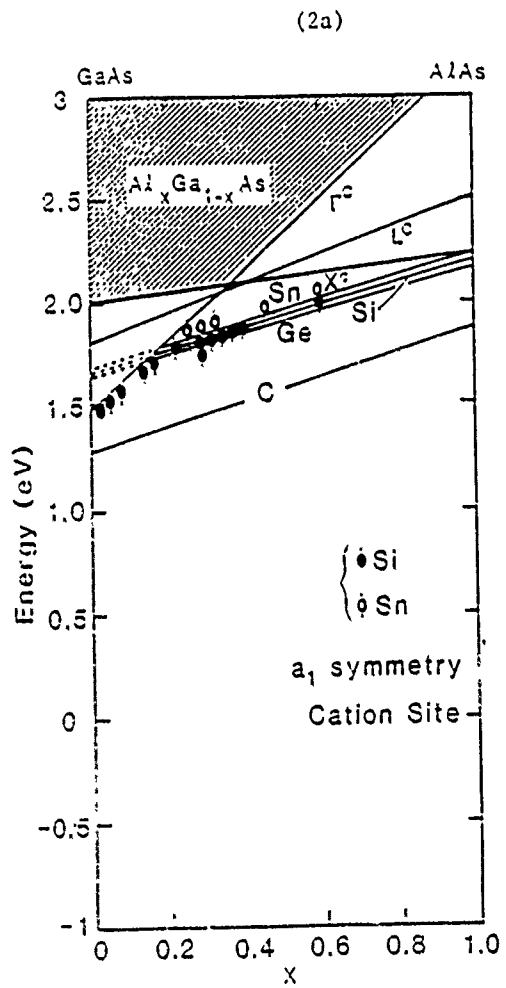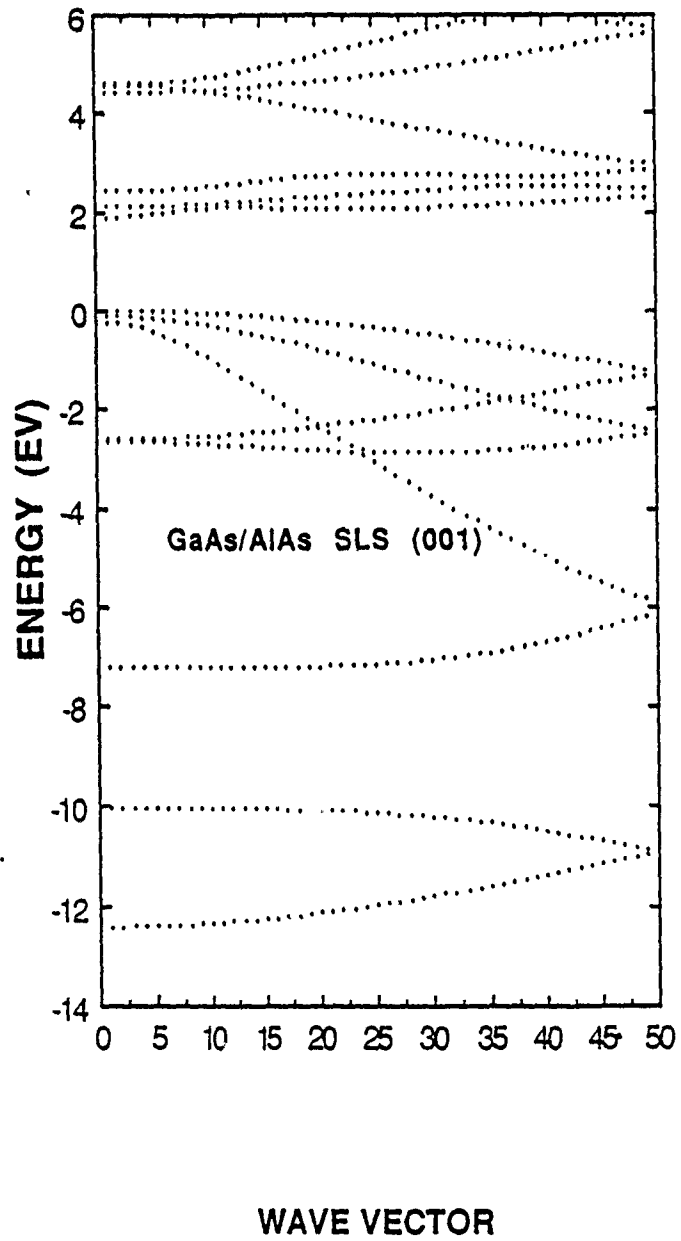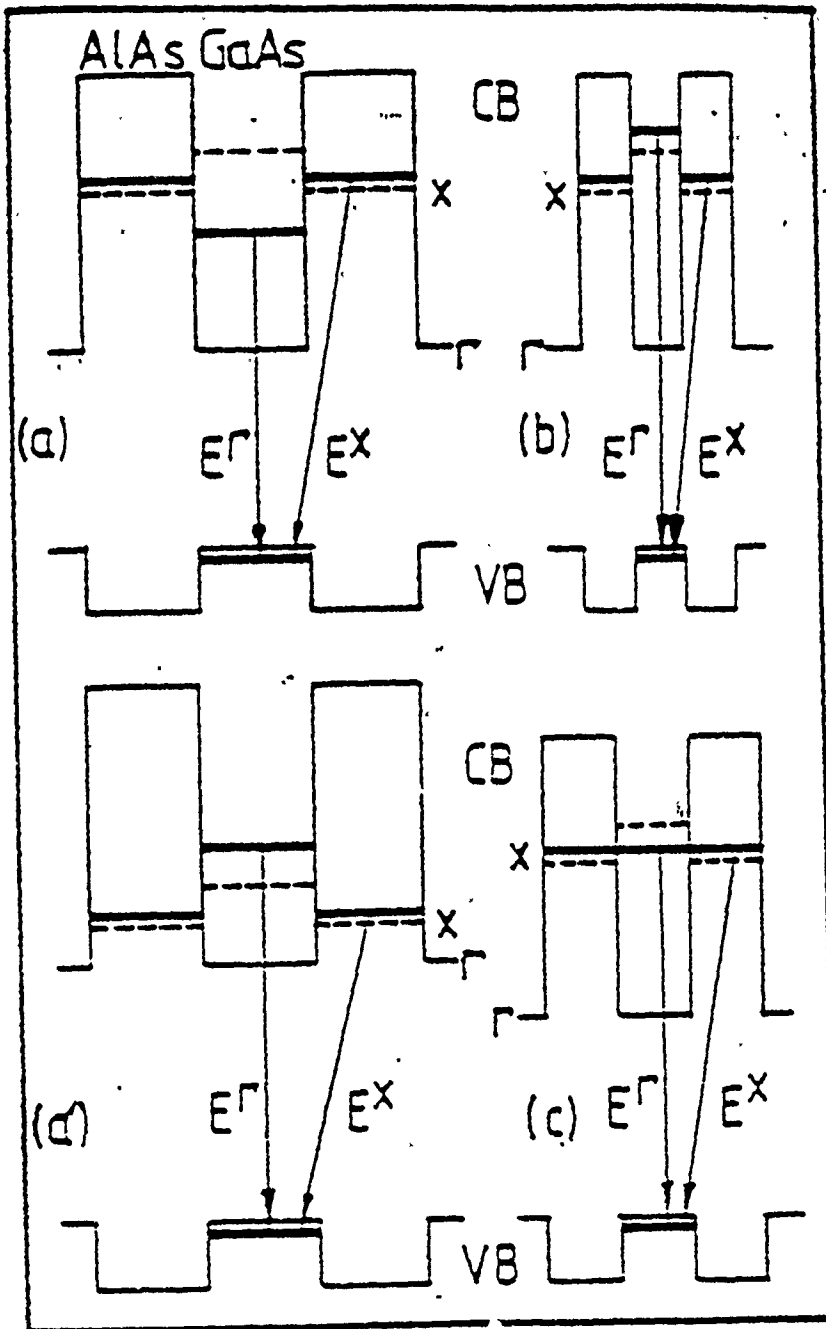
Fig. 4

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Sensor Integration Issues in Robotic Rapid Aircraft Turnaround

Prepared by: John S. Bay, Ph.D.

Academic Rank: Assistant Professor

Department and University: Bradley Department of Electrical Engineering

Virginia Polytechnic Institute and State University

Research Location: WRDC/FIVMB

WPAFB, OH 45433

USAF Researcher: Mangal D. Chawla, Ph.D.

Date: August 1, 1990

Contract No.: F49620-88-C-0053

Sensor Integration Issues in Robotic Rapid Aircraft Turnaround

by

John S. Bay

ABSTRACT

Concepts for robotic rapid aircraft turnaround are examined for the anticipated sensor specification and data processing requirements. Of particular concern are the inspection, monitoring, and supervision of robotic operations in an integrated combat turn where there is the possibility of chemical and biological hazards. Because of the harsh environment expected, it is suggested that the robotic refueling operations exploit infrared sensing systems, including thermal imaging and range sensing, as opposed to ultrasonic measurement techniques. Parts mating operations should also incorporate active and passive compliance devices, as well as force and torque sensing. For some operations, such as ordnance loading, it is suggested that force-amplifying telemanipulators be considered over full autonomy. Also, some sensing apparatus should not be mounted to the robotic components themselves, but are more appropriately fixed in space. These fixed sensors should include vision and floor-mounted load cells. For laboratory feasibility studies and demonstration, tactile sensor research is suggested, along with a data processing technique based on distributed computation of the extended Kalman filter.

I. INTRODUCTION

A. The Problem

During wartime, a critical consideration in the generation of sorties for tactical fighter aircraft is the time required for ground servicing. An integrated combat turnaround (ICT) is the procedure used by the U. S. Air Force for maximizing this aircraft turnaround rate. During this procedure, the aircraft must be marshalled to the servicing area, the wheels must be chocked, the aircraft must be inspected for any damage, and electrically grounded and bonded to the refueling vehicle. Following a well specified set of rules, the aircraft is then refueled and re-armed with missiles and bombs before it returned to the runway for the next mission.

Of additional concern in the operation is the possibility of chemical and biological (CB) warfare hazards on the flight line. In such a situation, all equipment and manpower must be protected from these CB agents, and must be decontaminated subsequent to exposure. Unfortunately, the protection gear worn by the ground crew consists of several layers of bulky material which severely restricts their dexterity and endurance. Ideally, the turnaround is targeted for 45 min. servicing time, but the crew members can suffer heat stress during this time in the protective suits. Furthermore, especially in missile loading, arming, and validation, intricate close tolerance dexterous maneuvers are required. These procedures are exceedingly difficult with multi-layer gloves.

In 1988, Battelle Memorial Institute in Columbus, Ohio, completed an exhaustive report on the possibility of using robotic manipulators for performing some or all of the ICT tasks [Smith]. The report recommended near- and far-term scenarios for using robotics to reduce crew exposure to the CB agents, and to expedite the automated process by concurrent execution of the tasks involved.

Of particular importance to this work, the report recommended an autonomous boom to refuel the aircraft, taking advantage of the increased flow rate of the aircraft's aerial refueling port over the conventional ground refueling port. Refueling would otherwise follow Air Force guidelines for "hot-pit" refueling (engines running). Automated ordnance loading would be facilitated by smart robotic "jammers;" vehicles which could be operated by remote control pendants, and which could position the missile or bomb in its correct position for mounting and arming. To facilitate the rest of the ordnance loading procedure, it was recommended that the aircraft and missile be re-engineered with "quick-connect" connections, reducing the manual dexterity required.

B. Summer Faculty Research Program

In the past, my own work has focused on the interpretation of noisy sensor data in the intelligent perception of complex objects by robotic manipulators. I am also in the process of

investigating the properties and availability of various sensors in order to outfit my own robotics laboratory at Virginia Polytechnic Institute and State University. It is therefore appropriate that I assist the Flight Dynamics Laboratory in the sensor selection, signal processing, and interpretation of sensor data in the context of robotic solutions to the aircraft turnaround problem. The particular tasks are described below.

II. OBJECTIVES

Assuming that the near-term robotic turnaround functions recommended by the Battelle report are to be implemented, it is the purpose of this effort to examine technical issues in the feasibility of robotic implementation. In particular:

1) Given the current technological state of robotics techniques and the available commercial robotics and automation hardware, recommendations will be made as to the practicality of some of the planned robotic turnaround functions.

2) In the laboratory, a concept demonstration cell has been set up for illustration of the fundamental aspects of robotic turnaround. Based on the evaluation of the practicality of the final implementation, long and short range equipment acquisitions will be analyzed and recommended.

3) With special attention to the particular difficulties inherent in noisy, outdoor/hangar environments, sensor selection and signal processing techniques will be suggested.

4) For the sensor-intensive application envisioned, it is likely that multi-sensor fusion techniques will be required for practical implementation. These techniques will be examined from the current literature, and suggested methods will be chosen, given the sensor assortment previously selected.

5) In conjunction with the multi-sensor pick-up and processing problem above, issues in robotic perception of geometry and automated exploration will be examined for possible use in the robotic aircraft turnaround functions.

The remainder of this report will present a comparison of possibly applicable sensor technologies. Special attention is paid to features relevant to the rapid aircraft turnaround problem, both in the lab and in the field. Due to space limitations, detailed analysis of processing techniques and some other technical data is given as reference to available literature only.

III. SENSOR TECHNOLOGY

A. Ultrasonic Sensors

In situations where lighting conditions are unfavorable for visual inspection, some other sensing modality is necessary. One popular type of non-contact sensor for such uses is the ultrasonic ranger. Ultrasonic ranging systems can give very accurate measurements with relatively compact and inexpensive equipment. Drawbacks include poor directionality and noise immunity.

Several companies market ultrasonic sensing equipment, but the basic hardware configuration is similar for all. In fact, many different sensing packages are based on the same Polaroid electrostatic transducer. The Polariod transducer is light and portable, consisting of a thin (gold) metallic sheet mounted in a disc-shaped plastic housing, .328 inches thick and 1.69 inches in diameter. The necessary driving and signal processing electronics as well as the portable battery packs are separate and attached by wires.

The sensor operates by emitting a "chirp," lasting approx. 1 microsecond (ms), containing 56 pulses at 300 volts. The pulses are at frequencies ranging from 49.1 to 60 kHz. (This is to avoid possible cancellation of certain frequencies by surfaces with special topographical characteristics.) The transducer then acts as a microphone, picking up the echo from its own chirp. The elapsed time between emission and detection is used with the speed of sound to compute the length of the echo's path. Nominal operating conditions in the Polaroid specification sheet are: temperature, 32°-140°F and relative humidity, 5%-95%. The Polaroid ultrasonic sensor experimenter's kit includes circuitry which enables the unit to measure distances in the range 0.9 to 35 feet.

Ultrasonic ranging systems are relatively simple and inexpensive ($200 - $1000) to implement. The obvious application for robotic environments is in object detection for automatic collision avoidance. This is particularly true for mobile robots and autonomous guided vehicles (AGV's), in which case the sensors operate as a "seeing eye dog" for the robot. Difficulties arise in the implementation on articulated robots due to the presence of the arms themselves. Because of the wide field of view of the sensors, parts of one arm might be identified as an obstacle by a sensor mounted on another arm. Furthermore, if multiple sensors are used, the signals transmitted from one sensor might be erroneously received and interpretted by a different sensor.

In the rapid aircraft turnaround operations, such detectors might be employed to detect obstacles, both static and dynamic. For example, a static obstacle is presented by the wing itself when the required operation is fuel vent verification, since the current scenario has an overhead robot arm reaching under the wing to check vent flow. Dynamic obstacles are presented by moving objects, and might include humans, vehicles, and the robot arm itself. An ultrasonic ranger on each arm segment could detect imminent collision and activate obstacle avoidance algorithms.

1. IMPLEMENTATION CONSIDERATIONS - As shown in [Polaroid], the beam shape of the sonic chirp has a primary lobe approximately 20° wide. Detection circuitry can be readily designed so that objects of width only .007″ [Contaq] can be sensed. Therefore, small objects can be detected in a relatively wide field, but the objects cannot be further localized. The standard configuration is therefore most useful for ranging flat objects in a sparse field. Cluttered environments are not easily mapped, since the signal received indicates only the distance to the nearest object. This situation can be improved upon with the use of ultrasonic horns, which effectively narrow the beam width to improve directionality (see datasheets from Massa Products Corp., Hingham, MA 02043). These structures are available commercially and effectively narrow the beam to less than 10°. A similar effect can be produced with the use of higher pulse frequencies, which disperse less than the approximately 50 kHz pulses normally used. Some improved imaging can also be achieved by using two separate transducers or with scanning sonar and extensive signal processing and filtering [Beckerman], but this process might be too slow for real-time robotics operations.

The operating temperature range mentioned above is to assure nominal accuracy of approximately 1% for the detection circuitry. In reality, the useful operating range is somewhat larger, with some degradation in accuracy. Application notes from the Contaq Technologies Corporation (Bristol, VT) also mention the possibility of error due to object temperature. The problem is actually a result of air turbulence due to thermal currents near the surface of the object. A surface which is correctly ranged when cold might go undetected when hot, since the reflected beam is dispersed and refracted by the turbulent air. Of course, in a flight-line situation, there may be considerable air currents due to nearby jet engines, and these currents may affect the accuracy of the reading.

Another obvious possible source of error in the flight-line environment is the jet engine noise itself. While the noise developed from the jet engine depends on a great many factors, including nozzle diameter, ambient temperature, nozzle air flow velocity, and angle from the jet axis, the sound power levels can easily exceed 140 dB near a turbojet. Peak power occurs in the 150-1200 Hz band, with considerable roll off past 10 kHz [Beranek]. Most commercial noise meters have no appreciable frequency response above 10 kHz, especially so at 50 kHz, where the ultrasonic pulses are. In the ideal case, the ultrasonic pulse and the jet noise are sufficiently separated in frequency that interference would be unexpected, or at least easily filtered. However, the power level of the hot-pit servicing environment is so high that the physical characteristics of the transducer itself may be altered, changing the frequency response even at high frequencies through a nonlinear effect. As mentioned, the transducer is composed of a thin gold film, and is not very robust to the mechanical shock of jet blast.

The sonic reflectance properties of a surface may also affect sensor readings. Flat, smooth surfaces should be oriented perpendicular to the beam path for best readings. The Contaq notes suggest 10° or less deviation from this perpendicular, but note that sufficiently rough surfaces should reflect enough of the beam in a wide path that detection is possible and ranging is accurate.

2. EXPERIMENT - On Friday, June 30, 1990, the polaroid ultrasonic experimenter's kit was tested on a USAF C-130. The flight-line visit was ostensibly to observe cargo loading procedures, and was not an ICT environment, nor were there jet engines in operation in the vicinity. However, the auxiliary power unit (APU) on the aircraft was running and produced high sound levels and hot exhaust. The purpose of the ad hoc experiment was to test the accuracy of the sensor in the vicinity of the power unit. The unit was held by hand and pointed directly at the skin of the aircraft, behind the APU on the left side of the fuselage.

In the range of one to five feet, the sensor performed well, giving stable readings of the distance to the skin of the aircraft. Between six and ten feet, the sensor gave unpredictable results, showing stability for brief periods only (less than 10 seconds). It is in this range that the exhaust from the APU began blowing eratically into the path of the sound beam. After 10 feet, the sensor did not give stable readings of the range, despite the 35 ft. capability of the sensor and circuitry. In this range, the sensor's beam path was persistently interrupted by the hot exaust. With the APU's (and hence, the exhaust) off, the sensor had no discernible difficulty reading ranges of greater than six feet.

B. Laser Range Finders

For remote sensing purposes similar to those for which the ultrasonic range sensors were considered (obstacle detection, etc.), optical range sensors are also available. Most of the applicable optical range sensors use LED lasers (often near infrared: invisible to the human eye) and triangulation techniques to determine the distance to an object. Like ultrasonic rangers, these devices are compact (commonly smaller than a pack of cigarettes, sometimes as small as a matchbox), and have excellent resolution (10 - 50 μm).

Unlike ultrasonic sensors, though, they are highly directional. They operate by shining a laser spot on an object and measuring the location of the spot with a light-sensitive array. They require mounting at a nominal distance from the object, called a "stand-off," ranging from 1 in. to over 3 ft. Around this distance, displacements of ± 1 to 30 in. can be measured (approximately, depending on sensor model and configuration/electronics). The primary advantages of these sensors are i) they are immune to noise, air currents, and temperature variations; ii) multiple sensors laser range finders do not interfere with each other's operations, as ultrasonic sensors will; iii) they are

comparable in price to ultrasonic sensors ($300 - $600); and iv) they are resonably robust mechanically, being sealed packages which can be made resistant to CB agents. The disadvantages are: i) they require opaque surfaces which are not excessively shiny; ii) they require approximately the same ambient temperature and humidity conditions as the ultrasonic sensors, which precludes rainy days; and iii) they are technologically more sophisticated and complex than ultrasonic sensors, although they are still self-contained and are provided with the interface electronics.

C. Microwaves

A new technology which combines the desireable features of both the ultrasonic sensors and the optical rangers is the microwave ranger. These devices operate on the same basic principal as the ultrasonic sensors, but use high frequency (>1 GHz) electromagnetic waves rather than sonic waves. This technique provides better resolution and immunity to environmental factors than the ultrasonic devices, and have greater range than optical sensors. A factor to be investigated in the use of microwaves, though, is possible interference with aircraft and/or flightline electronics and communications. The devices are commercially available (General Microwave Corporation, Amityville, NY) and warrant further investigation.

D. Vision and Infrared Imaging

Conventional imaging systems using vidicon technology transform visible light reflected from a three-dimensional scene to the two-dimensional plane image. The obvious utility of this sensing modality is the remote location of parts and features within the field of view. A properly calibrated camera can be useful in many of the aircraft turnaround tasks.

Primarily, the imaging system used in the current demonstration facility and in the near-term automated scenario uses visual images for two purposes: to inspect fittings and surfaces for defects, wear, or damage; and to fix positions of workpieces such as the aerial refueling port, the grounding receptacles, the nose gear, the robot end effector itself, and visible obstacles. With overhead mounted cameras and vision processing software, the incoming aircraft can be identified and inspected before the refueling ports and other necessary workpoints are identified in three-space.

In the demonstration facility, a charge-coupled device (CCD) camera is also mounted to the last link of the robot, so that the local scene is always accessible. While vidicons are electron tubes, fundamentally cathode ray screens operated backwards, CCD's are solid-state devices fabricated on a semiconductor chip. This facilitates miniaturization and remote operation. Furthermore, CCD cameras have high resolution and low power consumption, making them ideal for robotic applications. where they have gained wide acceptance.

Whereas visible light occupies the wavelengths 0.4 - 0.75 μm in the electromagnetic spectrum (approximately the same band available from vidicon cameras), CCD cameras have useful sensitivity from 0.4 - 1.1 μm. These higher wavelengths extend into the "near" infrared region, suggesting that electromagnetic emissions from hot objects might be visible with CCD cameras. Indeed, CCD cameras are useful in detecting hot objects whose temperatures are $\geq$ 600°C. There are, however, many tasks in the turnaround operation for which temperature sensing might be useful, especially if the range of temperature resolution were extended to include objects cooler than 600°C. Such sensing systems are available in the form of thermal imagers.

Thermal imaging systems are common in the range 2.0 - 6.0 μm. These wavelengths correspond, roughly, to the range of wavelengths emitted by common objects (not necessarily ideal black-bodies) with temperatures of -20° to 2000°C (-4° to 3632°F). A typical system from Hughes Electronics Company, the Systems Series 2000, operates at 2.0 - 5.6 μm, can focus in the range 10 in. to infinity, has a 15° × 10° field of view, and plots temperature in a 16 color coding scheme. It can resolve temperatures within the range -20° to 2000°C within ± 0.5°C. Feature and options allow frame grabbing, isotherm viewing and "thermal filtering" (selecting only objects of particular temperatures for viewing), and magnetic storage on videotapes or hard disk recorders. It consists of an indium-antimonide detector housed in an argon gas coolant.

The possible uses of such a device are numerous:

- Initial inspection of the aircraft includes a check for brake temperature. If the brakes can melt a crayon designed to melt at 371°C (700°F), hot pit refueling is unsafe because of the risk of spilled fuel ignition. If they are over 649°C (1200°F), they are at risk of igniting the hydraulic fluid, and present imminent danger of brake fire. A thermal imager could detect brake temperature before any turnaround operations are performed.

- Throughout the refueling process, constant surveillance must be maintained on the fuel tank vent to ensure that air is properly venting, indicating proper fuel delivery rate. For fuel which has been stored in underground tanks. there is likely to be a temperature gradient between the venting vapors and the surrounding air. A gradient would also result from evaporation of the vapors, which will also be the case for liquid fuel spills onto the tarmac. Such a gradient would surely be within the 0.5° resolving capability of the sensors.

- A well established procedure among chemists is the detection of chemical element through their electromagnetic emission and absorption properties. It was reported in [Hudson] that this technique has been used for the detection of fuel leaks as well as the presence of poison

gases. For example, it was found that hydrogen gas absorbs radiation at a wavelength of 3.4 μm. Therefore, if spectral analysis of the scene is suspiciously devoid of this wavelength, a hydrocarbon fuel leak might be suspected. The Hughes system demonstrated in the laboratory on July 27, 1990 did not have this capability.

- Aside from detecting materials with temperature in the extreme ranges, the imaging system can also "see" objects with "normal" temperatures, so that some of the more conventional visual imaging tasks might be accomplished with a single imaging system; for example, coordinate fixation of the aircraft and localization of the refueling port. As stated in [Lloyd], "Thermal imagers are superior in performance to other types of passive-sensing electro-optical imaging devices when operability at any time of the day or night and under all weather conditions is the primary consideration." A problem always associated with visible wavelength imaging in military environments is the necessity to detect camouflaged objects. Since it has already been determined in the laboratory that a camouflaged aircraft is difficult to detect on a matching runway, and since a metallic plane returning from a recent sortie is certain to have a different temperature than the service area floor or ground surface, thermal imaging might be a valuable option.

Although the thermal imaging systems may have considerable advantages over the CCD or vidicon cameras, they also have their drawbacks. Because the sensors are temperature-sensitive, they must be cooled well below the temperature of the objects they image. This need is analogous to the difficulty in observing visual images with a bright light shining in one's eyes. Most thermal imagers are cooled with cryogenic fluids such as argon, dispensed from cylinders which last no more than several hours (system cost: $\geq$ \$25,000). There are, however, thermoelectrically cooled systems which require only continuous electrical power, although at a higher cost ($\sim$ \$50,000). The cooling systems also add some bulk and fragility, eliminating the possibility of mounting on small, fast robot arms. Gantry mounting or mounting on large, slow moving arms is a feasible option.

In addition, since the wavelengths received are considerably longer than visible light, spatial resolution is diminished, especially at the low temperature range. The most inexpensive systems are limited to a small number of colors ($\sim$ 16) in the display, so representation of large temperature ranges in one image will cost considerable temperature resolution. The Hughes systems can, however, change the display range and bias on-line, so selective smaller temperature gradients can be presented. Hughes also offers 128 and 256 color models, again at a higher price. Complete imaging systems are available, which include the image processing techniques common to many visual imagers (image subtraction, filtering, etc.).

Finally, it should be noted that thermal imagers are <u>not</u> cameras. it are thermal sensors whose inputs are scanned in from the field of view by mirrors [Lloyd]. They are considerably slower than CCDs, having time constants on the order of 100 ms.

E. Force and Tactile Sensors

In the event that desired visual information is unavailable due to poor lighting, occlusion, or surface reflectance problems, surface features of any object (including the refueling hose and aircraft surface), touch sensors can be used instead. The possible uses for force and other touch sensors are many, and include: grip verification and force control for the end effector, position sensing through detection of wheel contact on the floor or ground, surface exploration and feature detection on the plane's surface, perhaps for increased aircraft identification confidence or location of visually occluded connectors and ports.

A difficulty in the implementation of tactile sensors is commercial availability. Most research in this area is in development of the sensors themselves, so that selection of a commercial device is difficult, especially tactile array sensors. Luckily, in the large scale robotics and manipulation issues necessary here, sensitive sensor arrays are not crucial. Examination of commercial products reveals three types of sensors which are inexpensive and useful: the force sensing resistor, the force/torque sensing wrist module (a strain gage device), and the remote center compliance with lockout.

The force sensing resistor (FSR) is available from Interlink Electronics (Santa Barbara, CA). It is a film of semiconducting material whose resistance decreases as a power of the applied force. The manufacturer's information bulletins suggest that while the force sensing capability is useful, accurate force sensing should be left to strain gage-type devices. The FSR is thin, flexible, mechanically robust, and versatile. It is available in an "XYZ" configuration, which acts as a contact location sensor with resolution of .002″ in the X-Y plane, as well as a force sensor along the Z axis. The basic sensor can be used on gripper faces to monitor gripping, contact force and object weight. As an XYZ sensor, it becomes useful for object exploration and recognition as discussed to some extent in [Bay, Bay & Hemami]. In that work, it is stated that the signal processing techniques developed are most useful when accurate position sensing is available, but when force sensing is relatively inaccurate. This is exactly the situation with the Interlink sensor. Detailed data and customization information is available from Interlink. and a starter designer's kit will provide an assortment of 18 sensors and application notes for $79.95.

EXPERIMENT - Two sample FSRs were obtained from Interlink. and the mass-resistance profile shown in Figure 1 was obtained. It should be noted that the curve is a log-iog plot. demonstrating

that the sensor is a highly nonlinear device. A further problem discovered in laboratory experimentation with the sample FSRs was "creep." Though not specified in the manufacturer's data sheet, it was found that over a period of approximately ten minutes, the resistance of the sensor drops with a constant mass applied. It does decay to a constant and repeatable value, but appears noisy and unpredictable during the transients. At the time of this writing, experiments were in progress to reject the creep, either through electronic processing of the measurements, mechanical damping, or software compensation.

The force/torque sensing wrist senses the resultant forces and torques on the end effector or hand. It consists of as many as 16 strain gages mounted on each side of metallic cross-bridges which attach the hand to the wrist. With this redundancy, simple signal processing techniques can easily resolve a 3-axis torque and a 3-axis force (see Figure 2). The devices are widely used in robotics, since joint torques are difficult to measure and are invariably contaminated with nonlinear disturbances such as gear flexibility, backlash, hysteresis, and friction. In the aircraft turnaround operation, such a wrist will be necessary not only to monitor workpiece forces, but also to monitor contact forces, since forcing a part contact with the aircraft skin should be tightly monitored and controlled. Most force/torque sensors are provided with a dedicated controller, which provides the somewhat complex operation of strain gage signal pickup, amplification, conditioning, and response to requests for force/torque information from the host computer.

The remote center compliance (RCC) is a well known device intended to facilitate manipulation tasks, especially small parts mating. In principle, it provides a spring-dashpot combination along all six axes of hand motion, and is constructed so that it "gives" in reaction to applied force or moments. Inaccurate parts mating, for example, the peg-in-the-hole task, is accomplished through passive compliance rather than active control. This function will be useful for any pin insertion or latching tasks, including aircraft electrical grounding and refueling nozzle insertion. During refueling with a robotic boom or gantry robot, the function of the RCC is the same as that of the flexibility in the hose during an actual aerial refueling.

The "lockout" feature simply locks the compliance to reduce wobble and undesired flex during unobstructed and non-contact moves, thereby also facilitating use of a force/torque sensor in conjunction with an RCC. RCC's, as well as force-sensing wrists, are available from Assurance Technologies (Garner, NC) or PFA, Inc. (New Berlin, WI). Force/torque sensors cost $4000 - $10,000, depending on the payload capability ($\sim$ 15 lb./10 in.-lb. $\sim$ 350 lb./2100 in.-lb.), and RCC's are about $500, or $1800 with lockout. Some robotics literature describes a combination of the two devices in one module, the "instrumented RCC," but a commercially available product was not found.

IV. RECOMMENDATIONS

The following paragraphs summarize the final recommendations of the summer investigation, particularly with regard to the objectives listed above.

A. SENSORS - Because of the problems associated with ultrasonic transmitters/receivers as mentioned in section III.A. they are probably not suitable for multiple mountings on an articulated arm. Their use should be limited to range estimation by single mobile robots. Because of stringent safety requirements dictating accurate range sensing and obstacle avoidance, laser range finders are better suited to the flightline environment. Microwave devices should be further investigated. Because of the low tolerances allowed, it is strongly recommended that the laboratory upgrade to industrial grade robot arms, and equip them with force/torque sensing wrists and RCC devices. Contact applications around critical and expensive flightline equipment justify their cost. Tactile and gripper-mounted array sensors beyond basic force sensing should be implemented in the laboratory only if artificially intelligent autonomous manipulation or shape description (see D. ROBOTIC PERCEPTION ..., below) is desired. In this case, processing techniques are widely available in the voluminous robotics and image processing sources literature references. The CCD overhead and gripper-mounted vision systems are adequate for laboratory study and should be useful on the flightline. Their use should be continued, and augmented with additional vision processing software. An actual flightline operation might also benefit from thermal imaging, although the cost of these systems is prohibitive in the laboratory. A useful comparison of sensor technologies was found in [Espiau] and is summarized in Figure 3.

B. AUTOMATION of SELECTED TURNAROUND FUNCTIONS - As recommended by the Battelle report, refueling is the primary turnaround function currently suited to full automation. Bomb and missile loading is currently too intricate a task to be fully automated, although gross ordnance transfer operations are feasible with semi-autonomous guided vehicles. Automated reloading will not be a candidate for full automation until "quick connect" fittings are implemented, implying substantial re-design of the pylon, retaining, and arming structures on the plane and the missile/bomb. Automated visual/thermal inspection is a technologically feasible application for mobile robots and should be considered for integration with the refueling apparatus. Wheel chocking is best performed with re-design of the chocking device, perhaps with inflatable bladders over which the aircraft might be taxied prior to inflation. Grounding, umbilical connection, and initial inspection can be accomplished with the same overhead apparatus as the refueling operation, but might imply identifiable visual markers on the surface of the aircraft. Ground plug insertion and fuel hose mating requires some passive compliance device such as an RCC.

C. SIGNAL PROCESSING - It was discovered from the literature that current multi-sensor fusion techniques are driven uniquely by the particular assortment of sensors in the application. No accepted, unified multi-sensor integration paradigm exists, and should be the topic of further research. Sensor subsystems should become modular ("plug-in and go"). Until the time that this becomes possible, sensor fusion software will be custom-written for each application. Multi-sensor integration for problems similar to the refueling task are considered in the literature [Abidi] and would make a good guide for refueling operations. Obstacle avoidance procedures are common in robotics literature and will be necessary here. Especially relevant and potentially problematic is the problem of avoiding collision between the refueling boom and the aircraft. This is inherently a three-dimensional mapping and avoidance problem, as opposed to the much easier planar problem. A promising approach is to create a ficticious repulsive force exerted by the surface of any obstacle detected by the range sensors [Cheung], although this is most effective for fixed obstacles. The obstacles present then create an artificial "force field," repelling each link of the robot arm away from object surfaces. This is an elegant and passive technique and requires no global planning and mapping.

D. ROBOTIC PERCEPTION AND SHAPE DESCRIPTION - In remote locations or bare base operations, it has already been determined that camouflaged aircraft may be difficult to localize by camera. This may be problematic if insufficient data is available to precisely fix the physical coordinates of the various aircraft fixtures. To remedy this situation, salient features of the aircraft should be accurately fixed with other sensor inputs. First, if the inflatable bladder or floor-mounted automatic chock is used, this device could also be instrumented to provide approximate resting location of the aircraft tires. Then, if the end effector is equipped with either range or touch sensing, the surface of the aircraft can be explored, and measured geometric features matched with the known characteristics of the aircraft model. Sensor information useful for this procedure includes range, surface normal direction, curvature, and even force (for constrained motion on unknown surfaces). This information can then be fused with visual information so that a surface model can be fitted to the data, and coordinate references can be estimated. The procedures specified in [Bay & Hemami] will accomplish this estimation and the exploration tasks are being investigated [Bay]. Other promising approaches include neural network classifiers [Norwood]. It is recommended that hardware experimentation with tactile shape estimation be conducted on aircraft models in order to verify the theory in the laboratory.

E. OVERALL - The authors of the Battelle report [Smith] used a technique of assigning scores to various turnaround functions based on feasibility, potential time savings, etc. The combined score then ranked the functions in order of strength as candidates for full automation. Curiously, the

functions of bomb loading and AIM-9 missile loading ranked higher than refueling, but refueling was nevertheless chosen as the best candidate for immediate prototype development.

Based on the analysis of sensor equipment and manipulator hardware available, refueling does appear to be the best candidate for automation (along with the associated necessities of grounding, chocking, etc.). Next, noncontact inspection should be implemented. This can be done with currently available remote sensors. The primary difficulty with inspection is the comparative time required to scan the entire surface of the aircraft at high resolution, something humans do well. The primary advantage is the ability of the robotic system to exploit sensors (X-ray, etc), which can detect minute or invisible flaws.

Autonomous bomb and missile loading should be left for further study in the future. While mounting and arming the weapons requires intricate maneuvers and fine dexterity, handling the weapons requires high strength and close supervision. This wide range of capabilities is difficult to build into an autonomous system, and it requires technology currently unavailable. Although non-autonomous force amplifiers and human-in-the-loop mechanical assistance concepts [Kazerooni] show promise, these operations are so safety critical that more development and testing time is required.

It has been said that robotics can be defined as the marriage of artificial intelligence and movable physical objects. If this is true, the current state of industrial applied robotics is better defined as automation, which implies little intelligence. Although crew exposure to CB agents and physical workload advantages are immediately realizable, time and cost savings currently are only appreciable in repetetive tasks requiring little sensory perception and rational decision-making. Since the comprehensive rapid aircraft turnaround is such a complex, sensory intensive, low tolerance problem with little room for error, considerable further research is necessary.

Acknowledgements

REFERENCES

Abidi, M. A., and R. C. Gonzalez, "The Use of Multisensor Data for Robotic Applications," IEEE Transactions on Robotics and Automation, vol. 6, no. 2, April 1990, pp. 159-177.

Bay, John S., "A Fully Autonomous Active Sensor-Based Exploration Concept for Shape Sensing Robots," subm. to IEEE Transactions on Systems, Man, and Cybernetics, May 1990.

Bay, John S., and H. Hemami, "Dynamics of a Learning Controller for Surface Tracking Robots on Unknown Surfaces," Proceedings of the 1990 IEEE International Conference on Robotics and Automation, Cincinatti, OH, 1990, pp. 1910-1915.

Beckerman, M., and Oblow, E. M., "Treatment of Systematic Errors in the Processing of Wide-Angle Sonar Sensing Data for Robotic Navigation," IEEE Transactions on Robotics and Automation, vol. 6, no. 2, April 1990, pp. 136-145.

Beranek, Leo L., ed., Noise Reduction, New York, McGraw-Hill, 1960.

Cheung, E., and V. Lumelsky, "Motion Planning for a Whole-Sensitive Robot Arm Manipulator," Proceedings of the IEEE International Conference on Robotics and Automation, Cincinatti, OH, 1990, pp. 344-349.

Contaq Technologies Corporation, 15 Main Street, Bristol, Vermont, 05443, (803) 453-3332.

Espiau, B., "An Overview of Local Environment Sensing in Robotics Applications,", in Dario, Paolo, ed., Sensors and Sensory Systems for Advanced Robots, New York, Springer-Verlag, NATO ASI series vol. 43, 1988.

Hirzinger, G., and J. Dietrich, "MultisensoryFeedback Including Cooperative Robots," in Dario, Paolo, ed., Sensors and Sensory Systems for Advanced Robots, New York, Springer-Verlag, NATO ASI series vol. 43, 1988.

Hudson, Richard Jr., Infrared System Engineering, New York, Wiley-Interscience, 1969.

Kazerooni, H., "Human-Robot Interaction via the Transfer of Power and Information Signals," IEEE Transactions on Systems, Man, and Cybernetics, vol. 20, no. 2, March/April 1990, pp. 450-463.

Lippmann, Richard P., "An Introduction to Computing with Neural Nets," IEEE Acoustics, Speech, and Signal Processing Magazine, April 1987, pp. 4-22.

Lloyd, J. M., Thermal Imaging Systems, New York, Plenum Press, 1975.

Norwood, John, Peter Weiland, and J. B. Cheatham, "Robotic Navigation from Local Sensory Data Using Neural Networks," Proceedings of the International Robots & Vision Automation Conference, Detroit, Michigan, 1990, pp. 7-130 -- 7-138.

Polaroid Corporation, Ultrasonic Ranging System, Manual for Ultrasonic Ranging System Experimentation Kit.

Casals, Alicia, ed., Sensor Devices and Systems for Robotics, New York, Springer-Verlag, NATO ASI series, vol. 52, 1988.

Smith, John L., et al., Study of Robotic Concepts for Aircraft Turnaround, Technical report no. AFWAL-TR-88-3058, Battelle Memorial Institute for Flight Dynamics Laboratory, AF Wright Aeronautical Laboratories, WPAFB, OH, 1988.
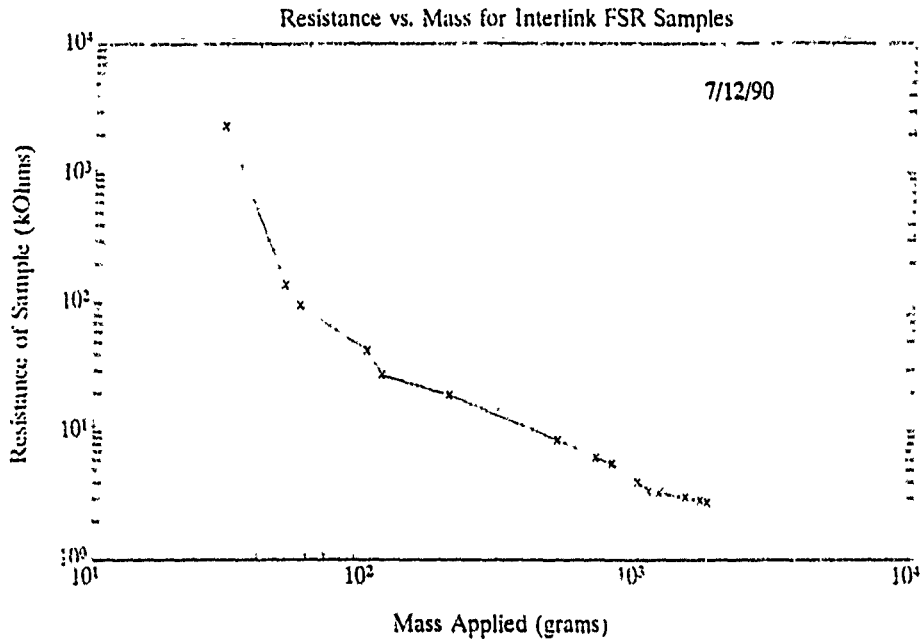
Figure 1. Resistance vs. mass applied for the sample Interlink FSR. Note the log-log scale and the outlier, probably due to the creep discussed in section III.
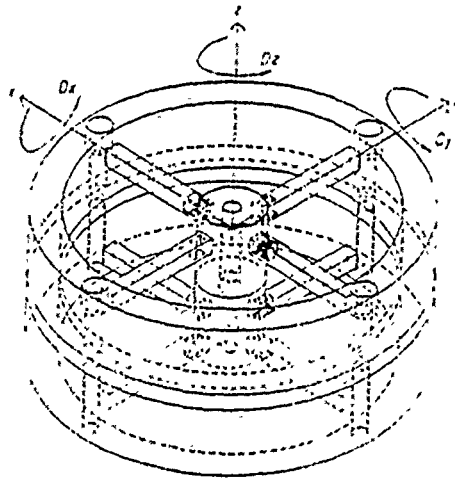
Figure 2. Structure of the strain-gage force/torque sensor representative of commercially available products. External processing extracts directional information from the bridge-connected strain-gage pairs [Hirzinger].

Figure 3. Summary of sensor types, including applicability information (from [Espiau]). The symbol "/" denotes "not applicable".

FINAL REPORT

Influence of Static and Dynamic Aeroelastic Constraints on the Optimal Structural Design of Flight Vehicles

Prepared by: Dr. Franklin E. Eastep and Anne Stephenson

Academic Rank: Professor   Graduate Student

Department and   Aerospace Engineering
   University: University of Dayton

Research Location: Flight Dynamics, Laboratory, Structures and
                    Dynamics

                    Division, Analysis/Optimization Branch

USAF Researcher: Dr. Vipperla B. Venkayya

Date: 1 November 1990

COntract No.:   F49620-88-C-0053

INFLUENCE OF STATIC AND DYNAMIC AEROELASTIC CONSTRAINTS OF THE OPTIMAL STRUCTURAL DESIGN OF FLIGHT VEHICLE STRUCTURES

Franklin E. Eastep, Ph.D.

Department of Aerospace Engineering
University of Dayton, Dayton, OH 45469-0227

ABSTRACT

This investigation focused upon the structural weight optimized design of a fighter-type wing of low aspect ratio using ASTROS. The optimal weight redesign of a preliminary finite element model representing the wing structure is obtained with the constraints on strength, control reversal and flutter imposed using both subsonic and supersonic aerodynamic theories. It is demonstrated that the optimization capabilities of the ASTRO procedure are well suited for the preliminary structure design environment. ASTROS gives to the structural designer the capability to develop unique solutions to the design problem facing flight vehicle structures with the many constraints. Recommendations are made to include a transonic aerodynamic formulation with ASTROS for the structural design of a flight vehicle over the entire Mach number regime.

ACKNOWLEDGEMENT

I. INTRODUCTION

An aircraft structural designer must consider aeroelastic instabilities (i.e. flutter, divergence and control reversal) in a addition to the strength requirements for the structural design of high performance aircraft. In particular, he must design a structure such that the maximum operational flight velocity is less than (by 15%) the velocity of the critical aeroelastic instability while still insuring satisfactory strength at the velocity of the critical aeroelastic instability. The critical aeroelastic instability is defined to be the least value of the flutter, divergence or control reversal velocities. In addition, the structural designer desires to adjust the structural sizes to minimize the structural weight while insuring that the maximum flight velocity is less than the critical aeroelastic instability velocity.

In recent years, structural optimization as needed and used by the aerospace industry has expanded in scope to include such additional disciplines as static and dynamic aeroelasticity, composite materials, aeroelastic tailoring, etc. One of the more promising multidiciplinary codes presently under development is the Automated Structural Optimization System (ASTROS) [1-3]. In this computer code, static, dynamic, and frequency response finite element structural modules, subsonic and supersonic steady and unsteady aerodynamic modules, and an optimization module are combined and allow for either analysis or optimized design of given aircraft configurations. Interfering surface aerodynamics are incorporated to handle the aerodynamic modeling of combinations of

wings, tails, canards, fuselages, and stores. Structures are represented by fully built-up finite element models, constructed from rod, membrane, shear, plate, and other elements. Static and dynamic aeroelastic capabilities include trim, lift effectiveness, aileron effectiveness, gust response, and flutter analysis. The optimization and aeroelasticity modules of this code were used as a tool for the structural optimization of fully built-up finite element wing models in subsonic and supersonic flow with strength as well as static and dynamic aeroelastic constraints.

First, the performance of the flutter analysis/optimization module was evaluated against results from the other methods and codes, i.e., the large scale finite element codes MSC/NASTRAN [4] and COSMIC-NASTRAN [5]. Here, a generic low aspect ratio fighter type wing model of reasonable complexity (86 nodal points) was analyzed for flutter by all three codes, then optimized for minimum weight with strength and various aeroelastic constraints by ASTROS using a small number of design variables (=5) and design variable heavy linking. The optimum designs were again analyzed by all three codes, i.e., free vibration mode shapes, flutter speeds, and flutter mode shapes were determined and the results from the different codes were compared. Also, for some of the cases, the free mode shapes, the flutter speeds, and the flutter mode shapes were monitored from iteration to iteration during the course of the optimization, together with the constraint behavior to understand the performance of the optimization routine especially during steps in the optimization when various different constraints were active at the same time. In the investigation, the strength constraints

were incorporated in the form of a 9-g symmetric pull-up maneuver, the static aeroelastic constraints were based on control effectiveness and aileron reversal, and the dynamic aeroelastic constraint was represented by a minimum flutter speed.

II. OBJECTIVES

The objective of this study is to demonstrate that the optimization capabilities of ASTROS are well suited for the preliminary structural design environment. The constraints of strength, control reversal and flutter speed are imposed on a preliminary finite element structural model of fighter-type wing to obtain a weight optimized wing structure. It is desired to show the ability for ASTROS to simultaneously consider many constraints from each of several disciplines allowing the structural designer to develop non-intuitive solutions of the complex design problem placed on modern flight vehicle structures.

III. EQUATIONS FOR STRUCTURAL AND AEROELASTIC DESIGN

Consider the swept and tapered thin wing shown in Figure 1 which is typical of a wing investigated herein. The underlying structure of the wing is represented using finite elements as shown in Figure 2 which is typical of a wing built-up structure. In addition, also shown in Figure 1 is an aerodynamic flap located near the wing tip whose purpose, in conjunction with a flap located on the opposing wing, is to create a rolling moment for rolling of the flight vehicle.

Begin the investigation by first determining the static aeroelastic deformation of the wing produced by the aerodynamic

forces acting on the wing. The equation of equilibrium for the bending and twisting of the wing is written in matrix form conforming to the finite element formulation as:

$$[K]\{u\} = q[A]\{u\} + q[A^R]\{\alpha^R\} + q[A^F]\{\beta\} \tag{1}$$

where $\{u\}$ represents the bending and twisting deformation at the nodal points of the finite element model $\{\alpha^R\}$ is some initial angle of attack of the wing and $\{\beta\}$ is flap setting angle. The parameter q is the dynamic pressure defined as $1/2\ \rho_\infty Uo^2$. In equation (1) the matrix [K] is the structural stiffness and [A] is the aerodynamic flexibility matrix representing the aerodynamic lift and moment generated by the deformation of the wing beyond that which depends on the initial angle of attack and flap setting. The equilibrium equation (1) can be written in the usual form for determining the deformation $\{u\}$ as

$$[[k] - q[A]]\{u\} = q[A^R]\{\alpha^R\} + q[A^F]\{\beta\}. \tag{2}$$

1. Wing Divergence Velocity

The wing divergence velocity is independent of the initial wing angle of attack and the flap setting angle. At the divergence velocity there is a balancing between the aerodynamic force produced by wing deformation and the structural resorting force represented by the product of the stiffness and wing deformation. To determine the divergence dynamic pressure or velocity equation

(2) is written in homogenous form.

$$[[K] - q_{DIV} [A]] \{u\} = \{o\} \qquad (3)$$

Equation (3) is an eigen-problem, from which, the lowest eigen-value can be determined and represents the divergence dynamic pressure. Additionally, the accompanying deformation shape, the eigen-vector can be determined. It should be noted that the aerodynamic influence metric [A] in general is not a symmetric matrix which requires special numerical treatment of the eigen-problem represented by equation (3) because of the possibility of complex eigenvalues. An inverse-power method was used for the eigenvalue extraction from equation (3). The divergence velocity can then be written as:

$$U_{o_{DIV}} = \sqrt{\frac{2q_{DIV}}{\rho_\infty}} \qquad (4)$$

2. Lift Effectiveness

Lift effectiveness is defined to be the ratio of wing lift produced on a flexible wing to that produced on a rigid wing at a defined value of dynamic pressure. The lift effectiveness is computed at a certain velocity where the flap setting angle, β is zero. The lift effectiveness is then:

$$E^L = \frac{L_{rigid} + L_{flex}}{L_{rigid}} = \frac{q\{h\}^T[A^R]\{\alpha^R\} + q\{h\}^T[A]\{u\}}{q\{h\}^T[A^R]\{\alpha^R\}} \qquad (5)$$

where the matrix {h} are aerodynamic panel lengths. The wing deformation vector {u} in equation (5) is computed from equation (2) with the flap setting { β } equated to zero as:

$$\{u\} = q[[K] - q[A]]^{-1}[A^R]\{\alpha^R\} \qquad (6)$$

Substituting equation (6) into equation (5) allows one to determine the lift effectiveness.

$$E^L = \frac{\{h\}^T[[I] + q[A][[K] - q[A]]^{-1}][A^R]\{\alpha^R\}}{\{h\}^T[A^R]\{\alpha^R\}} \qquad (7)$$

3. Control-Surface Effectiveness

The purpose of the flap shown in Figure 1 located near the wing tip is to produce an aerodynamic rolling moment of the wing. The deformation of the wing surface will produce additional rolling moments beyond those produced if the wing was rigid. The rolling moment effectiveness is defined to be the ratio of the rolling moment resulting from a flap rotation on flexible wing to the rolling moment resulting from a flap rotation on a rigid wing. The control effectiveness is computed for a specified velocity with initial wing angle of attack equated to zero in equation (2)

$$E^c = \frac{M_{xrigid} + M_{xflex}}{M_{xrigid}} = \frac{q\{h\}^T[A^F]\{\beta\} + q\{h\}^T[A]\{u\}}{q\{h\}^c[A^F]\{\beta\}} \qquad (8)$$

where the matrix $\{h\}^T$ are products of panel lengths and moment arms to the nodal points of the structural finite element model. Additionally, it should be noted that the matrix $\{ \beta \}$ has a non-zero value _only_ for elements where the flap is located. Once again the deformation vector $\{u\}$ can be determined from the equilibrium equation (2) but in this case the initial angle of attack is equated to zero.

$$\{u\} = q[[K] - q[A]]^{-1}[A^F]\{\beta\} \qquad (9)$$

Substituting equation (9) into equation (8) the flap effectiveness for creating a rolling moment on a flexible wing is determined as:

$$E^c = \frac{\{h\}^T[[I] + q[A][[K] - q[A]]^{-1}][A^F]\{\beta\}}{\{h\}^T[A^F]\{\beta\}} \qquad (10)$$

The flap control effectiveness can thus be determined for any specified value of velocity or dynamic pressure.

4. Control-Surface Reversal

From consideration of equation (10) the possibility exists that there is a velocity or dynamic pressure where the control effectiveness E^c is zero. The velocity at which the control effectiveness is zero is called the control reversal velocity for the rolling moment produced by a flap deflection on a flexible wing

is zero and any further increase of velocity will produce a rolling moment tending to roll the wing in a direction opposite to the roll produced by a flap rotation on a rigid wing. The reversal velocity can be determined from a graphical display of equation (10), E^c for various values of dynamic pressure.

5. Flutter Analysis

To determine the flutter speed of a flight vehicle wing shown in Figures 1 and 2 it is necessary to obtain the equation of motion representing deformation of the structure when subjected to the unsteady aerodynamics forces. Two aerodynamic methods are available in ASTROS, the doublet lattice method for subsonic velocities and the potential gradient method for supersonic velocities; unfortunately there is not an aerodynamic method for transonic velocities. The flutter stability analysis is based on the p-k method with the equation of motion governing wing deformation as

$$\left[\left(\frac{U_o}{b} \right)^2 P^2 [M] + [K] - q \left(\frac{p}{k} [Q^I] + Q^R \right) \right] \{u\} = \{o\} \tag{11}$$

where p is a complex eigenvalue, [M] is the consistent mass matrix and $[Q] = [Q^R] + i [Q^I]$ is the complex aerodynamic coefficient matrix. The flutter speed is determined by plotting the complex eigenvalue on a v-g diagram.

6. Design Problem

The concept of optimization involves the alteration of design

variables (thickness) to achieve desired results. The optimization goal is to find a vector $\vec{X}$ of n design variables x_i, $(i=1,2...,n)$, which will minimize a multivariable function $f(x)$ subject to m constraints:

$$g_j(x) = G_j(x) - \overline{G}_j \geq o(j = 1,2,...,m) \tag{12}$$

and side constraints

$$x_i^l \leq x_i \leq x_i^u \ (i = 1,2,...,n) \tag{13}$$

Here the design variable vector $\vec{X}$ controls the finite element sizes, the constraints $G_j(x)$ are strength, aeroelastic divergence, control effectiveness and flutter speed, $\overline{G}_j$ are the limits, $f(x)$ represents the structural weight and the superscripts land u represent the lower and upper bounds.

VI. NUMERICAL RESULTS AND DISCUSSION

Two fighter type wing structural models shown in Figure 3 were selected to be representative of wing of a low aspect ratio. The preliminary design model was selected for weight optimization under strength as well as static and dynamic aeroelastic constraints for a reasonable sized model to allow parametric investigations. The large scale model would likely be used in the later phases of the structural design process.

Initially, the sizes and locations of the structural elements of a 10 spar, 3 rib wing shown in Figure 3 were selected and as a

nominal structural model. Additionally, concentrated weights were placed at the structural nodal points to simulate the non-structural weights to represent fuel, actuators and store mass. The aerodynamics modeling for the nominal and optimized structure was selected to be 36 aerodynamics boxes with 6 chord divisions and 6 span divisions.

1. Nominal Wing Structure

The nominal wing structure was used with ASTROS to determine the individual stresses in the structural elements resulting from a 9-g pull-up at a Mach no of 0.85 at sea level. Additionally, the flutter speed of approximately 32,500 in/sec for the nominal model was found.

Finally the roll effectiveness for a roll control system was determined for the nominal wing structure at a dynamic pressure near the control reversal dynamic pressure. The variation of control effectiveness and aileron effectiveness of the nominal structure is displayed in Figure 4. As indicated in Figure 4 the reversal dynamic pressure for the nominal model was approximately $45 \#\backslash in^2$.

As a side investigation to determine the reliability of ASTROS to predict a valid flutter speed, the flutter speed as predicted using the p-k method of ASTROS was compared to the flutter speed predicted using both the k and p-k methods of COSMIC-NASTRAN. As shown in Figure 5; the v-g diagram of the critical flutter mode of the nominal structure, all three methods predict exactly the same flutter velocity even though all three methods yield very different

damping values at velocities other than at flutter. In addition, comparison of flutter velocities predicted by three methods is displayed in Figure 6 for an optimized structural model (structural weight less than the nominal model using a preassigned flutter speed constraint). The v-g diagram, Figure 6 again demonstrates the same prediction of flutter velocity by the three techniques and for the optimal structural model the damping values, g, are approximately the same for all three techniques at all velocities. The displayed results of Figures 5 and 6 demonstrate that the p-k flutter techniques of ASTROS predicts the correct flutter velocity in the analysis or design phase of ASTROS.

2. Optimized Wing Structure

The fighter wing structural model was resized and optimized using ASTROS with both single and multiple constraints active at any given time as shown in Table I. The initial optimal structure model was obtained for a 9-g symmetric pull-up maneuver with a Von-Mises stress constraint with prescribed stress yield values. The following optimization studies were conducted using as minimum sizes for the individual structural members those obtained from the above strength optimized model.

Next, the structural model was resized and optimized using a constraint of an improvement of the reversal dynamic pressure from $45\#/in^2$ to $52\#/in^2$. The increase in the reversal dynamic pressure was accomplished while the structural weight of the optimized wing was reduced from 497# (nominal) to 409# (optimal) as shown in Table I. As a comparison problem with the reversal velocity used as a

constraint, the structure was resized to yield the same reversal dynamic pressure as the nominal structure and weight optimized to reduce the structural weight further to 347#. The corresponding flutter velocities for these two cases with constraints on reversal dynamic pressure are indicated in Table I and are 3.23 x 10^4(in/sec) and 2.97 x 10^4(in/sec) respectively.

Finally the structural model was resized and optimized using multiple constraints. In this case it was required that the reversal dynamic pressure be the same as reversal of the nominal structure and the flutter speed be increased to 3.1 x 10^4 in/sec. This increase of flutter speed is beyond that obtained when only a single constraint on reversal pressure was imposed. In this manner the structural designer has the added advantage of precisely placing the velocity of certain aeroelastic instabilities relative to other velocities of aeroelastic instabilities. Here we desired to make improvements in the flutter velocity which is larger that the reversal velocity while the structure is weight optimized. In this particular case the weight of the optimized structure was obtained as 438#.

The previous optimization studies were accomplished in the subsonic Mach number regime (i.e. M_∞ = 0.85) using a doublet lattice aerodynamic formulation. Using subsonic aerodynamic theories resulted in the prediction of aeroelastic instability velocities in the supersonic regime. To remove this inconsistence in problem formulation a supersonic aerodynamic formulation, called a potential gradient method, was used in an optimization study. The Mach number was selected to be 1.2 and the structure was

TABLE I

WEIGHT OPTIMIZED WING WITH VARIOUS CONSTRAINTS
(ALTITUDE - SEA LEVEL)

MODEL	CONSTRAINT	REVERSAL-q(#/in²)	FLUTTER SPEED(IN/SEC)	STRUCTURAL
NOMINAL	NONE (M=0.85) (M=1.2)	45 55	3.25×10^4 3.09×10^4	497 497
OPTIMAL	STRENGTH (M=0.85) STRENGTH (M=1.2)	20 24		243
OPTIMAL	REVERSAL INCREASE M=0.85	52	3.23×10^4	409
OPTIMAL	REVERSAL = NOMINAL REVERSAL (M=0.85) REVERSAL (M=1.2)	45 45	2.97×10^4 3.01×10^4	349 447
OPTIMAL	FLUTTER (M=0.85) FLUTTER (M=1.2)	46 53	3.25×10^4 3.25×10^4	426 447
OPTIMAL	REVERSAL = NOMINAL (M=0.85) FLUTTER = INCREASED	45	3.3×10^4	438
OPTIMAL	STRENGTH, REV, FLUTTER (M=0.85) STRENGTH, REV, FLUTTER (M=1.2)	45 45	3.10×10^4 3.10×10^4	400 445

resized and optimized with single and multiple constraints also shown in Table I. The selected constraints for M=1.2 were similar to the constraints selected for M=0.85 and the weight reductions for the optimized structure is shown in Table I.

V. CONCLUSIONS

The examples presented in this investigation demonstrate that the optimization capabilities of the ASTROS procedure are well suited for the preliminary design environment. Any number of constraints of strength, divergence, control reversal and flutter can be imposed on general finite element structural models of flight vehicles. The ability to simultaneously consider many constraints from each of several disciplines allows the structural designer to develop non-intuitive solutions to the complex design problem placed on modern flight vehicle structures.

This investigation focused upon the structural weight optimal design of a fighter-type wing of low-aspect ratio. The finite element representation of the structure was sufficiently large and the design variables large enough to serve as a preliminary design model. The optimal weight redesign of the wing structure was obtained with constraints on strength, control reversal and flutter imposed using both subsonic and supersonic aerodynamic theories. The weight savings for the various imposed constraints is shown in Table I.

VI. RECOMMENDATIONS

At the present time only the linear aerodynamic theories of the sub and supersonic Mach number regimes are available for use by

ASTROS in the optimal weight design of flight vehicle structures. The critical flight regime for structural design to preclude the flutter instability is the "transonic" Mach number regime. In this regime the aerodynamic equations are non-linear and not available for use by ASTROS in the very regime that is most critical. It is recommended that a transonic aerodynamic formulation be incorporated with ASTROS to obtain the weight optimized design of flight vehicle structures for all Mach number regimes. The aerodynamic formulation exists at the present time as XTRAN3 being developed by the AF Flight Dynamics Lab. XTRAN3 is a numerical solution scheme using finite differencing method to solve the non-linear aerodynamic flow equations.

REFERENCES

1.   Johnson, E.H., and Venkayya, V.B., "Automated Structural Optimization System (ASTROS), Volume I - Theoretical Manual", AFWAL-TR-88-3928/I, Air Force Wright Aeronautical, December 1988.

2.   Neill, D.J., Johnson, E.H., and Herendeen, D.L., "Automated Structural Optimization System (ASTROS), Volume II - User's Manual", AFWAL-TR-88-3028/II, Air Force Wright Aeronautical Laboratories, December 1988.

3.   Johnson, E.H., and Neill, D.J., "Automated Structural Optimization System (ASTROS), Volume III - Applications Manual", AFWAL-TR-88-3018/III, Air Force Wright Aeronautical Laboratories, December 1988.

4.   Rodden, W.P., Editor, MSC/NASTRAN Handbook for Aeroelastic Analysis, The MacNeal-Schwendler Corporation, Los Angeles, California, 1987.

5.   "NASTRAN User's Manual", NASA SP-222(08), June 1986.

6.   Striz, A.G., and Venkayya, V.B., "Influence of Structural and Aerodynamic Modeling on Flutter Analysis", Proceedings, 31st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Long Beach, California, April 1990, pp. 110-118.
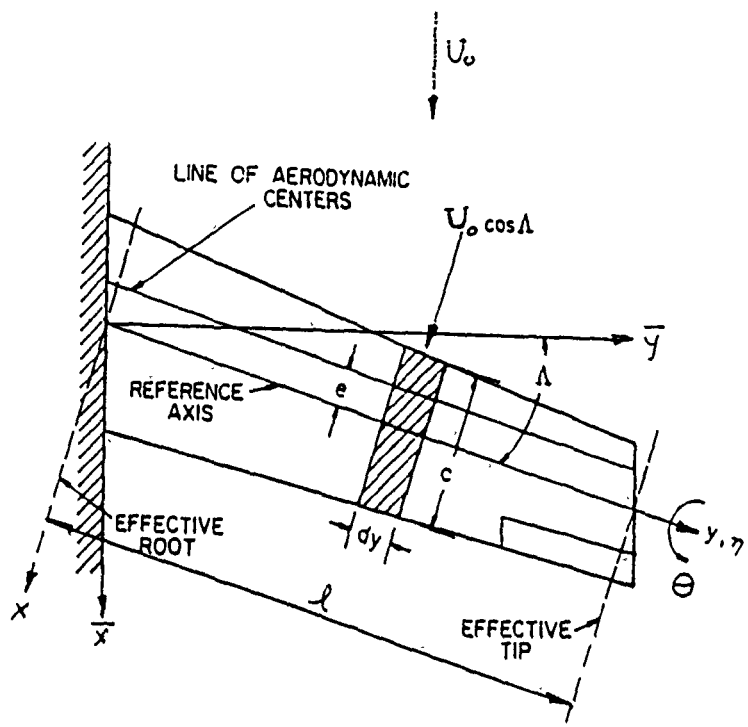
Figure 1. Idealized Wing Planform Geometry.
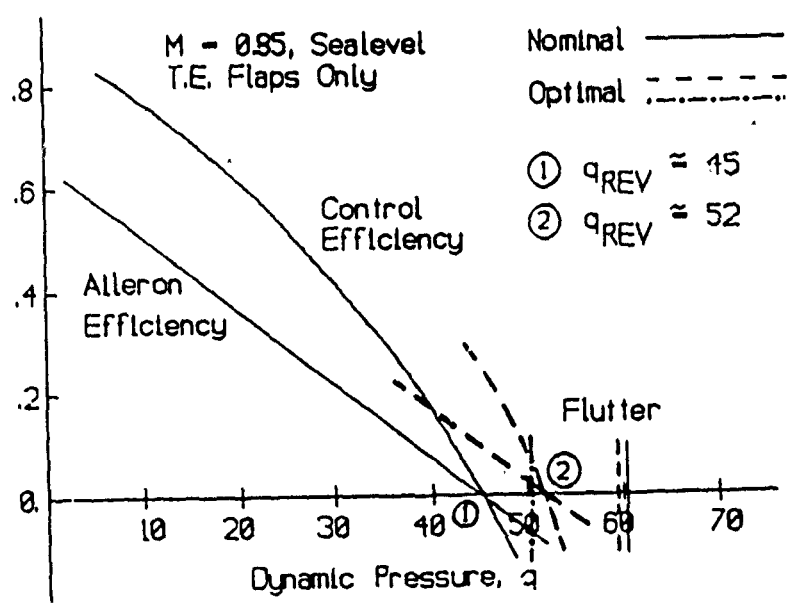
Figure 4. Control Aileron Efficiency versus Dynamic Pressure
(M = 0.85, sealevel).
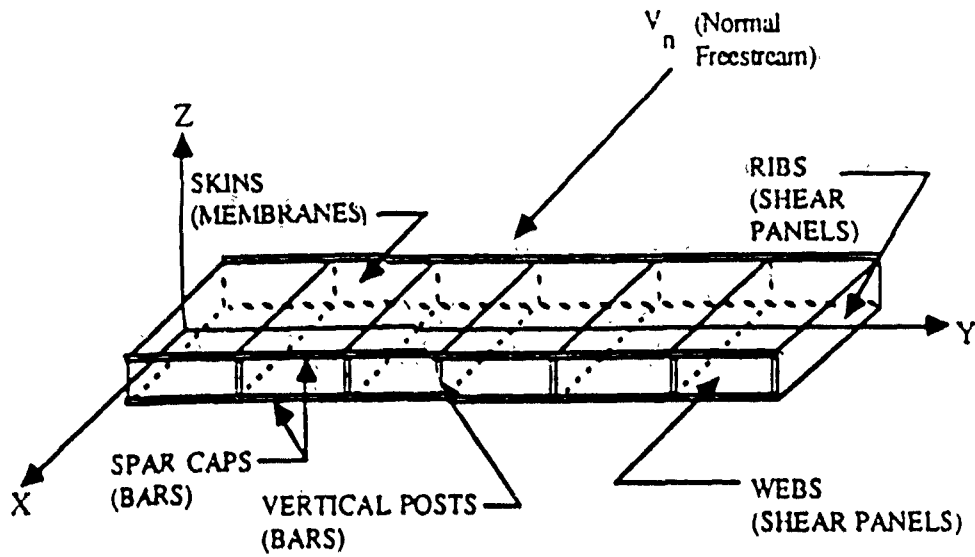
Figure 2. Built-up Wing Structure

1. Preliminary Design Wing Model
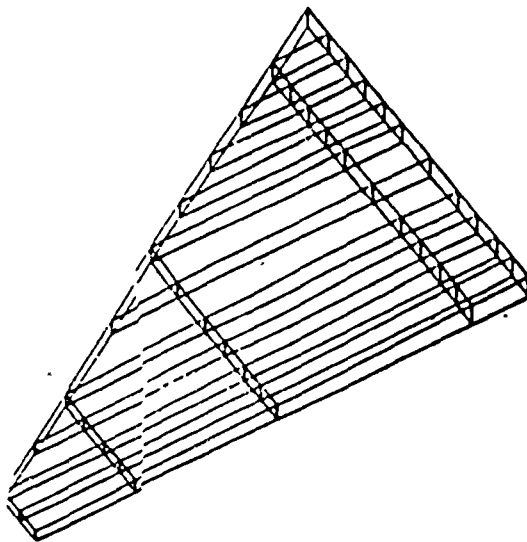
2. Large Scale Wing Model
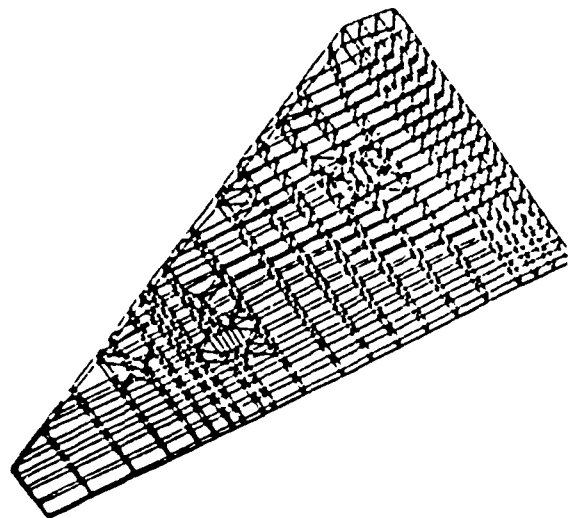
Figure 3.   Finite Element Wing Models for Optimization.

10-SPAR FIGHTER WING

MACH 0.85 AT SEALEVEL
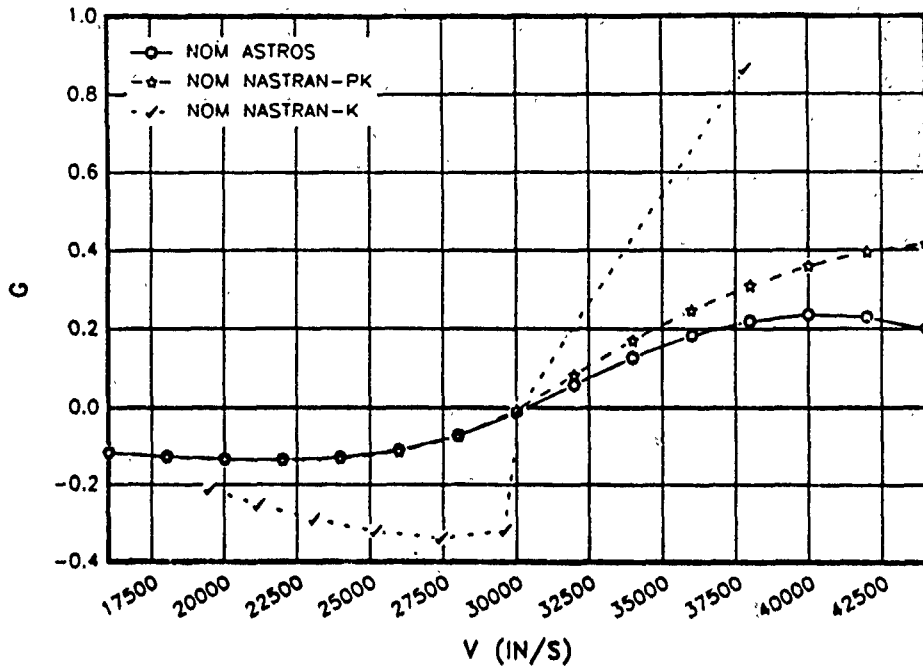
Figure 5. Damping versus Velocity of the Critical Flutter
Mode - Nominal Structure.

10-SPAR FIGHTER WING

MACH 0.85 AT SEALEVEL
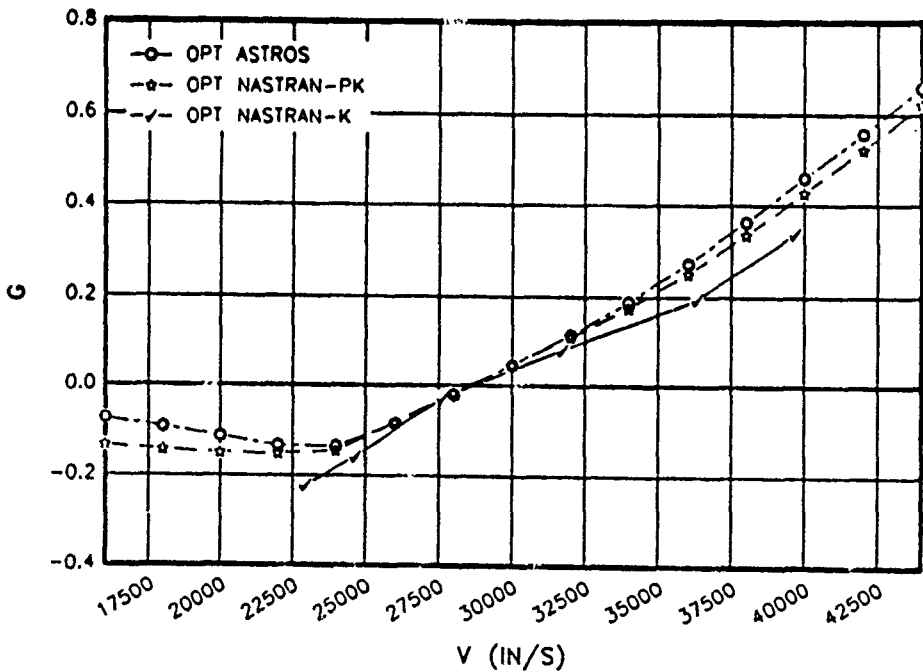
Figure 6. Damping versus Velocity of the Critical Flutter
Mode - Optimal Structure.

99-20

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Location of Crack Tips by Acoustic Emission for
Application To Smart Structures

Prepared by:	Marvin A. Hamstad, PhD
Academic Rank:	Professor
Department and	Engineering Department
University:	University of Denver
Research Location:	WRDC Flight Dynamics Laboratory Wright-Patterson AFB (Dayton) OH 45433-6553
USAF Researcher:	George P. Sendeckyj, PhD
Date:	17 Sept 90
Contract No:	F49620-85-C-0053

Location of Crack Tips by Acoustic Emission
for Application to Smart Structures

by

Marvin A. Hamstad

ABSTRACT

The use of commercial acoustic emission (AE) equipment for location of crack tips in a fatigue test environment was studied. The results from six channels of commercial AE equipment were compared to those derived from waveforms obtained by a two-channel transient recorder. Prior to fatigue cycle monitoring, the AE wave propagation characteristics were extensively studied using pencil lead breaks in a center notch. For the fatigue studies center crack samples of 2024-T351 aluminum were used with hydraulic grips to eliminate extraneous noise over the whole fatigue cycle. Results show that waveform approaches are superior to standard AE systems for location of crack tips. Results also indicated that reopening of a closed crack generates much more AE than that generated at the crack tips. This "unclosure" AE may be of potential use to detect cracks and characterize crack length.

I. INTRODUCTION

The Fatigue, Fracture and Reliability Group of the Flight Dynamics Laboratory of the Wright Research and Development Center is interested in the concept of smart structures. This terminology refers to the ability to nearly continuously sense conditions and states such as loads and structural integrity in aircraft structures [Sendeckyj and Paul, 1989]. The inputs of the sensors into properly programmed computers lead directly to corrective actions thereby by-passing scheduled nondestructive evaluation. Acoustic emission (AE) technology provides one approach to the network of sensors and computers necessary for a smart structure.

AE technology·is concerned with the monitoring of stress waves which are generated by rapid local energy releases in solid materials. The AE signals are generated by a variety of sources such as impacts, phase changes, crack growth, friction, inclusion-particle fracture, and other microdamage sources.

AE technology has a number of useful and potentially unique features for smart structures.applications. First, AE is a passive technique in that stress waves generated at sources throughout a structure propagate to the sensor location. This feature limits the density of sensors required to monitor a structure. Second, AE is a whole volume sensing technique. Third, AE data from multiple sensors can be used to locate the spatial position of the sources within the structure. Fourth, the source mechanism may be identified by use of advanced statistical concepts with waveform feature vectors. Fifth, AE is essentially a real time sensing technique. Sixth, AE is a microscopic technique with high sensitivity. And seventh, AE is at its very basis a measure of the response of a structure to stress.

II. OBJECTIVES OF THE RESEARCH EFFORT

Since a goal of the smart structures concept is to bypass the usual scheduled nondestructive inspections, the application of AE would make its most valuable contribution if crack lengths could be measured by following the position of crack tips in a fatigue environment. AE technology as normally practiced in structures has only been concerned with detecting and

locating the _vicinity_ of a crack. Thus, the smart structures application requires an extension of the technology to locate crack tips. With this perspective in mind, the primary objective of this research was to examine the ability of normal commercial AE instrumentation to distinguish and estimate the difference in lengths of cracks in metal samples in a laboratory environment.

III. EXPERIMENTAL APPROACH:

a) Acoustic Emission Equipment

The existing commercial AE equipment was a 32 channel SPARTAN AE system (CPM software) manufactured by Physical Acoustics Corporation. The operation of this equipment was characterized and checked for calibration by use of an AE Simulator manufactured by Acoustic Emission Associates.

Using a true root-mean-square meter the parasitic noise of the AE system was measured as a function of the system gain. It was found that certain system gains have relatively high parasitic noise and thus give poor sensitivity. To enhance the sensitivity, a system gain of 36 db was chosen so that the parasitic noise was only a small part of the preamplifier noise. The bandpass of the AE preamplifier (Physical Acoustics Corporation) and the Spartan system was 200-400 KHz. The threshold was selected to be a fixed value of 42 dB which was some 3 dB above the background preamplifier noise at the gain of 60 dB. In one test the threshold was set at 40 dB by mistake. Physical Acoustics u30 resonant sensors (diameter about 0.375 in.) were used since they are reasonably small with good sensitivity. All AE data was taken in the high speed mode and post-test processed or analyzed by hand from hit listings.

b) Simulated AE Studies

Simulated AE was used to study wave propagation characteristics and determine the propagation velocity in the specimens. The source of the simulated AE was a standard Pentel pencil lead break (0.3mm dia, 2H hardness). For these studies a 3.62 x 16.0 x 0.258 in. 2024-T351 aluminum specimen was prepared with a 2.0 in. long by 0.12 in. wide (with rounded ends) center through notch (perpendicular to the long direction). The notch allowed lead breaks on its surface. Figure 1 shows the geometry of lead break positions as well as the sensor locations used for various parts of the research. For these studies vacuum grease was used as the AE sensor

couplant, and the sensors were held in place by rubber bands. Since lead breaks have approximately constant amplitude, to simulate the fact that real AE signals vary in peak amplitude the location studies were made with a number of different thresholds relative to the peak amplitude. In addition to use of the AE system waveforms were also taken with two different Nicolet transient recorders (digitization rates of 2 and 10 MHz with 8K memory). The waveforms allowed detailed studies to determine proper settings of the AE system Rearm Time (RTO), Rise Time Time Out (RTTO), and Single Channel Event Time Out (SCETO) parameters. Based on typical waveforms, the values selected were RTTO = 20us, SCETO = 1000us, and RTO = 40 us.

c) Extraneous Noise Tests

Several approaches were taken to deal with extraneous AE noise. First, to eliminate the potential for rotation of the specimen relative to the grips a symmetric center crack specimen was used. Second, smooth faced hydraulic grips were used which apply the transverse load prior to the axial loads. Third, an uncracked aluminum sample (made with grip marks left by the precracking grips) was cycled at the typical crosshead rates used to test the cracked samples. Two or three sensors were mounted with rubber bands and vacuum grease or by adhesive bonding on the sample during these tests. One sensor was placed at the center of the sample for some tests along with two others near the grip edges for all tests. The load to the sample was monitored by the AE System parametrics during these tests. Table 1 shows the number of events for each increasing and decreasing load cycle. Examination of the AE hit listings showed that the few extraneous noise events generated hit one of the sensors near the grips first. This result validated that use of guard sensors would allow identification of any extraneous noise events. Fourth, the sensor array selected for the cracked samples included two guard sensors in addition to the four sensors placed around the crack region (see figure 1).

d) Center Cracked Samples With Load Cycling

The 16.0 x 3.61 x 0.258 in. test samples were prepared from 2024–T351 aluminum plate stock. A 0.50 in. long through slot in the 3.61 in. direction was electron discharge machined (EDM) in the center of each sample. The approximately 0.015 in. wide slot was extended by fatiguing

in an MTS machine at 10 Hz using a load shedding program so that Kmax remained near the AE test level of 10 ksi$\sqrt{in}$ with R=0.01. During the last 0.020 in. of crack growth at each tip the maximum and minimum loads were kept at the value used for the AE monitored fatigue cycles. Table 2 gives the crack lengths of each sample at the conclusion of precracking. Since the starter notch was the same for each crack, the three samples provided three different lengths over which closure related AE could be generated as well as three different crack tip positions. The flat precracks indicated that the cracks were grown in plane strain. Table 3 gives the yield strength, and fracture toughness of typical 2024–T351 material [Damage Tolerant Design Handbook, 1983]. The same reference also has data which predicts that the crack growth rate at the end of precracking should be between 0.1–0.2 um per cycle. The crack growth rate was measured at about 0.4 um per cycle during the last approximately 0.020 in. of growth.

For the AE tests the specimens were moved to a 100,000 lb capacity Instron machine, which is much quieter than the hydraulic MTS machine with servovalve attached directly to the actuator. Since it was impossible to move the sample and grip it with exactly the same alignment in the new machine, each sample in the Instron was cycled an additional 630 cycles at a rate of about 0.4 Hz at the maximum and minimum loads used at the end of the crack sharpening procedure. This procedure grew each crack tip an additional nominal 0.005 in. and resulted in the crack tip adjusting to the slightly different alignment.

The AE test procedure consisted of 6 to 12 cycles at the same load levels used in the resharpening procedure. The crosshead rates and the approximate times for one cycle for these tests are given in table 2. Following these cycles an additional cycle with an overload above the previous peak load level was imposed. The load factors for the overloads are given in table 2. The purpose of the overload was to generate a number of events located at the crack tips as the plastic zone size was increased. During the AE tests, six AE sensors were used as shown in figure 1. The sensors on either side of the crack were the primary monitoring sensors, while the two sensors near the hydraulic grips were used to identify any AE signals which originated outside the crack region. The use of edge mounted sensors in some tests versus surface mounted sensors was based on

an earlier observation that in rod type samples sensors mounted on a specimen face perpendicular to the major stresses released at the AE source gave better location results [Hamstad et al., 1986]. Also, the edge sensors were more sensitive by about 4-6 dB in the wave propagation studies. The sensors were bonded in place using an epoxy mix (Cole Parmer #8778). After sensor installation lead breaks, about 0.5 in from the center of each sensor, were used to check sensitivity. A signal proportional to the load cell output was connected to the parametrics input of the AE system so that the stress intensity level for each AE event was recorded in the data set.

IV. RESULTS AND DISCUSSION OF RESULTS:

   a) Simulated AE Studies

        The velocity of the first arrival wave packet (first negative peak) was measured to be 214,000 in/s. This value corresponds to the in-plane extensional mode [Gorman, 1990]. Using this velocity the expected differences in arrival times, delta T, for lead breaks at positions 1, 2 and 3 (see figure 1) in the notch were calculated. Table 4 shows these theory results as well as the experimental values based on lead breaks for one array of two sensors (indicated in figure 1). The mean delta T values as well as the maximum and minimum values for four to six lead breaks are given as a function of the approximate difference measured in decibels between the peak of the AE signal and the threshold, delta dB. There are three main observations to be made from this data. First, the difference between the experimental values and the calculated value tends to increase as delta dB decreases. This means that the effective propagation velocity is decreasing as the fixed threshold, which detects the arrival of the AE signal, moves further into the analog signal. Second, the maximum and minimum values of delta T also show a wider range as delta dB decreases. This result implies that there is more variability in exactly which positive or negative half-cycle crosses the AE system threshold. Third, the mean value for the symmetric (with respect to both specimen geometry and sensor locations) lead break position (position 1) does not change much with decreasing delta dB. Results for the other sensor arrays were similar except that nearer positions gave delta T values closer to calculated

values for a given delta dB, and as already mentioned edge sensor positions experienced a 4–6 dB higher amplitude signal than surface mounted sensors.

The above results are clearly elucidated by the AE waveforms for lead break signals. Figure 2 shows typical waveforms at two sensors for a non symmetrical lead break. Since the AE system has a fixed threshold that triggers either on positive or negative signals, the AE system is triggered not on the first deviation of the signal from the background noise but instead at some point within the first wave packet, as the threshold is raised. Since for a non symmetric lead break location the AE signals at the different sensors are not in phase, and have different amplitudes and envelopes, differences between the measured delta T values and the calculated theoretical values arise. For this experiment if the two channels trigger out of phase by one-half cycle, the error is about 1.5 us and a full cycle gives an error of about 3 us. These errors are much larger than the ± one clock-tick (0.25 us) error of the system. Since the individual cycles of the waveform both rise and fall within the first packet, even greater errors can occur, leading to differences greater than 3 us between calculated and experimental values. Since a symmetric lead break generates in-phase and similar amplitudes at the sensors, the experimental values for delta T tend to be close to the calculated values, but the values can still vary by up to a cycle due to differences in sensitivity in the two AE channels. One other important observation from the transient recorder records can be seen in table 5. Here the calculated theoretical delta T value is compared with the delta T values which can be obtained from the two channel transient recorder by making in phase measurements of the delta T value for a non symmetric lead break. Again, it is seen that the effective propagation velocity increases even for inphase measurements as the triggering point is moved into the waveform.

The implications of the above results for the crack-tip location phase of this research are several-fold. First, only at delta dB's (between peak and threshold) greater than say 18 dB can we expect the AE measurement system to give reasonable values for delta T values. Even in these cases sensitivity differences between channels or the known lack of AE source radiation symmetry [Scruby, 1985] can lead to errors on the order of 1.5 us to 3 us or more. These errors when translated to distance can lead to

significant deviations in crack tip location considering the 0.214 in/us velocity. Second, the delta T values may need correction by use of a variable velocity for different delta dB values. Third, transient recorder waveforms gathered in the crack tip experiments can be expected to provide more reasonable results because in-phase delta T values can be determined. But, again an error in the crack tip location will result unless a variable velocity of propagation is used. The above results are all related to the fact that the peak amplitude of an AE signal is about 40-55 dB greater than the peak amplitude of the first half cycle deviation from background noise in the current experiments. This fact means that for all of the AE signals from the crack tip sources, waveforms will have the early part of the AE signal buried in the preamplifier noise because the peak amplitudes are not large enough. Hence the majority of delta T's which are determined will come from points somewhere within the AE waveform and not from the first arrival of the signal.

b) Center Cracked Samples With Load Cycling

The crack tip AE source, which will generate AE at peak loads or during an overload, is expected to be inclusion particle fracture in the plastic zone in 2024-T251 aluminum [Scala and Cousland, 1983]. Crack growth during fatigue is not expected to be a source at the sensitivities used in these experiments because the micro processes involved take too long to generate significant AE. Since relatively few events occur at the crack tip during fatigue cycling to a fixed level [McBride and Maclachan, 1982, Scala and Cousland, 1985], overload AE was studied to accomplish the crack tip location objective of this work.

Standard AE location algorithms for a four sensor rectangular array can be used to examine the capability of AE technology to distinguish the three different crack lengths used in these experiments. But, due to the errors in delta T values which are expected as a result of out of phase arrival times, such an approach is not expected to lead to reliable results. Figure 3 shows the result is as expected. Hence, the approach was changed to analyze the data in a more meaningful fashion. Using event listings, hand calculations were made of delta T values that would locate a crack tip event. The procedure was to first use the sensors symmetrically located above and below each crack tip to calculate delta T values. For crack tip

events these values should be nearly zero assuming symmetrical source radiation. Next the delta T value related to the location along the crack length was determined using the two sensors above (below) the crack. Table 6 gives the raw data from an event listing, and the sample calculations for an event with a relatively large delta dB between the threshold and the peak amplitude. Table 7 shows this result as well as data more representative of the fact that most of the crack tip events had delta dB values considerably smaller than the sample event in table 6.

The relative ability to locate crack tips by use of AE data analyzed with the above procedure is shown in table 7. This table gives delta T values based on the following approaches: i) calculated theoretical values using the measured extensional velocity; ii) measured values using surface lead breaks at the crack tips for two different thresholds; iii) AE system measurements (analyzed as in table 6) from overload events; and iv) measurements taken from transient recorder records (10 MHz) from events during overloads (table 7b). If the delta dB level is large enough, the three different lengths can be distinguished by the commercial AE system data. But, since in general the delta T values are too large, the crack length values calculated from these delta T's would tend to be larger than the actual ones. There is an additional very significant result to be gained from this table. The generally wide deviations from the expected value of near zero for delta T_{18-13} and delta T_{19-15} is not expected based on the wave propagation results. The likely explanation is the previously noted lack of uniform radiation of sound energy from the AE sources. This condition implies an additional source of error for fixed threshold approaches.

On the other hand, table 7b shows that the transient recorder measured delta T values can be much closer to the calculated theory values. The values listed in this table were determined by measuring the delta T between the first positive or negative peak (positive peaks in both channels or vice versa) of the AE signal above the noise. As can be seen in the table, there are some difficulties with the current transient recorder approach as well. The main problem is that it is difficult to determine which are the two inphase peaks from the two AE sensors. Since the real AE signals have nonsymmetrical radiation of energy, the determina-

tion is not trivial, and the above approach may not have determined the inphase points. Some preliminary data, based on analysis of a few hard copy waveforms from the long crack case, indicates that in some cases a small precursor wave can lead to errors in determining the inphase arrivals using the above approach. The limited studies here did not fully identify a standard approach to this inphase determination. It may be easier to use a different type of AE sensor that will simplify the problem [Gorman, 1990]. But such an approach must maintain sensitivity and applicability to real structures.

In addition to AE from the crack tip at peak loads, a second period of AE occurred on ꞈch fatigue cycle. Due to the low R ratio (R = 0.01) crack closure was present for the lower loads of each cycle. Crack closure did not generate significant AE at the sensitivities used. But, the "unclosure" deformation generated significant amounts of AE for each fatigue cycle. The source mechanism of this crack opening AE has not been a source of detailed study [Guozhi, 1985]. A potential source is the relatively brittle fracture of the cold welds, formed when the two crack surfaces come together as the load is reduced in the fatigue cycle. Table 8 gives the approximate number of "unclosure" events for each crack length as a function of rest time at minimum load. The approximate stress intensity range over which this AE was generated is also listed in this table. A number of significant observations can be made. First, the number of "unclosure" events increases significantly with increased rest time. Second, the number of "unclosure" events changes small amounts (for a fixed rest time) with increasing length of precrack (total crack length minus starter notch length). Another related observation is that the number of crack tip events per cycle was relatively few compared to the number of "unclosure" events. Typically the number of crack tip events was zero to three events per cycle. Attempts to apply the AE system source location approach used for crack tip events (see table 6) were plagued by the same errors in delta T values which are present in the crack tip events. Further, since the location of these "unclosure" events was not known, there is no standard to compare the results as was done for the crack tip events. Using transient recorder (10 MHz) records, delta T values were determined from in phase positions (assumed to be first same sign cycle

peaks) for the "unclosure" events. Table 9 shows these results for the long and short cracks along with the approximate load at which the events occurred. The results are not easy to interpret. In some cases the region of the starter notch corresponds to the delta T value. It may be that more sophisticated approaches are needed to find the true inphase positions. The limited number of waveforms examined here did not lead to such an approach. This potential unzipping of crack closure seems to represent an additional opportunity for crack detection and length characterization. There are two reasons why this source of AE should not be neglected. First, this source is much more numerous than crack tip events and thus provides more data. Second, this source is much more likely to continue to operate under spectrum loading which occurs in real applications. Some key questions still to be answered about this source include, over what length of the crack does it occur? Does this source occur in very ductile materials which have been observed to not emit AE with crack growth? What are the angular radiation properties of this source? And what exactly is the source mechanism?

V. STATUS OF CRACK-TIP LOCATION BY AE:

This research shows clearly that AE technology as practiced with commercial instrumentation currently available cannot accurately determine the crack length even under relatively ideal conditions (small sample, laboratory environment, and ideal sensor locations). The key factor that leads to errors is the measurement of AE signal arrival times by a fixed threshold.

The primary problem is the measurement technique and not the basic AE waveform data which can provide reasonable results. Using displacement type point contact sensors which allow identification of specific wave modes it has been shown that techniques using waveform recorders can quite accurately locate crack tips in relatively uniform laboratory size samples [Scruby, 1985]. Unfortunately, these sensors do not have high sensitivity. The approach of using waveform recorders has some difficulties. First, waveform recorders are not designed to handle the event rates in typical AE testing. If applied to AE instrumentation, current technology can overcome this difficulty. Second, low amplitude AE sources can sometimes

pose difficulties in finding the inphase arrivals at different sensors due to lack of symmetry in energy radiation from an acoustic source.

VI. CONCLUSIONS:

1) A center crack specimen, screw machine, and hydraulic grips leads to elimination of extraneous noise in fatigue crack propagation.

2) Commercial AE systems with fixed thresholds cannot in general accurately locate crack tips for most of the crack tip events in this case.

3) Use of waveform recorders to determine the arrival time with the same phase at each sensor can give reasonably accurate crack tip locations.

4) "Unclosure" AE dominates crack tip generated AE.

5) "Unclosure" AE may offer an additional approach to crack detection and length characterization.

VII. RECOMMENDATIONS:

1) Develop algorithms to rapidly locate the earliest possible in phase arrival times of digital AE waveforms for an array of AE sensors even in the presence of precursor waves or low signal to noise ratios.

2) Examine the use of ultrasonic type sensors or other sensors to make such an algorithm simpler [Gorman, 1990]

3) Study the waveform approach for larger test samples. First, for uniform thickness samples with a range of thicknesses and later for samples with nonuniform thicknesses and other appropriate geometry. Determine the size of areas which can be effectively monitored. Real AE sources may be necessary since non-symmetry of source radiation must be included in the study.

4) Investigate in more detail the "unclosure" AE source in both aluminum and more ductile materials. Also include in the study spectrum fatigue at realistic loading rates.

5) Document the sensitivity improvements available for AE sensors relative to the amplitude of AE sources of interest.

REFERENCES

Damage Tolerant Design Handbook, Vol. 3, MCIC-HB-OIR, Metals and Ceramics Information Center, Battelle Columbus Laboratories, Columbus Ohio, 1983, p. 7.5-1, 7.5-111.

Gorman, M. R., Plate Wave Acoustic Emission, Submitted for publication to the Journal of the Acoustical Society of America, 1990.

Guozhi, Lu, Fatigue Crack Closure Study, Journal of Acoustic Emission, Vol. 2, No. 2/3, 1985, pp. S203-S206.

Hamstad, M. A., R.D. Young and P. M. Thompson, Acoustic Emission Source Location During Four-Point Bend Tests in Alumina, Progress in Acoustic Emission III, Japanese Society of NDI, 1986, pp. 26-33.

McBride, S. L. and J. W. Maclachan, Acoustic Emission Due to Crack Growth, Crack Face Rubbing, and Structural Noise in the CC-130 Hercules Aircraft, J. of Acoustic Emission, 3 January 1984, pp. 1-10.

Scala, C. M. and S. McK. Cousland, Acoustic Emission during Fatigue Crack Propagation in Aluminum Alloys 2024 and 2124, Material Science and Engineering, Vol. 161, 1983, pp. 211-218.

Scala, C. M. and S. McK. Cousland, Acoustic Emission During Fatigue of Aluminum Alloy 2024: The Effect of an Overload, Materials Science and Engineering, Vol. 76, 1985, pp. 83-88.

Scruby, C. B., Quantitative Acoustic Emission Techniques, in Research Techniques in Nondestructive Testing, Vol. VIII, Edited by R. S. Sharpe, Academic Press, London, 1985, pp. 141-210.

Sendeckyj, G. P. and C. A. Paul, Some Smart Structures Concepts, Fiber Optic Smart Structures and Skins II, SPIE Proceedings, Vol. 1170, 1989, pp. 2-10.

Table 1   Extraneous AE during test with dummy sample (loading
between 200 lbs and 14,000 lbs).

                          Total Events

	Condition	Increasing Load	Decreasing Load
a.	Newly installed sample:		
	first cycle	12	9
	second cycle	4	2
	third cycle	6	2
b.	Sample reinstalled:		
	first cycle	19	19
	second cycle	2	9

Table 2 Conditions for center cracked specimens

Specimen	Crack length 2a (in)	Crosshead rate, in/min (cycle time, min)	Overload factor
Short Crack	0.741	0.005 (21)	1.18
Medium Crack	1.755	0.002 (18)	1.53
Long Crack	2.770	0.002 (7)	1.09

Table 3 Mechanical properties: 2024 – T351 aluminum.
        Yield strength: 54.5 ksi
        Plane strain fracture toughness: 35 ksi$\sqrt{in}$

Table 4 Typical lead break wave propagation times for edge
sensors at nominal 3 in. position (notched sample).

                Delta T, Mean and range (us)

Lead break
position and
theoretical
delta T, us

	52dB	30dB	18dB	6dB
1	-0.4	0.2	0.4	0.7
(0)	(0.3 to 0.5)	(0.2 to 0.3)	(0.3 to 0.6)	(0.2 to 2.5)
2	3.0	2.6	2.2	4.8
(2.4)	(2.6 to 4.0)	(2.4 to 2.7)	(1.0 to 2.8)	(−103 to 4.8)
3	5.2	5.2	4.4	28.5
(4.4)	(4.4 to 6.2)	(4.7 to 6.2)	(3.0 to 5.0)	(−96 to 28)

<u>Table 5</u> Change in delta T for inphase measurements for sur-
face lead breaks at crack tip (long crack sample) with edge sensors and
transient recorder measurement. Using times at peaks of half-cycle of
waveform.

                    delta T, us

First negative   8.8
First positive   9.3
Second negative   9.6
Second positive   9.9
Third negative   10.1
Third positive   10.5
Fourth negative   10.6
Fourth positive   10.8
Fifth negative   11.3
Fifth positive   10.4

        Theory (9.4 us)

<u>Table 6</u>   Sample calculation of delta T values using rectangular array of
           sensors, for an overload event (long crack sample with delta dB
           = 26dB)

                13 _____ 15
                18   19

Channel	Arrival time, us	delta T Calculation	Theory, absolute value, us
18	87.0	delta T_{18-13} = 87.0−89.0 = −2.0 us	(0)
13	89.0	delta T_{19-15} = 95.0−95.0 = 0 us	(0)
15	95.0	delta T_{13-15} = 89.0−95.0 = −6.0 us	(9.4)
19	95.0	delta T_{18-19} = 87.0−95.0 = −8.0 us	(9.4)

<u>Table 7a</u>   Some typical delta T values for overload events using data from
           AE measurement system (*largest event)

Crack length	Delta T, us				Delta dB
	18−13	19−15	13−15	18−19	
Short with	−1.5	2.6	7.3	3.2	2
edge sensors,	4.0	−1.2	2.0	7.2	4
theory delta T	0.6	−3.2	−6.0	−2.2	8
=2.8 us	1.0	−2.0	−4.6	−1.6	12
	−2.2	3.0	5.2	0.0	16
	0	0	−2.7	−2.7	18*
Tip lead break	2.0	0.6	3.0	4.4	14
Tip lead break	2.0	0.4	−3.0	−1.4	15
Tip lead break	0.3	0.6	3.0	2.7	45
Tip lead break	1.8	0.2	−4.6	−3.0	45

Medium with	8.0	0.4	7.0	14.6	3
top sensors,	0.2	3.2	-4.0	-7.0	8
theory delta T	-1.0	0.7	-7.7	-9.4	14
-5.5 us	10.0	0.5	6.0	15.5	19
	2.0	0.5	6.0	7.5	24
	0.0	0.6	-7.2	-7.8	24
	0.2	0.3	7.3	7.2	36*
Tip lead break	1.4	0.4	6.0	7.0	48
Tip lead break	0.8	0.4	-7.6	-7.2	50
Long with	5.0	0	12.2	7.2	6
edge sensors,	37.8	5.2	-11.6	21.0	9
theory delta T	5.7	5.5	9.8	10.0	11
-9.4 us	2.0	8.5	-8.5	-15.0	14
	1.3	0	-11.0	-9.7	15
	-2.0	0	-6.0	-8.0	26
	-2.4	0	12.0	9.8	30*
Tip lead break	1.3	0	-11.0	-9.7	15
Tip lead break	1.4	-0.3	10.3	12.0	15
Tip lead break	0	0	-10.2	-10.2	46
Tip lead break	-0.2	-0.8	9.8	10.8	44

Table 7b   Typical delta T values from transient recorder for overload events

Crack length	Delta T, us 15-13	Peak amplitude* mv
Short with	-9.5	35 to 46
edge sensors,	-3.1	66 to 70
theory delta T	-2.6	65 to 81
-2.8 us	2.0	162
	-4.0	185 to 193
	2.2	188 to 229
	-2.8	185 to 234
Tip lead break	-2.7	2,700
Tip lead break	2.8	2,600
Long with		
edge sensors	-9.1	33 to 39
theory delta T	-6.0	165 to 170
= 9.4 us	-2.0	46 to 50
	10.1	170 to 240
Tip lead break	-8.5	2,800
Tip lead break	9.0	2,900

*Preamplifier noise typically 16-20 mv pp.

Table 8   Typical number of crack opening ("unclosure") events and
           approximate load range of AE source operation versus rest time.

           Number of events/Rest time (min)

           Load range
           K_I, Ksi $\sqrt{in}$

Short Crack   0.7 to 2.8   86/1.4   161/58.6
Medium Crack   0.9 to 3.0   82/0.8   170/63   397/988
Long Crack   1.0 to 3.6   60/1.1   105/110.5

Table 9   Delta T location information for crack opening
           ("unclosure") events from transient recorder*

	Delta T 13-15 us	Peak amplitude, mv	Load, lbs
Short Crack	-3.8	82 to 99	695
	1.8	72 to 86	800
	-2.4	103 to 120	900
	-2.0	112 to 141	1210
	-2.2	188 to 190	1360
	2.0	205 to 229	1720
	1.8	196 to 239	2360
	2.2	286 to 327	2580
Long Crack	-1.2	30 to 60	200
	2.0	128 to 226	315
	-2.4	186 to 219	500
	-0.8	84 to 94	850
	-0.9	>320, saturated	1105
	-2.1	176 to 199	1325
	-2.4	275 to 309	1435
	1.8	>320, saturated	2210

* Theoretical delta T values 1.8, 2.8, 9.4 us for starter
  notch, short crack tip, and long crack tip respectively
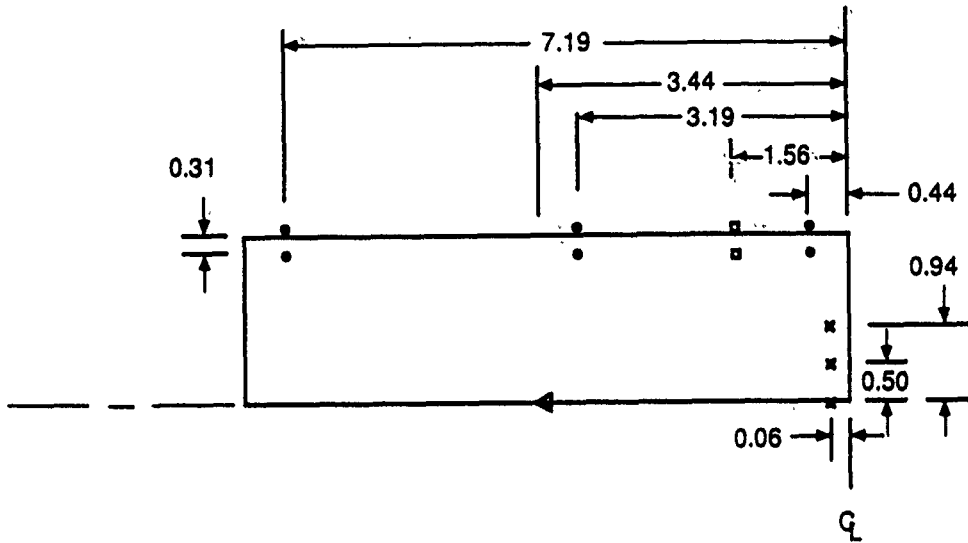
Figure 1   Location of centers of sensors and notch leadbreak positions, symmetrical about lines through notch or crack and perpendicular to crack. Key: x Lead break positions in notch, 0 propagation study sensor position, □ test sensor positions, ◁ guard sensor positions, and ₵ centerline of crack. Dimensions in inches, not to scale, one quadrant of sample.

Ch 1

Ch 2
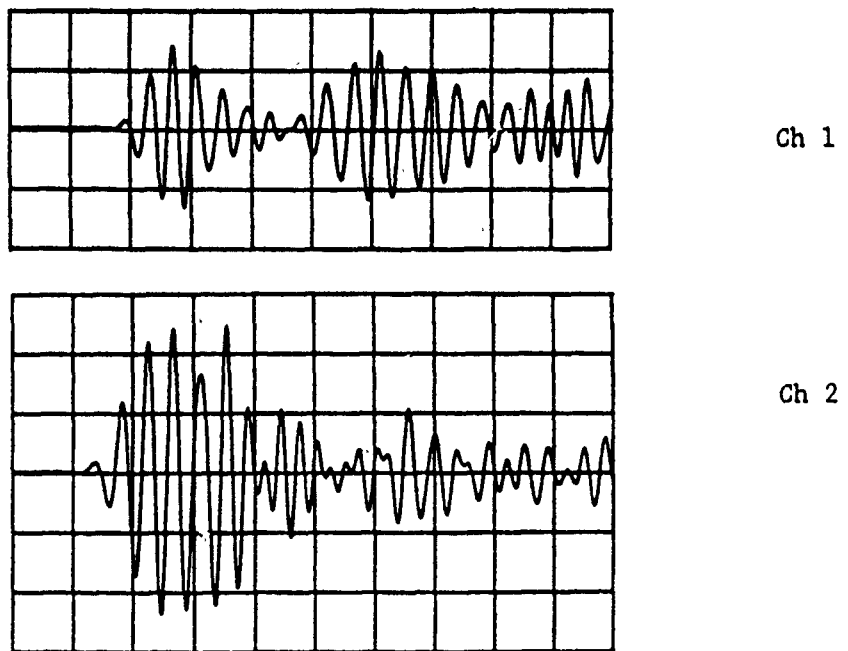
Figure 2   Typical waveforms for lead break at crack tip of long crack sample with edge sensors (Channel 1 near sensor, Channel 2 far sensor, at 40 dB gain, vertical scale 1v/div, horizontal scale 10us/div)

100-19
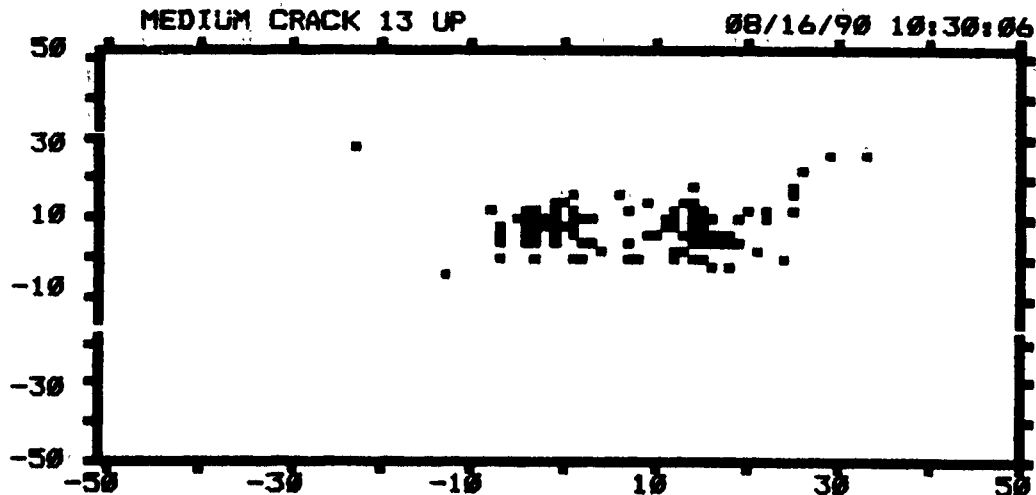
Figure 3 AE measurement system location of crack tip events from overload cycle for standard system algorithm for rectangular array on medium crack sample. Key: Event locations ■ ; coordinates of sensors 0,0; 13,0; 0,14; 13,14; units of axis in us.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

H_∞ Design Based on Loop Transfer Recovery and Loop Shaping

Prepared by:	Chin S. Hsu, Ph.D.	Jenny L. Rawson
Academic Rank:	Associate Professor	Graduate Student
Department and	Department of Electrical Engineering and Computer Science	
University:	Washington State University	
Research Location:	Flight Dynamics Laboratory WRDC/FIGC Wright-Patterson AFB, OH 45433	
USAF Researcher:	Siva S. Banda, Ph.D.	
Date:	17 Aug 90	
Contract No:	F49620-88-C-0053	

H_∞ Design Based on Loop Transfer Recovery and Loop Shaping

by

Chin S. Hsu
Jenny L. Rawson

ABSTRACT

This report addresses the issue of H_∞ loop transfer recovery and loop shaping when an H_∞ output feedback controller is used. A method of selecting the H_∞ design parameters to achieve asymptotic loop transfer recovery is presented. It is shown that the problem of approximate loop transfer recovery is equivalent to that of H_∞ state feedback design. A new H_∞ design procedure is also presented.

Acknowledgements

We wish to thank the Air Force Systems Command and the Air Force
Office of Scientific Research for sponsorship of this research.
Universal Energy Systems must be mentioned for their concern and help to
us in all administrative and directional aspects of this program.

Our experience was rewarding and enriching. John Bowlus provided
us with support, encouragement, and a truly enjoyable working
atmosphere. The help of Dr. Siva Banda and Dr. Hsi-Han Yeh was
invaluable for furnishing us with much-needed technical guidance. Their
concern and keen interest in our research progress was the key factor
which rendered our research effort into a pleasant and fruitful task.
We also wish to thank Capt. Sharon Heise for helping to make our work
easier and fun.

I. Introduction:

Robust multivariable control has been a main thrust of modern feedback design. Practical control systems must be designed to cope with various system uncertainties. One particular concern of robustness is the stability measure in terms of the gain and phase margins which are determined by the system's loop transfer functions. An effective approach to design a controller for robust control is the Linear Quadratic Gaussian/Loop Transfer Recovery (LQG/LTR) method. The LQG/LTR output feedback design provides a means of recovering the guaranteed margins which are an inherent feature of state feedback design via the Linear Quadratic Regulator (LQR).

The most recent robust control design is based on the H_∞ optimization method. The H_∞ method ensures that the closed-loop system is internally stable and the H_∞ norm of the closed-loop transfer function matrix is less than a prespecified scalar index. Just like the LQG/LTR method, suboptimal H_∞ controllers can be obtained by solving two Riccati equations.

A significant extension of the H_∞ design has been made at WRDC/FIGC. This is a powerful and new control design procedure using mixed H_2 and H_∞ optimization.

My research interests have been in the area of linear control systems with emphasis on developing numerical algorithms for industrial applications. My previous research on model reduction, system identification and optimal control, which serves as a general background

for robust control theory, contributed to my assignment to the Control Dynamics Branch, Flight Control Division.

II. OBJECTIVES OF THE RESEARCH EFFORT:

The objective of this summer research is to explore new design methodologies in robust multivariable control. Motivated by the new H_2/H_∞ controller design procedure, we focused our efforts on extending H_∞ design by incorporating loop transfer recovery and loop shaping. The goal of developing this new design approach, H_∞/LTR, is to provide control engineers with a procedure for designing practical robust control laws for aerospace applications in the presence of unwanted external disturbance.

Owing to the facts that suboptimal H_∞ controllers are observer-based and obtained from the solution of two Riccati equations, it is natural to investigate the possibility of designing H_∞ compensators using the combination of an H_∞-state feedback regulator and a full-order observer. The success of this new observer-based H_∞ design will yield an effective method of developing reduced-order H_∞ controllers with the use of reduced-order observers.

III. H_∞/LTR: A LOOP SHAPING METHOD FOR H_∞ OUTPUT FEEDBACK COMPENSATOR DESIGN:

1. Motivation.

State feedback with an H_∞ constraint as presented in [3] gives a static controller from the solution of a single Riccati equation. The objectives are to provide internal stability and to bound the H_∞ norm of the transfer function between a deterministic disturbance input and a

designated output. Additional benefits are a bound on the sensitivity function and guarantees on the gain and phase margins at the input to the plant that are even better than those given by linear quadratic regulators (LQR).

These benefits may not be maintained when output feedback is used. In the Linear Quadratic Gaussian/Loop Transfer Recovery method (LQG/LTR) [8, 9], an adjustable parameter in the filter Riccati equation is used to asymptotically recover the state feedback loop transfer function when output feedback is used. This allows the stability margins to be improved. In this paper, we investigate a method for asymptotic recovery of the H_∞ state feedback loop transfer function when H_∞ output feedback is used (H_∞/LTR). The goal is to regain the good properties of H_∞ state feedback.

One disadvantage of asymptotic LQG/LTR is that the feedback gains can become unacceptably high; the same problem can occur in H_∞/LTR. To avoid this, approximate H_∞/LTR can be used. This paper includes a method that is based on a Riccati equation which solves a dual state feedback problem.

In the following, section 2 includes a description of the feedback system, formal statements of the H_∞/LTR and loop shaping problems, and a review of the H_∞ state and output feedback methods. H_∞ asymptotic and approximate LTR methods are presented in section 3. In that section, it is also shown that some asymptotic loop shaping can be done in the state feedback case. Proofs for the theorems and lemmas can be found in [7].

Notation.

γ, δ, μ, ρ	positive numbers.
x, u, y, w, z, x_c	n, m, m, q, p, n-dimensional vectors.
A, B, C	$n \times n$, $n \times m$, $m \times n$ real-valued matrices.
D_1, D_2	$n \times q$, $m \times q$ real-valued matrices.
V_1, V_2	$D_1 D_1^T$, $D_2 D_2^T$.
$E_{1\infty}$, $E_{2\infty}$	$p \times n$, $p \times m$ real-valued matrices.
$R_{1\infty}$, $R_{2\infty}$	$E_{1\infty}^T E_{1\infty}$, $E_{2\infty}^T E_{2\infty}$.
Q, Y_∞	$n \times n$ symmetric, positive semi-definite matrices.
A_c, B_c, C_c	$n \times n$, $n \times m$, $m \times n$ real-valued matrices.
$G(s)$, $G_c(s)$	(A, B, C), (A_c, B_c, C_c).
$L_s(s)$, $L_o(s)$	(A, B, C_c), $G_c(s)G(s)$.
$R_s(s)$, $R_o(s)$	$I - L_s(s)$, $I - L_o(s)$.
$E(s)$	$L_s(s) - L_o(s)$.
$\tilde{A}$, $M(s)$	$\tilde{A} + \gamma^{-2}V_1 Y_\infty$, $(\tilde{A} - B_c C, B, C_c)$.

2. Preliminaries and Problem Statement.

The plant to be controlled is described by the following equations.

$$\dot{x}(t) = Ax(t) + Bu(t) + D_1 w(t), \quad x(0) = x_0.$$

$$y(t) = Cx(t) + D_2 w(t) \qquad\qquad (1)$$

$$z(t) = E_{1\infty}x(t) + E_{2\infty}u(t)$$

A goal of H_∞ design is to bound the H_∞ norm between the deterministic disturbance $w(t)$ and the output $z(t)$ while internally stabilizing the closed-loop system. To do this, the following assumptions will be made.

<u>Assumptions</u>:

1. (A, B) and (A, D_1) are stabilizable pairs;

2. (A, C) and $(A, E_{1\infty})$ are detectable pairs;

3. $D_1 D_2^T = 0$, $E_{1\infty}^T E_{2\infty} = 0$;

4. $V_2 = D_2 D_2^T > 0$, $R_{2\infty} = E_{2\infty}^T E_{2\infty} > 0$.

Two problems are addressed in this paper. The first is to design an output feedback compensator to recover the loop transfer function as designed in a state feedback problem. The second is to shape the state feedback loop transfer function. These are both stated formally below.

<u>Problem A</u>: H_∞ <u>Loop Transfer Recovery</u>.

Find design parameters D_1 and D_2 such that output feedback asymptotically gives the same loop transfer function at the input as does state feedback.

<u>Problem B</u>: H_∞ <u>Loop Shaping</u>.

Find design parameters $E_{1\infty}$ and $E_{2\infty}$ such that H_∞ state feedback gives a desired loop transfer function.

Solutions to these two problems are presented in Section 3. The remainder of this section is devoted to a review of the standard state and output feedback design methods.

<u>H_∞ State Feedback</u>. [1, 3]

In this case, $C = I$ and $D_2 = 0$. Static state feedback is used, with one Riccati equation to be solved. This equation is:

$$0 = Y_\infty A + A^T Y_\infty + Y_\infty(\gamma^{-2} V_1 - B R_{2\infty}^{-1} B^T) Y_\infty + R_{1\infty}. \qquad (2)$$

The control law is then $u(t) = C_c x(t)$ where $C_c = -R_{2\infty}^{-1} B^T Y_\infty$. Let $H_s(s)$ be the transfer function between w and z. The closed-loop system is

stable and $\|H_s(s)\|_\infty < \gamma$ iff $Y_\infty \geq 0$ and $A + (\gamma^{-2}V_1 - BR_{2\infty}^{-1}B^T)Y_{i_0}$ is stable.

Transfer functions of interest are the loop transfer function $L_s(s) = C_c\Phi(s)B$ and the return difference matrix $R_s(s) = I - L_s(s)$, where $\Phi(s) = (sI - A)^{-1}$. It can be shown that

$$R_s^T(-j\omega)R_{2\infty}R_s(j\omega) = R_{2\infty} + B^T(-j\omega - A^T)^{-1}R_{1\infty}(j\omega - A)^{-1}B$$
$$+ \gamma^{-2}B^T(-j\omega - A^T)^{-1}Y_\infty V_1 Y_\infty(j\omega - A)^{-1}B \quad (3)$$
$$\geq R_{2\infty}.$$

This can be used to determine guaranteed gain and phase margins much as is done with LQR [4, 8, 9]. Here there is the added benefit of an extra positive semi-definite term on the right hand side of equation (3) to give even better margins than those of LQR [2].

<u>H_∞ Output Feedback</u>. [1, 6]

A dynamic compensator is used and two Riccati equations must be solved when the states are not available. The design equations are:

$$0 = Y_\infty A + A^T Y_\infty + Y_\infty(\gamma^{-2}V_1 - BR_{2\infty}^{-1}B^T)Y_\infty + R_{1\infty}$$
$$0 = AQ + QA^T + Q(\gamma^{-2}R_{1\infty} - C^T V_2^{-1}C)Q + V_1 \quad (4)$$

And the compensator is:

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t)$$
$$u(t) = C_c x_c(t) \quad (5)$$

where,

$$A_c = A - B_c C + BC_c + \gamma^{-2}V_1 Y_\infty$$
$$B_c = (I - \gamma^{-2}QY_\infty)^{-1}QC^T V_2^{-1} \quad (6)$$
$$C_c = -R_{2\infty}^{-1}B^T Y_\infty.$$

Note that C_c is the same as for state feedback. Figure 1 is a block diagram of the closed-loop system.
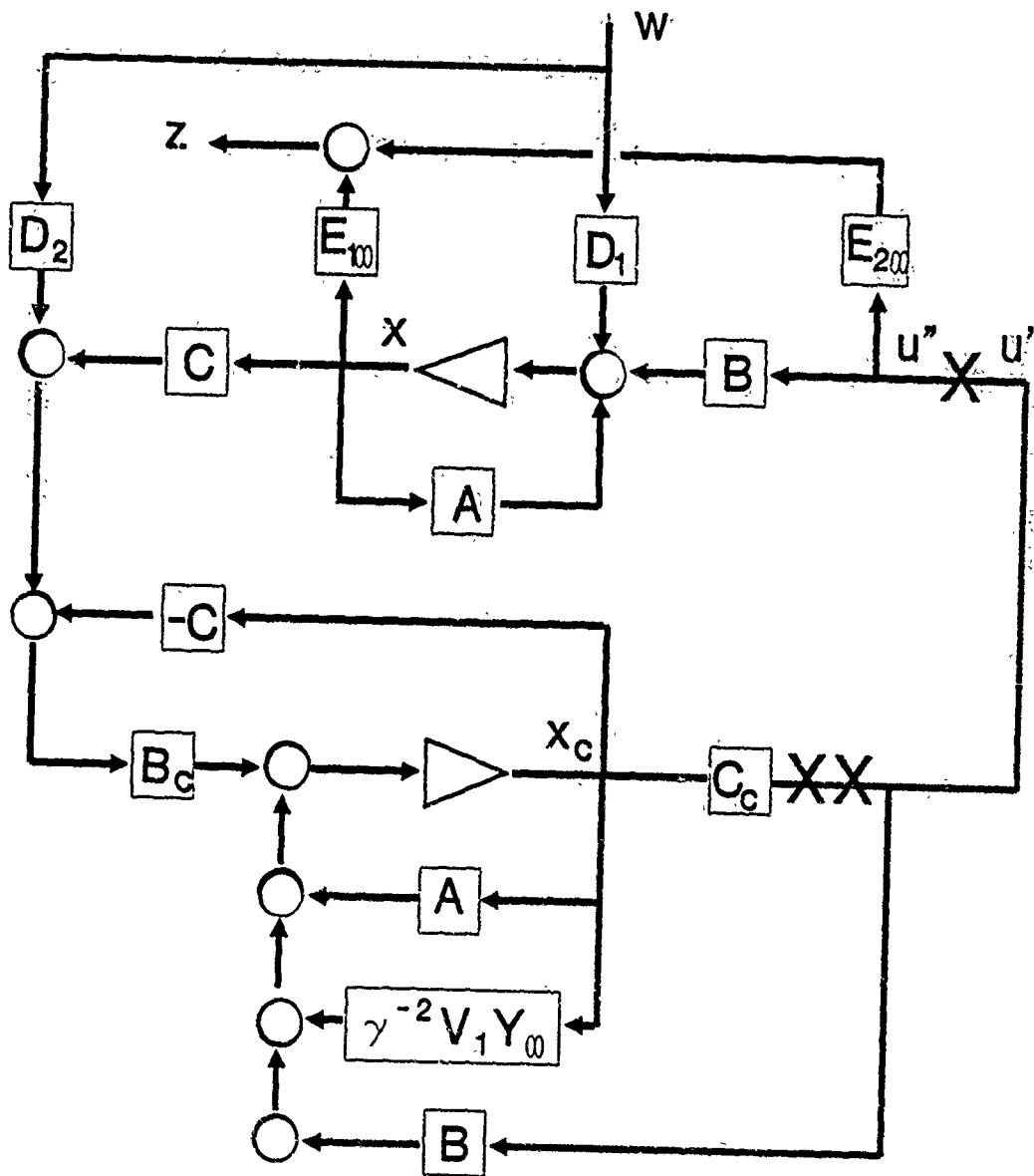
Figure 1.   Closed-Loop Control System with H_∞ Compensator.

Let $H_0(s)$ be the transfer function between w and z. There will be a stabilizing solution such that $\|H_0(s)\|_\infty < \gamma$ iff Q, $Y_\infty \geq 0$, $\rho(QY_\infty) \leq \gamma^2$, ($\rho(\cdot)$ denoting the spectral radius of a matrix), and both $A + (\gamma^{-2}V_1 - BR_{2\infty}^{-1}B^T)Y_\infty$ and $A + Q(\gamma^{-2}R_{1\infty} - C^TV_2^{-1}C)$ are stable.

As will be shown in the next section, there may not be any guarantees on the gain and phase margins at the input to the plant.

3. H_∞/LTR Design Methodology

3.1 LTR Conditions

Before presenting the actual design theory, we would like to examine some conditions for exact LTR. This is done in the next lemma and two theorems.

<u>Lemma 1</u>: Define $\Phi_c(s) = (sI - A_c)^{-1}$, $L_0(s) = C_c\Phi_c(s)B_cC\Phi(s)B$ (the loop transfer function from u″ to u′) and $R_0(s) = I - L_0(s)$. Then,

$$R_0^T(-j\omega)R_{2\infty}R_0(j\omega) = R_s^T(-j\omega)R_{2\infty}R_s(j\omega) - L_s^T(-j\omega)R_{2\infty}L_s(j\omega)$$
$$+ L_0^T(-j\omega)R_{2\infty}L_0(j\omega) \qquad (7).$$

<u>Proof</u>. Follows by using the definitions of $R_s(s)$, $R_0(s)$, $L_s(s)$ and $L_0(s)$. ∎

This lemma suggests that H_∞ output feedback may or may not give phase and gain margins which are as good as those guaranteed for state feedback. If they are not, then the use of LTR is indicated.

<u>Theorem 1</u>: (H_∞/LTR Condition)

With $V_1 = D_1D_1^T = BVB^T$, $V > 0$, then $L_0(s) = L_s(s)$ iff

$$B_c[I + C\Phi(s)B_c]^{-1}C\Phi(s)B = B. \qquad (8)$$

∎

<u>Note 1</u>:

$$B_c[I + C\Phi(s)B_c]^{-1}C\Phi(s)B = B \Leftrightarrow [I + C\Phi(s)B_c]^{-1}B = 0 \Leftrightarrow B = 0 \ .$$

This indicates that B_c will have to be chosen carefully in order for exact recovery to occur.

<u>Theorem 2</u>: (Mismatch Condition)

Again let $V_1 = D_1D_1{}^T = BVB^T$, $V > 0$. Define $E(s) = L_s(s) - L_o(s)$ and $M(s) = C_c(sI - A + B_cC - \gamma^{-2}BVB^TY_\infty)^{-1}B$. Then

$$E(s) = M(s)[I - M(s)]^{-1}[I - (C_c + \gamma^{-2}VB^TY_\infty)\Phi(s)B], \qquad (9)$$

and

$$E(s) = 0 \text{ iff } M(s) = 0.$$

∎

<u>Note 2</u>: $M(s)$ is the transfer function of the compensator with the loop broken at **XX** in figure 1.

<u>Note 3</u>: The above lemma and two theorems are equivalent.

The following lemma gives an expression for the sensitivity function at the plant input, and shows that with LTR, the sensitivity for output feedback approaches that for state feedback. If $R_{2_\infty} = \mu^2 I$, then equation (3) can be used to show that $\|S_s(s)\|_\infty \leq 1$. Thus, the upper bound on $\|S_o(s)\|_\infty$ can be made arbitrarily close to 1 by decreasing ρ.

<u>Lemma 2</u>: Let $S_o(s) = R_o^{-1}(s)$, $S_s(s) = R_s^{-1}(s)$. Then,

$$S_o(s) = [R_s(s) - \gamma^{-2}M(s)VB^TY_\infty\Phi(s)B]^{-1}[I - M(s)]. \qquad (10)$$

And, if $M(s) = 0$, then

$$S_o(s) = S_s(s).$$

The proof requires simple manipulations. ∎

3.2 LTR Method

The weights for LTR are $V_1 = D_1 D_1^T = BVB^T$, $V > 0$, and $D_2 = \rho \tilde{D}_2$, $\rho > 0$ where we require that $\tilde{D}_2 \tilde{D}_2^T > 0$ and $D_1 \tilde{D}_2^T = 0$. The design equations become:

$$0 = Y_\infty A + A^T Y_\infty + Y_\infty (\gamma^{-2} BVB^T - BR_{2\infty}^{-1} B^T) Y_\infty + R_{1\infty} \tag{11}$$

$$0 = AQ + QA^T + Q(\gamma^{-2} R_{1\infty} - \rho^{-2} C^T (\tilde{D}_2 \tilde{D}_2^T)^{-1} C) Q + BVB^T \tag{12}$$

And the compensator is:

$$A_c = A - B_c C + BC_c + \gamma^{-2} V_1 Y_\infty$$
$$B_c = (I - \gamma^{-2} Q Y_\infty)^{-1} \rho^{-2} Q C^T (\tilde{D}_2 \tilde{D}_2^T)^{-1} \tag{13}$$
$$C_c = -R_{2\infty}^{-1} B^T Y_\infty.$$

The next theorem shows that the state feedback loop transfer function is recovered as $\rho \to 0$.

<u>Theorem 3</u>: If $G(s) = C\Phi(s)B$ and $G_c(s) = C_c \Phi_c(s) B_c$, then

$$\lim_{\rho \to 0} G_c(s) = C_c \Phi(s) BG^{-1}(s)$$

That is,

$$\lim_{\rho \to 0} L_0(s) = \lim_{\rho \to 0} G_c(s)G(s) = C_c \Phi(s) B = L_s(s). \qquad \blacksquare$$

<u>Note 4</u>: If for state feedback $\|H_s(s)\|_\infty < \gamma$, then there exists a $\rho_0 > 0$ such that with output feedback $\|H_0(s)\|_\infty < \gamma$ for all $0 < \rho < \rho_0$.

3.3 Approximate LTR

Asymptotic LTR is an iterative procedure, with adjustments of the design parameter ρ alternating with solutions of the Riccati equations and checks on the stability margins. To avoid this, an approximate LTR procedure can be used instead, where the norm of the error $E(s)$ is constrained by an upper bound.

Theorem 4:

a)   Approximate LTR can be achieved if for a frequency band of interest, $\bar{\sigma}(E(j\omega)) < \alpha$, $\alpha > 0$. This is equivalent to requiring that $\sup\bar{\sigma}(M(j\omega)) < \delta$ for some $\delta > 0$, where δ is a function of α and $\bar{\sigma}(I-(C_c+\gamma^{-2}VB^TY_\infty)\Phi(j\omega)B)$.

b)   Define $\tilde{A} = A + \gamma^{-2}V_1Y_\infty$. If $(\tilde{A}^T - C^TCX, B^T)$ is a detectable pair and there exists a solution $X \geq 0$ to the Riccati equation

$$0 = \tilde{A}X + X\tilde{A}^T + X(\delta^{-2}C_c{}^TC_c - C^TC)X + BB^T, \tag{14}$$

then, $\sup\bar{\sigma}(M(j\omega)) = \|M(s)\|_\infty < \delta.$ ∎

Note 5: Approximate LTR is obtained by solving a second H_∞ state feedback problem.

3.4 Asymptotic Loop Shaping with State Feedback

In this section, we present a method for shaping the loop transfer function when state feedback is used.

Suppose that the desired loop transfer function is $-C_D(sI - A)^{-1}B$. To select the state feedback gain such that the actual loop transfer function shape approaches this one and the closed-loop transfer function satisfies the H_∞ constraint, use the following weights:

$E_{1_\infty} = C_D$, (A, C_D) a detectable pair;

$E_{2_\infty} = \mu\tilde{E}_{2_\infty}$, $\tilde{E}_{2_\infty}{}^T\tilde{E}_{2_\infty} = I$, $C_D{}^T\tilde{E}_{2_\infty} = 0$;

$D_1D_1{}^T = BVB^T$, $V > 0$.

Then, the design equation becomes:

$$0 = Y_\infty A + A^TY_\infty + Y_\infty(\gamma^{-2}BVB^T - \mu^{-2}BB^T)Y_\infty + C_D{}^TC_D \tag{15}$$

And the feedback gain is $C_c = - \mu^{-2}B^TY_\infty.$

<u>Theorem 5</u>: If $\gamma^2/\mu^2 > \rho(V)$, then there is a solution $Y_\infty \geq 0$ to (15). Furthermore,

$$\lim_{\mu \to 0} \mu C_c(sI - A)^{-1}B = -C_D(sI - A)^{-1}B.$$

■

3.5 Discussion

The H_∞/LTR procedure is analogous to LQG/LTR. In fact, it can easily be shown that equations (8), (11), (12) and (13) reduce to those for LQG/LTR [8] as $\gamma \to \infty$. In addition, (2) and (8) are identical in form to the "Kalman Inequality" and "Doyle-Stein Condition" respectively [9].

As with LQG/LTR, there is a dual procedure for H_∞/LTR at the output of the plant. The H_∞/LTR method can be extended to minimum phase, non-square plants if the number of outputs exceeds the number of inputs for recovery at the plant input, or the reverse for recovery at the output.

Note that this procedure differs from McFarlane and Glover's H_∞ loop shaping [5] where the singular value plots of the open loop transfer function are shaped.

3.6 Design Procedure.

The algorithm for loop shaping and H_∞/LTR is shown in the flow chart of Figure 2.
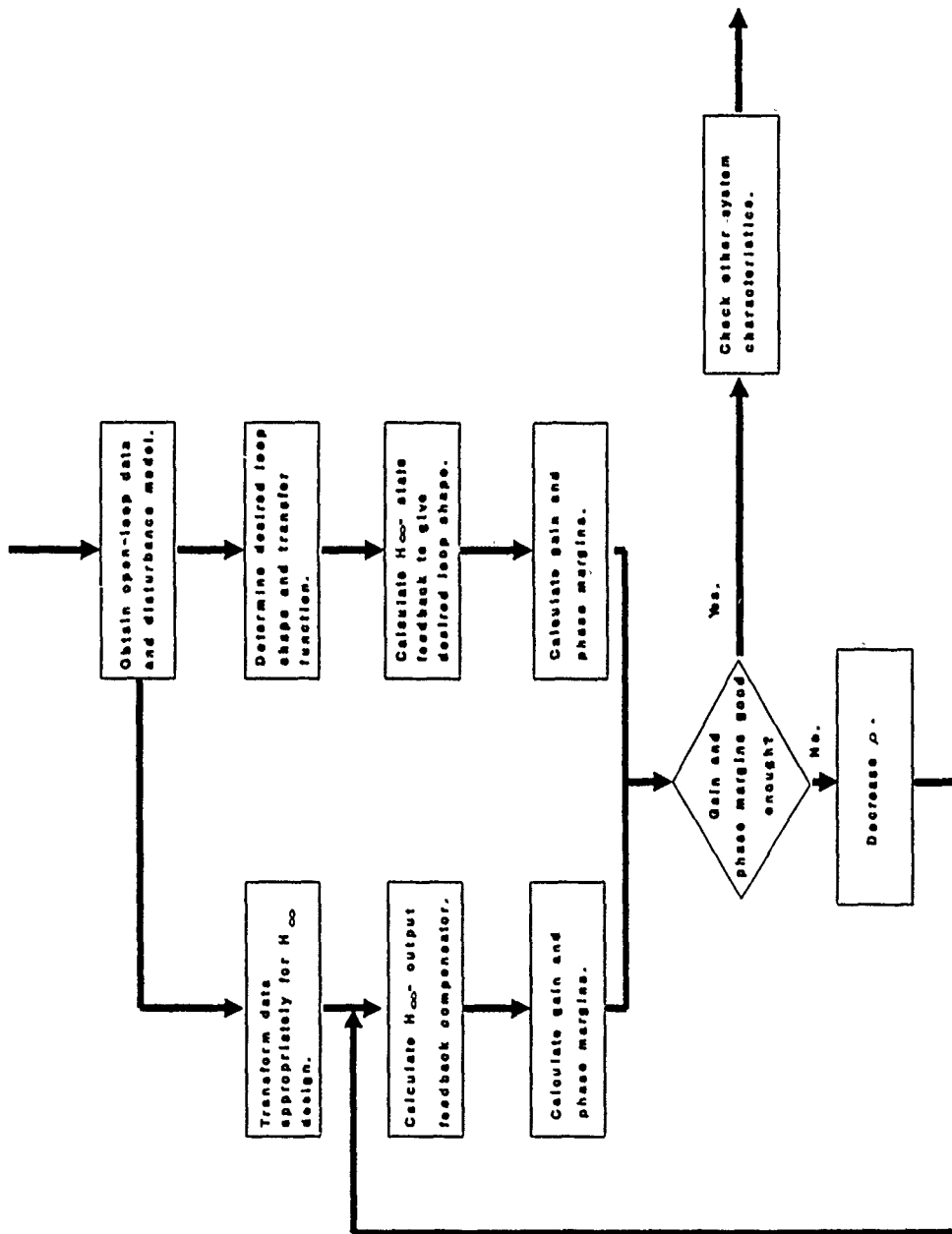
Figure 2. H_∞/LTR Design with Loop Shaping.

IV. H_∞ DESIGN--A NEW APPROACH:

It is known that H_∞ suboptimal controllers can be obtained by solving two Riccati equations and that the compensator is observer-based [1, 9]. In this section, we propose a direct approach to H_∞ design using observer theory. This new approach has the advantage of generating reduced-order controllers.

Consider the following linear system:

$\dot{x}(t) = Ax(t) + B_1w(t) + B_2u(t)$, $x(0) = x_0$.

$z(t) = C_1x(t) + D_{11}w(t) + D_{12}u(t)$ \hfill (16)

$y(t) = C_2x(t) + D_{21}w(t)$

Lemma 3: The closed-loop transfer function matrix (TFM) from w to z is

$$T_{SF}(s) = D_{11} + (C_1 - D_{12}K)(sI - A + B_2K)^{-1}B_1 \hfill (17)$$

if state feedback is used; i.e. $y(t) = x(t)$ and $u(t) = -Kx(t)$. ∎

Theorem 6: Suppose that an observer, with observer gain L, is used with (16). It is of the form:

$$\dot{x}_c(t) = (A - LC_2)x_c(t) + Ly(t) + B_2u(t) \hfill (18)$$

$$u(t) = -Kx_c(t)$$

then closed-loop TFM from w to z is

$$T(s) = T_{SF}(s) + F(s)(sI - A + LC_2)^{-1}(B_1 - LD_{21})$$

where,

$$F(s) = D_{12}K + (C_1 - D_{12}K)(sI - A + B_2K)^{-1}B_2K \hfill (19)$$

∎

Based on the foregoing result, an H_∞ design procedure is as follows:

1.   Determine a state feedback gain K such that $\|T_{SF}(s)\|_\infty < \gamma$. This can be done by solving one Riccati equation [4].

2.   Determine an observer gain L such that $\|T(s) - T_{SF}(s)\|_\infty < \delta$.

Note 6: The procedure in step 2 is dual to an H_∞-state feedback problem (i.e., H_∞-filtering), and as such, a Riccati equation can be found to determine the observer gain L. Progress is being made to demonstrate this design procedure.

Note 7: If, instead of a full-order observer (18), a reduced-order observer is used, then the above design procedure will result in a reduced-order H_∞ controller.

V.   RECOMMENDATIONS:

a)   The straight application of H_∞ design may be of limited use. It is more advisable to use the proposed method delineated in this report for developing robust contol laws since the H_∞/LTR design parameters can now be properly selected to achieve approximate loop transfer recovery in conjunction with loop shaping. The H_∞/LTR design can be considered as a subproblem of H_2/H_∞/LTR which is currently being developed at WRDC/FIGCA. Choice of control design methodologies is obviously dependent on the performance specifications of the control system. It is believed that there are distinct advantages and disadvantages for each design method encompassing LQG/LTR, H_∞/LTR and H_2/H_∞/LTR among others.

b)   It is generally accepted that low-order controllers are preferred to high-order controllers, given comparable performance. The H_∞ design procedures presently being used give a compensator of the same dimension as the given plant. The issue of controller reduction is of paramount importance to be addressed so that the H_∞ design and its H_2/H_∞ extensions can be fully fruitful in aerospace applications. As envisaged in this report, the significance of the H_∞ compensator being an observer-based controller cannot be overemphasized. Firstly, a reduced-order controller design can be developed using the available theory pertaining to reduced-order observers. Secondly, doubly coprime factorization (DCF) is a proven tool to represent observer-based compensators. Thus, the controller reduction method based on DCF should be fully exploited to develop reduced-order H_∞ suboptimal controllers. The preliminary results as presented in this report pave the way to continued research in this area. It is also of practical importance to evaluate the merit of other controller reduction methods in the framework of H_∞ and H_∞/LTR output compensator design.

c)   The mixed H_2/H_∞ compensator design requires the solution of three modified Riccati equations, two of which are tightly coupled. There are needs for effective numerical algorithms for solving these equations and for understanding their qualitative properties. It has been found in this summer reserch that under certain conditions the H_2/H_∞ design requires only two coupled Riccati equations. It is suggested that these equations be used as a starting point for studying the qualitative behavior of the three Riccati equations of general H_2/H_∞ design.

REFERENCES

1. Doyle, J. C., K. Glover, P. P. Khargonekar and B. A. Francis, "State-space solutions to standard H_2 and H_∞ control problems," IEEE Trans. Automat. Control vol AC-34 (1989) 831-847.

2. Fujita, M., K. Uchida and F. Matsumura, "Gain perturbation tolerance in H_∞ state feedback control," Int. J. Control vol 51 (1990) 315-328.

3. Khargonekar, P. P., I. R. Petersen and M. A. Rotea, "H_∞-optimal control with state feedback," IEEE Trans. Automat. Control vol AC-33 (1988) 786-788.

4. Maciejowski, J. M., Multivariable Feedback Design (Addison-Wesley Publishing Company, Wokingham, England, 1989).

5. McFarlane, D. C. and K. Glover, Robust Controller Design Using Normalized Coprime Factor Plant Descriptions (Springer-Verlag, Berlin, 1990).

6. Mustafa, D., "H_∞-characteristic value .," Proc. 28th Conf. Decision and Control, Tampa, FL (1989) 1483-1487.

7. Rawson, J. L., C. S. Hsu, H. H. Yeh and S. S. Banda, "H_∞/LTR: A loop shaping method for H_∞ output feedback compensator design," submitted for publication.

8. Ridgely, D. B. and S. S. Banda, Introduction to Robust Multivariable Control, Technical Report: AFWAL-TR-85-3102 (Flight Dynamics Laboratory, Air Force Wright Aeronautical Laboratory, AFSC, USAF, 1985).

9. Stein, G. and M. Athans, "The LQG/LTR procedure for multivariable feedback control design," IEEE Trans. Automat. Control vol AC-32 (1987) 105-114.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems,Inc.

FINAL REPORT

A FEASIBILITY STUDY ON INTERFACING ASTROS WITH NAVGRAPH

Prepared by: Ming-Shu Hsu, Ph.D.

Academic Rank: Associate Professor

Department and   Mechanical Engineering Department
University: University of Portland

Research Location: WRDC/FIBRA
                          Wright-Patterson AFB
                          Dayton, OH 45433

USAF Researcher: V. B. Venkayya, Ph.D.

Date: July 20, 1990

Contract No: F49620-88-C-0053

A FEASIBILITY STUDY ON INTERFACING ASTROS WITH NAVGRAPH

by

Ming-Shu Hsu

ABSTRACT

A computer program referred to as ASTROS (Automated STRuctural Optimization System) was developed under contract by the Flight Dynamics Laboratory at Wright-Patterson AFB, Ohio. ASTROS employs the well-known "Automated Design Synthesis" (ADS) procedure and optimality criteria methods, in addition to the finite element analysis, to provide an optimal design for interdisciplinary applications. Since its first introduction in 1987, ASTROS has received great response and a growing user population, and has set a revolutionary milestone in the field of aerospace structural analysis and design. However, the lack of a pre and post processor makes it inconvenient in preparing the model input data and in interpreting the results. This project investigated the feasibility of interfacing ASTROS with NAVGRAPH which is a general purpose geometry modeling and mesh generation computer graphics package. Three phases of development were recommended for the short and long term goals.

ACKNOWLEDGEMENTS

I. **INTRODUCTION**

Computer-Aided Design (CAD) systems are currently used in virtually all sectors of engineering design practices. This has improved the design quality and shortened the development time. The conventional CAD system takes the finite element model as input data and produces the analysis results. It is the designer's resposibility to analyze the results to make sure that the design criteria are satisfied, otherwise a redesign procedure will be repeated. In the field of aerospace structural design, because of the nature of its complexity, it is very difficult, if not impossible, to achieve an optimal design without the aid of an automated procedure.

To improve the design procedure, a new computer program referred to as ASTROS (Automated STRuctural Optimization System) has been developed under contract by the Flight Dynamics Laboratory at Wright-Patterson Air Force Base, Ohio. ASTROS employs the well-known "Automated Design Synthesis" (ADS) procedure and optimality criteria methods, in addition to the conventional finite element analysis, to provide an optimal design for interdisciplinary applications.

Since its first introduction in 1987, ASTROS has received great response and a growing user population, and has set a revolutionary milestone in the field of aerospace structural analysis and design. However, the lack of a pre and post processor makes it inconvenient in preparing the model input data and in interpreting the results. It is desired to interface ASTROS with a computer graphics package in order to enhance its pre and post processing.

My research interests have been in the area of integration of automated design and manufacturing systems. In 1989, I established a Computer Integrated Manufacturing (CIM) program in the School of Engineering at the University of Portland. The objective of this program is to study interfacing a variety of hardware and software in an interdisciplinary environment. My background in engineering (e.g. finite element, optimization) and computer science (e.g. computer graphics, data base, network), and my experience in systems integration (e.g. developing a postprocessor for a CAM system) contributed to my assignment for this project.

II. OBJECTIVE OF THE RESEARCH EFFORT

Currently there is no pre and post processor directly available to ASTROS. Indirectly, because the finite element analysis module contained in ASTROS is mostly the same as NASTRAN, the ASTROS user may take NASTRAN input bulk data which is created from any NASTRAN preprocessor and with a minor modification to create an input file for ASTROS. On the other hand, the ASTROS user may extract results from the ASTROS database and create an OUTPUT2 file through a Translator. The OUTPUT2 file, in turn, can be used by a NASTRAN postprocessor to generate graphical displays of results.

There are many computer graphics packages capable of serving as a pre and post processor for ASTROS. Among them, the Structural Analysis Group in the Flight Dynamics Laboratory (FIBRA) has preliminary chosen three: PATRAN, I-DEAS, and NAVGRAPH. The interface of the first two packages with ASTROS and the development of the Translator mentioned above are being develoed under a contract with ASIAC.

The third package, NAVGRAPH, developed under contract by the NAVY, is a general purpose finite element/finite difference pre and post processing software package that combines geometry modeling and mesh generation with computer graphics.

The objective of my assignment as a participant in the 1990 Summer Faculty Research Program (SFRP), therefore, was to investigate the feasibility of interfacing NAVGRAPH with ASTROS as a pre and postprocessor.

III. **EVALUATIONS**

Prior to this study, many engineers in the NAVY and the AIR FORCE had experience using NAVGRAPH for pre and post processing of both small and large structures models. A collection of their comments, suggestions, and the responses from B.Y.U.(NAVGRAPH developer) are attached in the APPENDIX for reference.

In my evaluation of NAVGRAPH, graphical capabilities, command funtionalities, and user friendliness were evaluated from the user's point of view while the degree of easiness or difficulty on the interface development was emphasized from the programmer's view. In the following comments and suggestions, items a, b, and c relate to the former aspect while items d, e, f, and g relate to the latter.

a. There are three documents supplied by B.Y.U. - User's Manual, Programmer's Manual, and Tutorial Manual. The Programmer's Manual was recently revised to include Database and Geometry sections. Both Programmer's and Tutorial Manuals are helpful to the user, however, the User's Manual lacks in providing an overall guidance for using the software package. It is recommended to enhance the User's Manual with a

Procedure Guidance which will list a sequence of steps (commands) to be followed for pre and post processing.

b. NAVGRAPH provides the great flexibility of using Global commands which can be selected by a key word from any other menu. In contrast, all the other commands can only be selected from the present menu. This causes great inefficiency when the pointer needs to be moved from one branch to another to select a command. Since all commands are built in a hierarchical structure, it is suggested to establish "hierarchical" Global commands at different levels.

c. On the X-Windows version of NAVGRAPH, commands may be selected by the mouse, but the data still must be typed in from keyboard. It may be more convenient for the user if the data, either ·digital numbers or characters, could also be input from the mouse.

d. As mentioned in Section II, ASTROS may be interfaced with a pre and post processor via an OUTPUT2 file. Generating the OUTPUT2 file, therefore, is a key element in the interface development. To understand the format used in the OUTPUT2 file a FORTRAN program has been written to read the binary OUTPUT2 file and to print the data blocks in ASCII form.

e. The Reverse Formatting module of NAVGRAPH can only process the following seven data blocks from the OUTPUT2 file — EXEQIN, GPDT, GEOM2, OUGV1, OEF1, OES1, AND OES1A. These seven data blocks are used for postprocessing geometry and F.E. models, and static analysis results. The postprocessors for dynamic analysis, aeroelasticity, and optimization are not available at this time.

f. NAVGRAPH consists of about 2000 routines written in the FORTRAN language and employs a dynamic memory allocation database developed by Sandia National Laboratory. One

difficulty the user may encounter in reading these routines is that the variable names are inconsistant throughout the program. No other major obstacle has been found for interface development.

g. In general, NAVGRAPH has the capability to perform the postprocessing ASTROS needs in both the analysis and design (optimization) modules. The interface development for the analysis module may be accomplished by some modifications in the routines, but it may require extensive changes in the program for the design module.

IV. **RECOMMENDATIONS**

The development of interfacing ASTROS with NAVGRAPH may be carried out in three phases. The first two phases provide a short term goal while the third phase is considered as a long term project.

a. Phase I:
The objectives of this phase are two fold: i) to provide a quick solution to the immediate need of postprocessing the two basic types of analysis - Staics and Dynamics. ii) to use Dynamic Analysis as a testbed to study the NAVGRAPH program for the next phase development.

By adding a Translator between ASTROS and NAVGRAPH as shown in Figure 1, the two databases of these two programs remain unchanged. This approach provides a quick solution to interface with minimun changes in both programs. The detailed work of this phase may include:
   1. To develop a "Translator" which can extract data from the ASTROS database and generate an OUTPUT2 file.
   2. To understand NAVGRAPH thoroughly, especially in the

data base management.

3. To modify "Reverse Formatting" and "Animation" modules in NAVGRAPH.

b. Phase II:

The objective of this phase is to make NAVGRAPH a complete pre and postprocessor for ASTROS by adding the capability of processing data from the Aeroelasticity and Optimization modules. Conceptually, this phase may look like an extension of Phase I by only adding more capabilities to the Translator and postprocessor. However, the complexity of modification is much greater than Phase I. It is not clear at this point that the OUTPUT2 file approach should be kept or be eliminated. At the completion of this phase, ASTROS and NAVGRAPH will still remain as two separate programs as shown in Figures 1 and 2. The detailed work may include:

1. To modify NAVGRAPH's preprocessor so that it can generate input data in ASTROS's format.
2. To resolve the approach to be taken for this phase development.
3. To expand the Translator to include output from the Aeroelasticity and Optimization modules, if necessary.
4. To study the graphical needs of these two modules.
5. To understand the ASTROS data base structure.
6. To include output from the Aeroelasticity and Optimization modules in NAVGRAPH's postprocessor.

c. Phase III:

The objective of this phase is to merge ASTROS and NAVGRAPH into one single package. The data files transfer between them, then, will be transparent to the user. The merging of these two programs requires the elimination of one of the two databases as shown in Figure 3. Since ASTROS employs a superior structured data base, it is recommended to revise NAVGRAPH to adopt ASTROS's database. This phase requires an

extensive effort to study the methodology of converting from one database to another, and to develop the code. The detailed work involved in this phase is beyond the scope of this project and cannot be outlined here.

V. **PROPOSED SCHEDULE**

```
                    Oct.        June   Aug.          June   Aug.
                    1990        1991   1991          1992   1992
                      |           |      |             |      |
                      |           |      |             |      |
Phase I               |   RIP     |      |             |      |
     Development      |-----------|      |             |      |
                      |           |      |             |      |
Phase II              |           | SFRP |             |      |
   Feasibility Study  |           |------|    RIP      |      |
   Optimization       |           |      |-------------| SFRP |
   Aeroelasticity     |           |      |             |------|
                      |           |      |             |      |
Phase III             |           |   | SBIR           |      |
   Feasibility Study  |           |   |------|    SBIR  |      |
   Development        |           |   |      |----- ⌇⌇ ----|   |
                      |           |   |      |             |   |
                      |           |   |      |             |   |
                              Aug.      Jan.          Aug.
                              1991      1992          1994
```
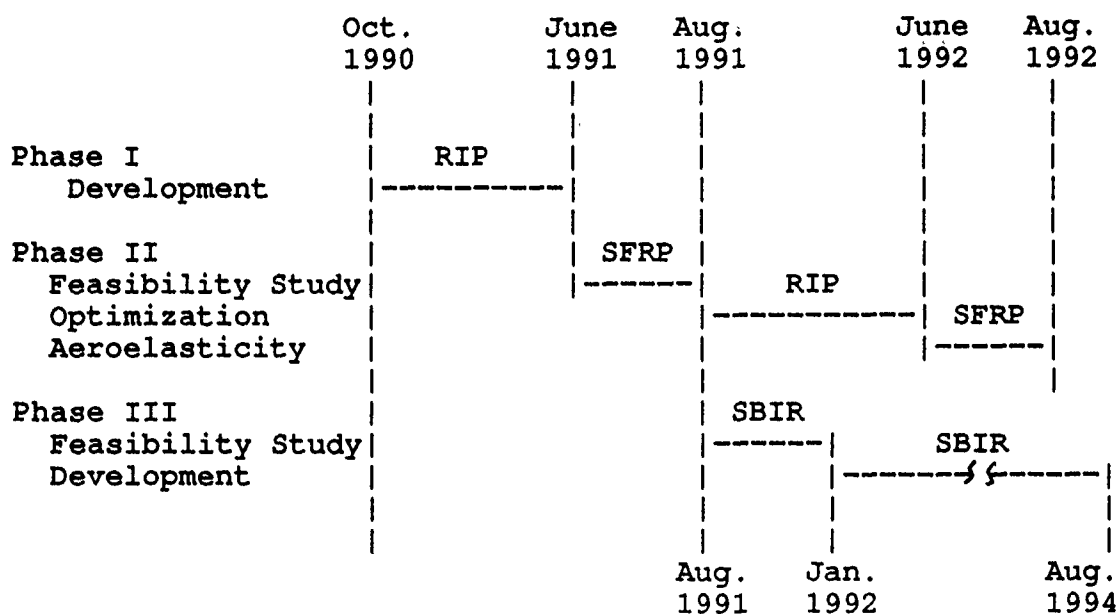
     RIP: Research Initiation Program
    SFRP: Summer Faculty Research Program
    SBIR: Small Business Innovation Research
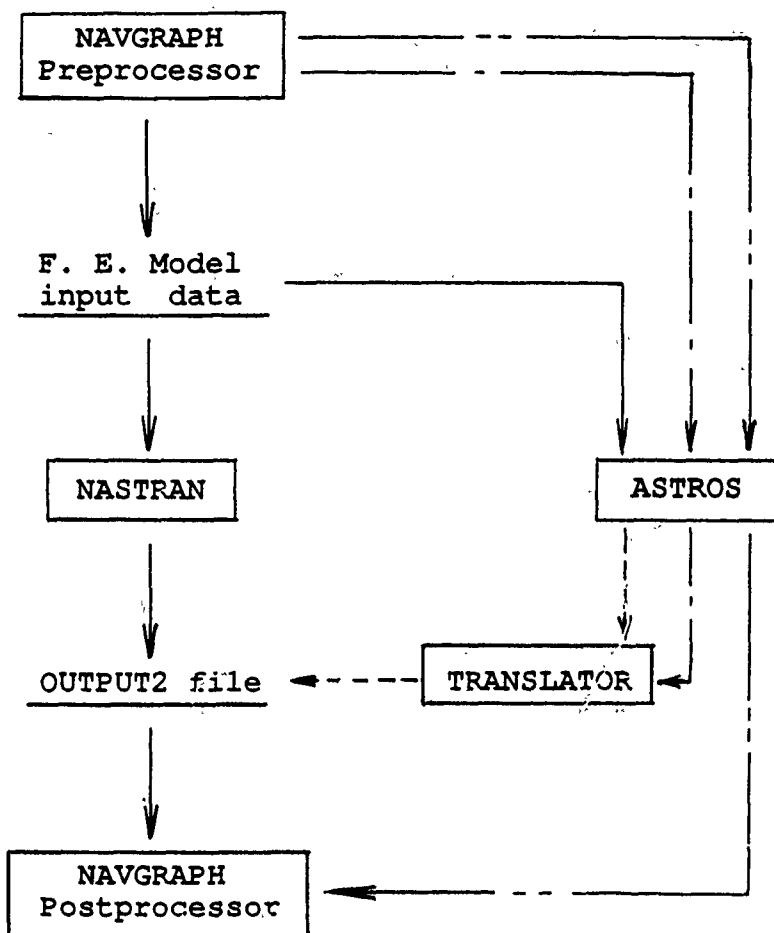
Figure 1. Three phases of interface development
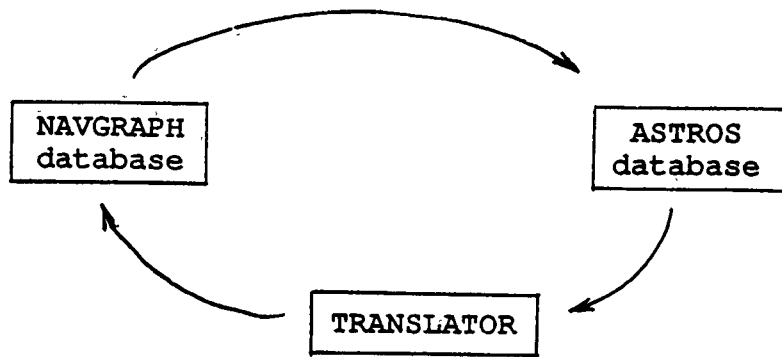
Figure 2.   Phase II development
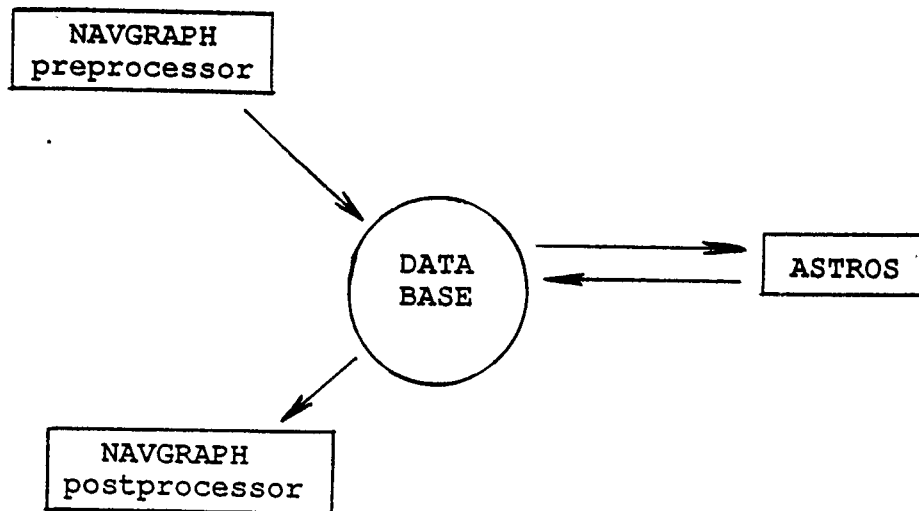
Figure 3.   Phase III development

APPENDIX

NAVGRAPH EVALUATION

The majority of these comments have been based on my reading through the entire user manual as well as the examples manual. I also spent some time working with the program on the VAX. Due to the fact that I've only had a few weeks to work with NAVGRAPH, some of these comments may actually be unnecessary and are simply things that I have missed. Please feel free to call me for any clarification or further discussion.

Matt Mells
(316) 433-3322
FTS 297-3322

1. Calling out colors by name or number would be easier than having to figure out RGB combinations to color the different items plotted to the screen.

2. The ability to color specified ranges of or individual items of an item as opposed to the entire range.
   Eg. elements 1 thru 20 color red
       elements 25 thru 40 color blue

3. Similar to #2 above with the GLOB DISP ENTI ( ). It would be nice to display entities by a specified ID or range of ID numbers. Eg. plotting elements 1 thru 100 with a single command.

4. From what I've read in the GLOB GROUP ADD options, I can accomplish what was mentioned in 2 & 3 above, however, it would really be helpful to be able to add or erase by applying specified ranges within the list.

5. I'm not totally clear on the differences between remove, erase, and delete.

6. When entering limits for the various options, it would be nice to be able to select any ID's or range of ID's desired as opposed to just upper and lower bounds. I would also prefer to have the option of adding these numbers at the end of the command string as opposed to waiting for the prompts to come up. This applies to a lot of the options in NAVGRAPH that require data input.

7. Would like to see arrows instead of little x marks to indicate the parametric directions of the geometry. It makes it much less confusing with a 3-D model.

8. For the DEFI -- SWEEP commands, it would be nice to also incorporate the use of vectors, grids or nodes to define the rotational axes as opposed to just using lines at present.

9. I would like to see some additional line options:

   a) a fillet option between both lines and surfaces
   b) a line option to extend the end of a line
   c) a line option to draw a straight line tangent to two
      other lines (arcs)
   d) a line option to draw a tangent line between a point and
      another line (arc)

10. I would like to see the extrude option expanded to incorporate twist or rotation about the axes it is being extruded out of

11. When using SOLID REVERSE commands, it would be desireable to only reverse the solid and not the lines or surfaces used to create it. Reversing these as well can lead to a lot of confusion.

12. I would like to see an option which would create a solid normal to a surface by specifying that normal thickness.

13. Is it possible to make degenerate solids with the SOLID SURF command ? If not, it would be desireable.

14. Some kind of an indicator of which nodes are equivalenced is absolutely necessary. (color the nodes or circle them etc.) Also, it would be desireable to prompt the user if it is ok to delete the indicated nodes in case there is a tolerance problem which would ultimately ruin the model.

15. The overall menu system makes me slightly uncomfortable. I would like to enter the majority of the commands without actually having to step through the menu tree. This appears very tedious to me based on my familiarity with other pre-processors. I may be less critical of this as I become more adept at working the program. For example it would be a lot easier to type TURN ELEM OFF, or SET ELEM OFF, or better yet, ELEM OFF as opposed to GLOB DISP ENTI ELEM. The syntax is very unfriendly to me at this point and as I stated above, I don't like having to step through the menu levels.

MEMORANDUM:

From: E. Becker (NWC)
To: P. Matuis (DTRC)

subj: NAVGRAP. XOVIESTAR EVALUATION COMMENTS

1. Sorry this has taken so long to get to you, but it has been very busy around here for the last couple of months. I would like to respond to the comments made by Matt Melis first, and then get into the comments and suggestions made by some people here about the code.

2. I will address each of the comments made by Matt Melis as a separate line item. That way you can coordinate what I say to the list you sent me.

1. I agree! I just had to do this, and it isn't easy. It would be good to keep this in, but an easier way needs to be incorporated into the program. His suggestions of color identified by name or number is a common practice.

2. This would prove to be very useful and should not be that difficult to add to the existing code.

3. We've been saying this for at least a year, and maybe even 2 to 3. Why wasn't this ever done? I thought they listened to us better than this.

4. Again, this is a functionality comment, and would make it easier for the user, as well as more acceptable.

5. This can be explained easily. Erase takes entities off of the screen, but not out of the database. Delete takes it off of the screen and out of the database. Remove is for group entities.

6. When and where one enters the ID's for an entity is often code dependent. You can, for the most part, enter information in the code before it asks you for it.

7. I don't think that arrows would help this situation. Perhaps another way to do this needs to be found. One thing that could help the problem is less x-marks on surfaces and solids. You really only need to have three for a solid and two for a surface.

8. This is something for future upgrades to the code.

9. Future upgrade.

10. Future upgrade.

11. This is a recurring theme. Different levels of the geometry construction have always been to closely tied to lower level entities (for example, delete a solid and all the surfaces, lines, and points go with it).

12. This is already available in the extrude option for a solid construction.

13. Yes. Try two surfaces with a coincident line.

14. This is already in the code. Prompting the user should not be necessary.

15. Again, this is already possible. But, depending on how this is done, the code will leave you at different menu levels after a command has been entered. This raises the comment, once again, of global commands for all aspects of the code.

Now, for our comments and concerns about the code.

A. It seems to be in pretty good shape for the most part. The problem I mentioned on the phone is included. This happened after resequencing the model. Why are solid elements such a problem? or is it because I'm using degenerate elements in part of my model?

B. In my opinion, this code is close to being acceptable, but there are still some bugs to be worked out. Has BYU instituted code validation and testing procedures yet? We should not be the ones to do this all of the time.

C. Another person here has had the program crash on them often enough that they no longer want to use it. I don't blame them. They had this happen will plotting a hidden line plot and also while resequencing (I had the latter happen as well).

D. Why does the screen clear itself after a line merge option has been performed?

E. Very small circles and arcs are plotted on the screen as hexagons. This doesn't seem right, and is confusing for the user.

4. What is really needed at this point is a good introductory and tutorial course for the program. The Tutorial Manual is no longer current with the latest version of the code that I have. There are a lot of things about the code that are undocumented as far as I know. Will we be getting this available in the near future?

5. If you have any questions concerning these comments, please call me at 394-2070.

E. J. Becker
Eric J. Becker
Naval Surface Warfare Center
Reentry Systems Branch

1720.2:RWH
19 DEC 1989

MEMORANDUM: (PRELIMINARY)

FROM: Roger W. Hoffman, Composites Group
TO : Michael J. Chamistru, Performance Group

SUBJ: Preliminary Examination of NAVGRAPH Capabilities

The following list partially outlines capabilities which must be added to NAVGRAPH for successful implementation of the program:

1. The ability to create a backup file of the user's model while executing NAVGRAPH.

2. NAVGRAPH must automatically write a backup file when the program bombs or the system crashes.

3. The capability to show the distance between points and between nodes is necessary.

4. The LIST MATERIAL PROPERTIES command should be modified to output material properties assigned to the material numbers. The command currently lists material numbers and names assigned to those numbers.

5. The VIEW RESTORE command should reset the rotational transformation matrix. The command currently resets the viewing transformation matrix and should reset both matrices.

6. The capability to "break" lines, surfaces, and solids with respect to parametric values is necessary.

7. When deleting geometric entities, the prompts to delete lower "order" geometric entities should be segregated into separate commands. For example, the DELETE SURFACE command generates a prompt "Delete associated lines and Points ?". The prompt should be split into two prompts so the user can retain contruction lines and/or points.

8. The capability to globally delete entries should be added.

9. The PLOT command should be modified to plot a list of non-consecutively numbered entities.

10. The DELETE command should be modified to delete a range of entities and individual entities.

11. When applying boundary conditions, a range of nodes needs to be available for each D.O.F. specified.

Date: 12 Feb 90 08:50:00 EST
From: navgraph@wpafb-fdl.af.mil
Subject: Evaluation of Navgraph
To: "matula" (matula)

The following list outlines either difficulties that FIBRA had in exercising NAVGRAPH or desirable capabilities which should be added.

Difficulties:

FIBRA was not able to figure out how to define and display a GROUP or subset of a model either in the pre-processing or post-processing mode. Once a model has been generated one would like to define a subset of the model for display, particularly if a model is very detailed.

Desirable Capabilities:

To be able to display value plots of information in the database. For example, the x-coordinate of all or a subset of the nodes displayed at the nodes. In particular, for results data, to be able to display stress, displacement, or force data at selected nodes or elements.

Unable to query the results database, i.e., list stresses or nodal displacements while interactive or be able to select a subset for display only.

The Air Force, in many instances, receives NASTRAN models from contractors. It's essential that NAVGRAPH be able to generate a database from the NASTRAN data itself. FIBRA is aware that one can run NASTRAN, write an OUTPUT2 file containing arrays EQEXIN,GPDT, and GEOM2 and then do a reverse translation into NAVGRAPH to generate a database. However, this approach is very awkward.

WRDC/FIB has phased out many of it Tektronix terminals and is moving in the direction of Micro-Vaxes, Sun Workstations, IBM Workstations and Silicon Graphics terminals. Drivers for all these systems are needed.

Enhanced documentation:

Programmer's Manual

Several scenerios on the over all flow of NAVGRAPH for
Pre-processing only
and
Post-processing
should be available.

In general, WRDC/FIB is very pleased with NAVGRAPH. Several engineers in our division are either using NAVGRAPH or tracking its capabilities through our usage. Serious consideration is being given to using NAVGRAPH as a potential pre and post processor for ASTROS (A STructural Optimization System) being developed by WRDC/FIBR. Additional comments may be forthcoming as FIBR continues to use NAVGRAPH.

To: Petro Matula

From: Jim Nelson

Date: Mar 2, 1990

Subject: Response to Navy Evaluation of MOVIESTAR

This is a response to the letters of review which we have received to this point. I'll try and keep my comments to three basic categories: 1) Things which we have fixed already, 2) Things we view as future enhancements, and 3) Things which already exist in Navgraph, but can: be done in perhaps another way.

I have received four letters with comments/suggestions about the status of Navgraph. They are from Matt Mellis, Roger Hoffman, WRDC/FIBR (Vickie Tishler I believe), and Eric Becker. If I am missing one please let me know. I'll respond to each of their letters by item, so you may wish to have a copy of their letters handy. I'll also use a (1), (2), or (3) with additional comments where necessary.

Letter from Matt Mellis:

1 - (2)
2 - (1), (2), and (3) We have now implemented the ability to color finite elements by region. However this does not work for hidden line representations yet. We hope to be able to carry this idea through to hidden line and shaded drawings soon.
3 - (3) The user can control which regions are displayed, and which are not by using the GLOB DISP ENTI REGI command. By default regions are displayed after creation, however they can be turned off by the user. Using groups some control can be given over display only of a set of elements which do not make up an entire region. In addition the plot command can be used to plot elements 1-100. I don't understand what the problem was here?
4 - (2) We have added entry by range to many areas in Navgraph, and will continue to do so in areas such as this where it is advantageous.
5 - Eric Becker explained this well.
6 - Navgraph does let you put this at the end of the prompt string. You can enter a range along with a step value to step through the range (i.e. 1,10 2 from 1 to ten by increments of two), however only one range per command can be given. Perhaps in the future this will be modified.
7 - We don't intend to change this.
8 - (2)
9 - (2) An option to extend a line in it's tangential direction has already been implemented.
10 - (2)
11 - (2) I guess I'm not convinced that this is a very useful command anyway??
12 - (3) As Eric mentioned this can be done with the extrude command.
13 - (1) and (3).Yes, there were some problems with this for some degenerate cases.
14 - (1)
15 - In our opinion the majority of the commands in menus other than global do not need to be globally accessed, and therefore is not worth the cost of trying to define all possible combinations globally. Perhaps in cases such as delete this causes an inconvenience, however we feel the overall usability of the program is not severely hampered.

Eric Becker's additional comments.

3A - (1) We have worked at trying to make the resequencer dynamic however we do not understand the method used for storage (We are using the routines developed by Everstein out of Nastran), and therefore have been unsuccessful. We have improved it sc that approx 15000 quadrilateral element can be resequenced, and 1200 or so solid elements. We haven't determined why so many more queds can be resequenced, and ask that maybe we get a little more help from the people there who are experts with this resequencing code. We have gone through and

cleaned up all of the degenerate element creation problems, and don't think you will experience any more problems of this nature.
Yes, we try to do the best we can, however we appreciate and need the help of outside sources like yourselves as we cannot exhaustively test all options.

3B - (1) Many fixes to hidden line drawings have been implemented.
3C - (1) I couldn't stand that either!
3D -
3E - I'm not sure the source of this, but since we are trying to approximate circles/arcs with several straight line segments I can imagine difficulties can arise with small radii. We will look into this, however we do not see it as a serious limitation.

4 - I agree. Navgraph has a personality of it's own, and until people become familiar with there will probably always be some complaints and reluctance to switch from other familiar systems. We have just about completed a complete update and format enhancement of the User's Manual, and Tutorial. Besides the Database documentation which we left with you the last time we were there, we will have a theoretical manual for the geometric modeling portion of Navgraph. We intend to provide theoretical manuals on the meshing, and optimizing commands as soon as we can.

The letter from Roger Hoffman at DTRC's Structures Dept.

1 - (1) A save command has been added to the global options menu.
2 - It is our hope that journal files can be used to restore work lost in the event of a crash. Unfortunately this system is a little too sensitive to change, however we continue to try and work this out the best we can.
3 - (2) This will be implemented in the very near future, however will not be available on the tape I'm sending to Steve H. at this time.
4 - (3) We don't see what the problems are. Perhaps this is something that has been corrected since your latest version.
5 - (3) The View restore command restores the rotation/translation matrix to identity, and not the viewing matrix. The viewing matrix is set to default by changing the Look at position to 0,0,0 and the look from position at 0,0,1. Perhaps they should both be restored with the same command, but I'm sure there would be some who would not like that.
6 - (2) I see this being done soon.
7 - (2) We are currently working on this, it is a good idea.
8 - See 15 from Matt Mellis' letter.
9 - (3) This can be done with separate commands and turning auto-erase to off. Again it would be nice to allow the user to enter multiple ranges at any prompt, but would cost too much at this point to do it.
10 - (1),(2) We now keep the user in the delete command so he can do this, however he is prompted once for each range.
11 - (3) The nodes can be put into a group;

The letter from WRDC/FIBR (Vickie Tishler). I will number responses to her comments one per paragraph on her letter.

Difficulties:
1 - (3) See number 3 on Matt Mellis' letter above.

Desirable Capabilities:
1 - (1) We think this is a good idea and plan to implement it.
2 - (1) Another important feature which we have already implemented.
3 - (2) We really feel that with the database documentation and a little help from us your people can get this implemented much quicker than we could hope to.
4,5 - (1) We think you will be very pleased with the X-Windows version and it's portability to the various workstations listed.
6 - See number four from Eric Becker's comments.
7 - Thank you.

102-17

sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

conducted by

UNIVERSAL ENERGY SYSTEMS, INC

FINAL REPORT

THEORETICAL MODELING OF THE PERFORATION OF LAMINATED PLATES

BY RIGID PROJECTILES

Prepared by: David Hui

Academic Rank: Associate Professor

Department: Mechanical Engineering

University: University of New Orleans

Research Location: WRDC/FIVST

                         Flight Dynamics Laboratory

                         Wright Patterson Air Force Base, Ohio 45433

US Air Force Researchers: Greg Czarnecki, Effort Focal Point

Date: 9/30/1990

Work Unite No.   24020261

Contract No.   F49620-88-C-0053

THEORETICAL MODELING OF THE PERFORATION OF LAMINATED PLATES BY
RIGID PROJECTILES   by David Hui

1. INTRODUCTION

The present work is concerned with the penetration and perforation
of laminated plates made of composite materials. The velocity of the
projectile of interest generally lies within the ordnance range, that is,
from 300 m/sec to 1000 m/sec. In additon to the many complicated phenomena
which are associated with the perforation of isotropic-homogeneous metallic
plates, one would need to consider additional vastly different in laminated
targets and they can be grouped into three stages of failure: matrix
cracking, delamination and fiber breakage. None of these three stages
of perforation of composite materials is understood since one has to incorporate
the complex interaction between fiber and matrix and the wave propagation
and high strain-rate properties of composites are not well understood.
Further, due to the hugh amount of structural parameters (number of layers,
lamination sequence, plate thickness, size and shape of the projectile
and size and boundary condition of the plates) and material parameters of
the composite plate, it is expensive and practically impossible to provide
extensive experimental tests involving all these parameters. As of today,
existing theories for the perforation of composite plates are empirical
or semi-empirical in nature; thus, the theories are not predictive and
would require some measurements (for example, the weight of the plug, etc.)
The aim of the present work is to develope a predictive theory which is based
on the fundamental principles of mechanics.   The theory is necessarily
a preliminary one due to the enormous complexity of the perforation problem.

The theory being developed is based on the energy balance method. In order to assess the validity of the applicability of the "metallic" theories to composite material plates, it is desirable to summarize the existing metallic theories which are based on the energy concept. Of course, some of the observations and conclusions from the metallic theories remain valid for composite plates. For example, the type of perforation can be classified as essentially plugging or essentially petalling, or a combination of both types; sharp-nose projectile within the ordnance velocity range tend to favor petalling while blunt-nose projectile or thick plates favor plugging. The remaining part of this section is devoted to a brief survey of the metallic theories using energy method.

Starting from the classical work by Taylor in 1948 on the theory of the enlargement of a circular hole in a thin plastic sheet, Thomson (1955) was the first who presented a quasi-dynamical theory for the petalling perforation of plates for rigid projectiles with conical or ogival head, using the energy balance method. Thomson's theory contained an algebraic error and it was later corrected by Sodha and Jain (1958) who also presented the residual velocity expression. Further, Johnson et al. (1973) examined the same problem but included the dynamic effects. Using the results of the bending of triangular cantilever beam under a tip concentrated force (Parkes 1955, 1958 and Johnson 1972), Landkof and Goldsmith (1985) found that the initial stage involve the formation of the triangular cantelever beams due to crack initiation, followed by plastic hinge motion and petal bending and they presented the residual velocity. Finally, Woodward (1978) extended Thomson's work to include the work done in bending and he was able to explain the transition from dishing to ductile hole formation as plate thickness increases.

Recht and Ipson (1963) were the first to present a theory of plugging for metallic plates using the energy concepts. The energy balance equation is, assuming a rigid projectile,

$$(1/2)m_p v_p^2 = E_{DH} + E_s^* + (1/2) (m_p + m^*) (v_r)^2 \qquad (1)$$

where m_p and m^* are mass of the projectile and plug respectively, v_p and v_r are velocity of the projectile just before impact and residual velocity respectively, E_{DH} is the energy required for deformation and heating and E_s^* is the energy required to shear the plug out. A mathematical model for the plugging perforation of plates was presented by Awerbuch and Bodner (1974) who considered the three stages (i) compressive stage where the projectile acquire added mass from the plate with no shearing of the plug (ii) both compression and shearing of plug occur (iii) plug continue to shear as a rigid body without compression of plug. The three-stage theory by Awerbuch and Bodner was later refined to five stages by Ravid and Bodner (1983)

(1) dynamic plastic penetration
(2) bulge formation
(3) bulge advancement
(4) plug formation and exit
(5) projectile exit

Further insights into the plugging process was presented by Woodward (1987) who considered dishing, stretching and plugging deformations where no post-perforation deformation measurements are required (predictive theory). The application of the above five-stage model to ballistic penetration to ceramic armor or ceramic-metallic armor was presented by (Ravid, Bodner and Holcman 1987, 1989).

2. PERFORATION OF LAMINATED COMPOSITE MATERIAL PLATES

As a first approximation, a lamianted composite material plate can be considered as stacking of the same number of layer of metallic plate where each layer of metallic plate is of equal areal density as each laminated layer. These metallic layers are "glue" together such that the shearing between adjacent layers is the same as the interlaminar shear stress of the laminated plate. Such modeling, especially with the inclusion of the interlaminar shear stress, has not been considered in the literature.

Marom and Bodner (1979) were the first to examine the plugging perforation of multi-layered beams which are separated or in contact. In the case of contacted beams, adjacent layers are free to slide (no glue). Marom and Bodner did not consider glued beams. Note that it is not obvious whether the the inclusion of glue would be advantageous (in a sense of defeating the projectile or reducing the residual velocity) over the no-glue case. The advantage of glue is to increase interlaminar shear resistance and the disadvantage is to reduce overall bending of the beam (note that overall bending may be an important energy absorbing mechanism). More discussions on this topic for unbonded layered plates were presented by Radin and Goldsmith (1988), using essentially the same equations as Maron and Bodner. The advantage of "bonded" versus "unbonded" plates is subject to some conrroversies since bonded plates has shear resistance between layers (advantage) but provide little overall deformation (disadvantage).

The energy balance equation for the perforation of bonded layered plate is, assuming a rigid blunt-nose projectile,

$$(1/2)m_p v_p^2 = E_{DH} + E_s^* + E_s + (1/2)(m_p+m^*)(v_r)^2 \qquad (2)$$

The above equation is identical to eqn. 1 with the inclusino of the energy for interlaminar shear between adjacent layers. Note that E_s vanishes in the case of unbonded layered plates or when the total thickness of bonded layered plates is small (so that thin plate theory applies).

In classical lamination theory, no account is taken of the interlaminar (perpendicular to the plane of the plate) stresses such as the interlaminar normal stress σ_z and the interlaminar shear stresses τ_{zx} and τ_{zy}. Rather, only the in-plane stresses σ_x, σ_y and τ_{xy} are considered (Jones 1975). Thus, classical lamination theory would not be able to predict some of the stresses which cause failure of a composite material. Jones (1975) showed that the existence of interlaminar stresses implies that the laminated plate can delamiante near free edges. Such free edge may be a hole generated by the projectile in the perforation process. Note that the interlaminar stresses died out rapidly away from the free edge and these stresses are significant only within one laminate thickness from the free edge(Pipes and Pagano 1970, Pagano and Pipes 1971 and Pipes and Daniel 1971). Finite element solutions for these interlaminar stresses as a function of the stacking sequence was presented by Halpin(1984), using three-dimensional elasticity equilibrium equations.

Assuming the worst case when all the layers are "delaminated" from each other, that is, adjacent "metallic" layers are free to slide with respect to each other, one can employ the energy balance equation to obtain the ballistic limit velocity. Setting the residual velocity of the rear layer (that is, the nth layer where n is the number of layers) to zero, the work needed to penetrate this layer is

$$W_n = (1/2)m_p v_{BL}^2$$

Note that the v_{BL} is computed for each layer such that the initial velocity for the ith layer is the residual velocity of the $i-1^{th}$ layer. Thus, one can start from the rear layer and compute the initial velocity for the $n-1^{th}$ layer, then the $n-2^{th}$ layer so that eventually,. the initial velocity for the first layer is obtained (this velocity is precisely the ballistic limit for the entire laminate if one starts with $v_r=0$ for the rear plate). It is postulated that,

$$v_{BL}(i^{th} \text{ layer}) < v_{BL}(i+1^{th} \text{ layer})$$

since the size of the hole tends to increase due to the added mass effect on the projectile. Since the layers are in contact with each other, the energy dissipated for the ith layer in the plugging process is due to the shearing of the plug as well as due to the compression of the remaining layers i^{th}, $i+1^{th}$,...,n-1, n^{th} layer:

$$E_c = \sigma_{u_1 t} (\pi R^2) (t_i + t_{i+1} + \ldots + t_n)$$

where σ_{ult} is the ultimate stress and R is the average hole size.

Following Marom and Bodner (1979), the energy transmitted to the ith layer in the shearing of the plug is the same as that for an isolated layer so that, for each layer,

$$E_s^* = (1/2) \; \frac{\overline{m}_p}{\overline{m}_p + m^*} \; m_p \; v_{BL}^2$$

where $\overline{m}_p$ is the added mass of the projectile. Note that the ballistic limit of the ith layer is related to the ballistic limit of a similar isolated layer by,

$$v_{BL}(i^{th} \text{ layer}) = v_{BL}(\text{isolated } i^{th} \text{ layer}) + \frac{\sigma_{ult} \; (\pi)(d^2)h(\overline{m}_p + m^*)}{2\overline{m}_p^2}$$

3. PARTITION OF ENERGY DISSIPATION IN COMPOSITE PLATES DUE TO PROJECTILE IMPACT AND DESIGN OF EXPERIMENTS

The total energy just before impact is $(1/2)\, m_p v_p^2$

The excess energy is $(1/2)\, m_p v_p - (1/2)\, m^* v_r^2 = E_X$

This energy is dissipated in the following form

(i)   Hertzian Contact and compression induced added mass E_H (using MTS)

(ii)   Plate bending E_B (measure using MTS and strain gauge to determine the deflection)

(iii) Delamination, including matrix cracking and fiber breakage E_D
           (using Hopkinson Bar   and then C-Scan)

(iv)   Shear Plugout E_S (using MTS and Hopkinson Bar)
              measure weight of plate before and after shooting

(v)   Petalling due to excessive circumferential stress
           and stretching due to inelastic elongation E_P
       (impact of plate with a pre-drilled tiny hole)

(vi)   Vibration, possibly finite amplitude vibration with damping E_V
           calculate based on material and structural properties, with
           special emphasis on time to first peak and logarithmic decrement
           from the displacement vs time curve, one can estimate the energy
           absorption.

$$E_X = E_H + E_B + E_D + E_S + E_P + E_V$$

us'ng Poly-vinylidene fluoride PVDF Piezopolymer sensor, one can

imbed these gauges in between layers and subject the composite material

sample under impact using the Hopkinson Bar apparatus

(i) Hertian Contact

$$\text{Force} = \text{const.} \left(\text{relative approach}\right)^{3/2} \quad \text{valid for quasi-static impact}$$

The energy dissipation begin with the stress concentration at the

contact surface, followed by compression induced added mass

This Hertian Contact phase will end with the formation of the

first crack starting from the contact surface.

The above power law coefficient should be determined for medium or

high speed impact.

relative approach
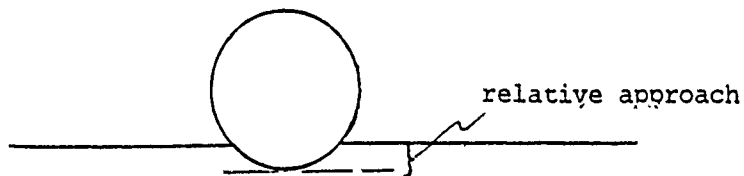
The above formula assume that the contact duration between the impactor

and the target is very long in comparison with their natural period of

vibration. Thus, it should be modified for very high speed impact

of plates so that the force is proportional to a different power.

The constantant of proportionality is a function of the geometric

and material properties or both the projectile and the laminated plate.

(ii) Plate Bending

Circular plate under concentrate load at the center

close form solution exists, assuming clamped boundary conditions, one can find the deflection and hence, the maximum stress located at the back surface of the line of impact. Equating this stress to the failure stress of the compostie material, which is usually slightly above the yield stress. One can then determine the impact force using Greszczuk method for low velocity impact.(see Shivakumar et al. 1985a,b)

(iii) Delamination

One can use very thin aluminum layers bonded togetner and test the bonded plate in a hopkinson bar setup.
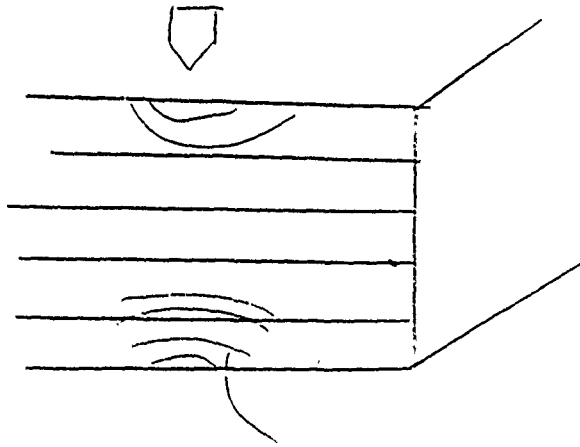
The initial wave is compressive until it reaches tne back surtace and this wave is reflected back as tensile wave. The tensile wave is responsible for the delamination nor separation of the layers. Such waves can be detected using imbedded sensors. The sensors will enable one to obtain the stresses versus time and compute the energy dissipation between adjacent layers.

(iv). Shear Plugout
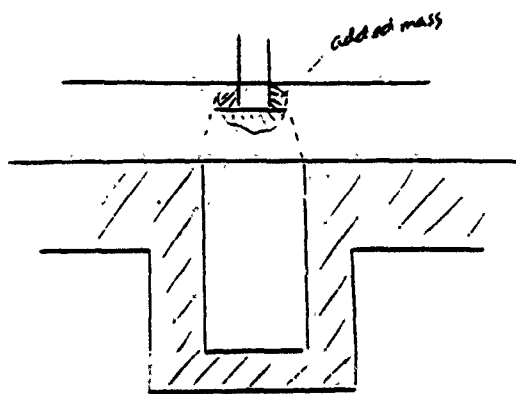
The sample is subjected to a flat nose projectile impacting the plate resting on a hat-shaped rigid support. The effects of shear plugout and petalling can be uncoupled by the introducing specially designed experiments. For example, the indentation of a flat punch, see above diagram using MTS for static test and hopkinson bar for dynamic test will enable one to determine the energy dissipation due to shear plugout.

(v) Petalling

As a rule of the thumb, thin plate tends to petal and thick plate tends to plug. Most plates (isotropic-homogeneous or laminated-fiber-reinforced) would fail under projectile impact in a combination of plugging and petalling. Petalling is due to the enlargement of the cylindrical surface due to the added mass, such that the noop stress exceed a critical value. We can determine the energy due to petalling by drilling a tiny hole and shoot the projectile at this hole using the MTS machine for the static case and hopkinson bar in the dynamic case. Note that the hopkinson bar apparatus would enable one to "aim" at the tiny hole precisely.

103-12
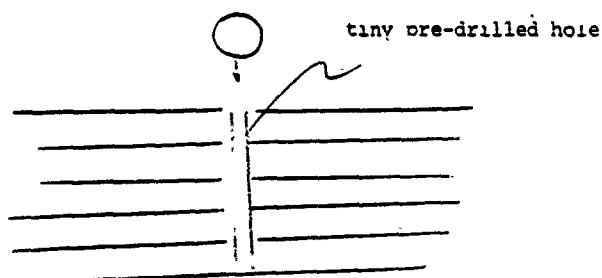
(vi) Finite Amplitude Vibration

The plate will vibrate under the initial condition of zero displacement
and a prescribed initial velocity. The structural parameters will be the
size of the plate, the total thickness of the plate and the boundary
conditions. The material parameters will be Young's modulus and Poisson's
ratio for the isotropic-homogeneous plate and there will be more such
elastic constants for a laminated plate. Further, one would need to estimate
the viscous damping (force is proportional to the velocity) coefficient
and the material damping (hysteresis damping) coefficient using the
complex modulus loss factor. Based on the displacement versus time curve
of the vibration, taking into account finite deflection terms, one can
determine the energy dissipation. One need to determine the energy dissipation
in the first few cycles since the plate tends to damped out quickly .
The first or second peak is likely to cause most of the damage since
they usually correspond to relatively large deflections.
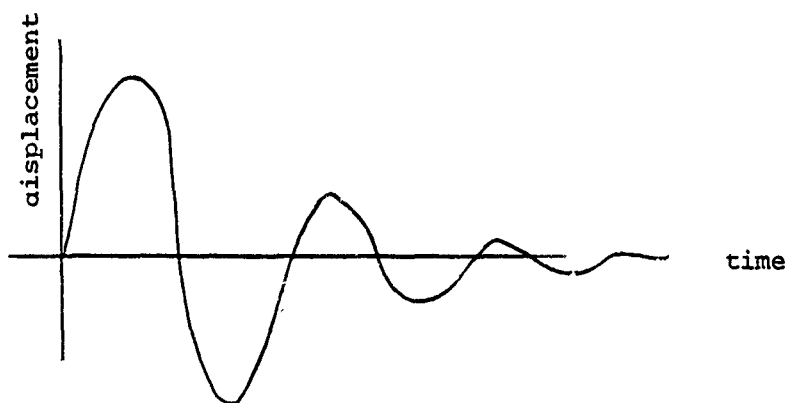
4. CONCLUSIONS

The present work deals with a preliminary attempt to formulate a theoretical modeling of the perforation of laminated plates by rigid projectile at ordnance velocity. Since the subject is an extremely complicated one, the present work would focus only on energy dissipation aspect on which the theoretical and the experimental programs are based upon. Although there are still many unresolve issues, it appears that the important issues are spelled out and they will be the subject of intense investigation in the coming years. Further refinement of the theroy is needed as the experimental data becomes available. Extensive tests using the Hopkinson bar apparatus, charpy impact tests and the static or drop tests for impact on laminated composite materials plates are planned.

5. Acknowledgements

6. REFERENCES

Awerbuch, J. and Bodner, S.R. (1974), "Analysis of the Mechanics of Perforation of Projectiles in Metallic Plates", Int. J. of Solids and Structures, Vol. 10, pp. 671-684.

Halpin, J.C. (1984), "Primer on Composite Materials: Analysis", revised version, Technomic Publishing Co., Inc.

Johnson, W. (1972), "Impact Strength of Materials", Edward Arnold Publishers Limited.

Johnson, W., Chitkara, N.R., Ibrahim, A.H. and Dasgupta, A.K. (1973), "Hole Flanging and Punching of Circular Plates with Conically Headed Cylindrical Punches", J. of Strain Analysis, Vol.8, No. 3, pp. 228-241.

Jones, R.M., (1975), "Mechanics of Composite Materials", Scripta Book Co. and McGraw-Hill.

Landkof, B. and Goldsmith, W. (1985), "Petalling of Thin, Metallic Plates during Penetration by Cylindro-Conical Projectiles", Int. J. of Solids and Structures, Vol. 21, No. 3, pp. 245-266.

Marom, I. and Bodner, S.R. (1979), "Projectile Perforation of Multi-Layered Beams", Int. J. of Mechanical Sciences, Vol. 21, pp. 489-504.

Pagano, N.J. and Pipes, R.B. (1971), "The Influence of Stacking Sequence on Laminate Strength", J. of Composite Materials, Vol. 5, January, pp. 50-57.

Parkes, E.W. (1955), "The Permanent Deformation of a Cantilever Struck Transversely at its Tip", Proc. of the Royal Society of London, Vol. 228, March, pp. 462-476.

Parkes, E.W. (1958), "The Permanent Deformation of an Encastre Beam Struck Transversely at any point in its Span", Proc. of the Institution of Civil Engineers, Vol. 10, pp. 277-304.

Pipes, R.B. and Pagano, N.J. (1970), "Interlaminar Stresses in Composite Laminates under Uniform Axial Extension", J. of Composite Materials, Vol. 4, October, pp. 538-548.

Pipes, R.B. and Daniel, I.M. (1971), "Moire Analysis of the Interlaminar Shear Edge Effect in Laminated Composites", J. of Composite Materials, Vol. S, April, pp. 255-259.

Radin, J. and Goldsmith, W. (1988), "Normal Projectile Penetration and Perforation of Layered Targets", Int. J. of Impact Engineering, Vol. 7, No. 2, pp. 229-259.

Ravid, M. and Bodner, S.R. (1983), "Dynamic Perforation of Viscoplastic Plates by Rigid Projectiles", Int. J. of Engineering Sciences, Vol. 21, No. 6, 1983, pp. 577-591.

Ravid, M., Bodner, S.R. and Holcman, I. (1987), "A Two-Dimensional Engineering Model for Perforation of Layered Targets", Proc. of the Tenth Int. Symposium on Ballistics, San Diego, Calif., October 27-29, Volume 2, pp. 1-9.

Ravid, M., Bodner, S.R. and Holcman, I. (1989), "Application of Two-Dimensional Analytical Models of Ballistic Penetration to Ceramic Armor", Proc. of the Eleventh Int. Symposium on Ballistics, Brussels, Belgium, May, 8pp.

Recht, R.F. and Ipson, T.W. (1963), "Ballistic Perforation Dynamics" ASME J. of Applied Mechanics, September, pp. 384-390.

Shivakumar, K.N., Elber, W. and Illg, W. (1985a), "Prediction of Low-Velocity Impact Damage in Thin Circular Laminates", AIAA Journal, Vol. 23, No. 3, March, pp. 442-449.

Shivakumar, K.N., Elber, W. and Illg, W. (1985b), "Prediction of Impact Force and Duration due to Low-Velocity Impact on Circular Composite Laminates", ASME J. of Applied Mechanics, Vol. 52, September, pp. 674-680.

Sodha, M.S. and Jain, V.K. (1958), "On Physics of Armor Penetration", J. of Applied Physics, Vol. 29, No. 12, December, pp.1769-1770.

Taylor, G.I. (1948), "The Formation and Enlargement of a Circular Hole in a Thin Plastic Sheet", The Quarterly Journal of Mechanics and Applied Mathematics, Vol. 1, pp. 103-124.

Thomson, W.T. (1955), "An Approximate Theory of Armor Penetration", J. of Applied Physics, Vol. 26, No. 1, January, pp. 80-82 and 919-920.

Woodward, R.L. (1978), "The Penetration of Metal Targets by Conical Projectiles", Int. J. of Mechanical Sciences, Vol. 20, pp. 349-359

Woodward, R.L. (1987), "A Structural Model for Thin Plate Perforation by Normal Impact of Blunt Projectiles", Int. J. of Impact Engineering, Vol.6,No.2, pp. 129-140.

Appendix A   Extension of Thomson's paper to spherical projectile

For a   spherical projectile, we have,
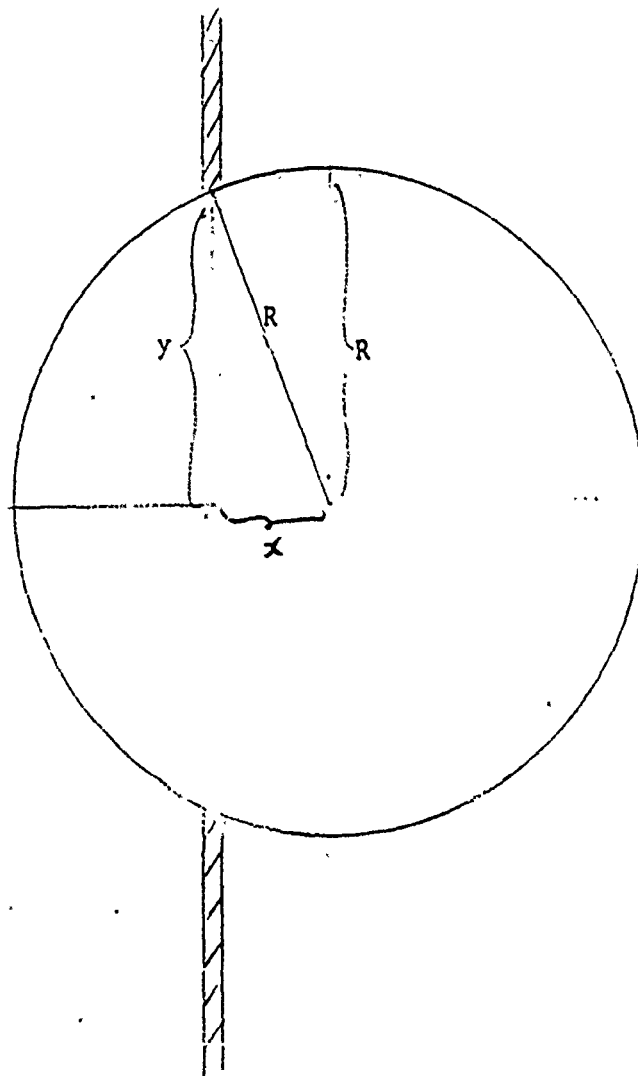
The aim of this appendix is to show that for a spherical projectile, the petalling equations provided by Thomson yields infinite work. This shows that the spherical projectile is actually a "blunt-nose" projectile and theoretically, plugging is the only possible perforation mechanisms and petalling is theoretically not possible, using Thomson's theory.

From Pythagoran theorem,

$$(R- Vt)^2 + y^2 = R^2$$

solving for y, we have,

$$y^2 = R^2 - (R-Vt)^2$$

$$\boxed{y^2 = 2RVt - V^2t^2}$$

that is,

$$y = (2RVt - V^2t^2)^{1/2}$$

Taking the first derivative, one obtains,

$$y,_t = (1/2)(2RVt - V^2t^2)^{-1/2} (2RV - 2V^2t)$$

that is,

$$\boxed{y,_t = (2RVt-V^2t^2)^{-1/2} (RV-V^2t)}$$

Taking the second derivative, one obtians,

$$\boxed{y,_{tt} = (-1/2) (2RV -V^2t^2)^{-3/2}(-2V^2t) - V^2 (2RV- V^2t^2)^{-1/2}}$$

$$x^2 + y^2 = R^2$$

$$y = (R^2 - x^2)^{1/2}$$

$$\frac{dy}{dx} = (1/2) \; (R^2 - x^2)^{-1/2} \; (-2x)$$

so that,

$$y,_x = (-x) \; (R^2 - x^2)^{-1/2}$$

$$y,_{xx} = - (R^2 - x^2)^{-1/2} + (x/2) \; (R^2 - x^2)^{-3/2} \; (-2x)$$

that is,

$$y,_{xx} = -(R^2 - x^2)^{-1/2} - (x^2) \; (R^2 - x^2)^{-3/2}$$

The integrals are,

$$I_1 = \int y^2 \; y,_{tt} \; dy$$

note that,

$$\frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt} = V \frac{dy}{dx}$$

$$y,_{tt} = (y,_t),_t = V(y,_t),_x = V^2 \; y,_{xx}$$

$$dy = \frac{dy}{dx} \; dx$$

$$I_1 = V^2 \int_{x=R}^{x=0} y^2 \; y,_{xx} \; y,_x \; dx \quad , \quad 2I_2 = 2V^2 \int_{x=R}^{x=0} y \; (y,_x)^3 \; dx$$

so that

$$I_1 + 2I_2 = V^2 \int_{x=R}^{x=0} y^2 \, y_{,xx} \, y_{,x} + 2y \, (y_{,x})^3 \; dx$$

$$= V^2 \int_{x=R}^{x=0} (-x) \left[-1 - (x^2)(R^2-x^2)^{-1} \right] + 2 \, (-x^3)(R^2-x^2)^{-1} \; dx$$

$$= V^2 \int_{x=R}^{x=0} x - (x^3)(R^2-x^2)^{-1} \; dx$$

$$= (-V^2 R^2/2) - (V^2/2) \int_{x=R}^{x=0} \frac{x^2}{R^2 - x^2} \; d(x^2)$$

$$Q = R^2 - x^2, \qquad \text{so that,} \qquad x^2 = R^2 - Q, \qquad d(x^2) = -dQ$$

so that,

$$I_1 + 2I_2 = (-V^2 R^2/2) - (V^2/2) \int_{Q=0}^{Q=R^2} \left(\frac{R^2 - Q}{Q} \right)(-1) \; dQ = \text{infinite}$$

Thus,

$$I_1 + 2I_2 = (-V^2 R^2/2) + (V^2/2) \left\{ R^2 \left[\ell n(Q) \right]_{Q=0}^{Q=R^2} - R^2 \right\}$$

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by

Universal Energy Systems, Inc.

FINAL REPORT

ACCELERATE FATIGUE TEST PROCEDURE FOR THE
STRUCTURAL POLYCARBONATE COMPONENT OF THE
F-16 CANOPY COMPOSITE MATERIAL

Prepared by:	Yulian B. Kin, Ph.D.
Academic Rank:	Associate Professor
Department and	Engineering Department
University:	Purdue University Calument
Research Location:	Flight Dynamics Laboratory
USAF Researcher:	Robert E. McCarty, Lorene V. Garrett
Date:	1 August 1990
Contract No:	F49620-88-C-0053

ACCELERATED FATIGUE TEST PROCEDURE FOR THE STRUCTURAL POLYCARBONATE COMPONENT OF THE F-16 CANOPY COMPOSITE MATERIAL

by

Yulian B. Kin

ABSTRACT

The long-term fatigue test procedure requires the breaking of 20 to 30 identically prepared specimens and one month to complete. Thus, manufacturers often do not perform a conventional fatigue test in spite of its obvious utility. Therefore, there is a definite need for an accelerated fatigue test which can be completed in approximately one day. The accelerated test procedure proposed can be developed on the basis of the data gained by the principal investigator during the conventional fatigue test run with the help of the UES mini-grant S-210-9M6-038 in 1989. The mini-grant was awarded to continue the research started by Yulian Kin during his summer appointment at Wright-Patterson Air Force Base in 1988.

ACKNOWLEDGEMENTS

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for the sponsorship of this research. Universal Energy Systems must be mentioned for their concern and help to me in all administrative, directional and social aspects during the summer appointment and the pre-summer visit.

I am grateful to Group Leader Robert McCarty and Aerospace Engineer Lorene Garrett. They provided support, attention, and a comfortable working atmosphere. I appreciate very much Lorene's help with my computer needs. Many thanks to Mrs Evelyn Schutte for the help in preparation of this report. Dr Arnold Mayer's interest in every phase of my summer and possible future projects served as a source of encouragement.

I.   INTRODUCTION:

There are assumptions based on preliminary tests that fatigue mechanical properties of the polycarbonate sheets are significantly vary from sheet to sheet and probably along the same sheet.

Therefore, it is very important to have a vehicle which permits to control (quickly and constantly) the quality of the polycarbonate sheets and detect possible deviations of the manufacturing process, in other words to control the stability of the manufacturing process.

The accelerated fatigue test can be utilized for such kind of control and the procedure of the test can be developed if some fatigue parameters for the tested materials or real parts are known from the conventional test.

The accelerated fatigue test can be also useful to make a preliminary estimate of new designs. For example, Flight Dynamics Laboratory is initiating the investigation of the fatigue characteristics of polycarbonate sheets with different thicknesses.

For the obvious economical reasons a conventional fatigue test should be run to confirm the expected possible output only if the accelerated test shows the strong dependency of fatigue strength on the polycarbonate sheet thickness.

II.   OBJECTIVES OF THE RESEARCH EFFORT:

The immediate objectives of the research were to:

1.   Develop a procedure including a regime program of the
accelerated fatigue test for the coupons cut from a
structural polycarbonate sheet of the F-16 canopy composite
material. The necessary fatigue parameters (fatigue
strength, knee points and fatigue curve slopes for 5%, 90%
and 95% probability fatigue curves) were computed from the
statistics determined by Yulian Kin during his SFRP in 1988
followed by a mini-grant program in 1989.

2.   Manufacture specimens in accordance with the procedure
given in [1].

3.   Start a conventional fatigue test using the procedure
given in [1]. It was my intention to complete 40% of this
test during the summer appointment and to continue at Purdue
University with funding from the mini-grant program.

4.   Run (if time permits) a preliminary accelerated fatigue
test using the procedure from item 1 and statistics from item
3 and from [1].

III. LOCATY'S METHOD FOR ACCELERATED DETERMINATION OF FATIGUE STRENGTH

1. Description of the Method.

The accelerated technique for the specified type of specimens can be reliably developed if the statistics of a conventional fatigue test of similar specimens are known. Therefore, the results of the preliminary run conventional fatigue test were used to assign certain parameters and estimate the error of the accelerated test.

Locaty's accelerated technique was used to start an investigation. It was assumed that time to run the proposed accelerated test would be within 10 hours.

The idea and description of the method had been given in appendix to [1] and repeats here in concise form for the convenience.

The method is based on the concept of cumulative fatigue damage or Palmgren - Minor rule considering $\sum \frac{n_i}{N_i} = 1\text{-}5$, where n_i is the number of cycles which specimen worked in the specified test regime, and N_i is the number of cycles which specimen could potentially work according to the fatigue curve received from the results or the conventional fatigue test of the same type of specimens.

The loading program and the treatment of results are
presented in Figures 1 and 2. Figure 1 shows three fatigue
curves (for example, 5%, 50%, and 95% probability of failure)
received from a conventional fatigue test.

During an accelerated test the specimen works, for example,
for 50,000 cycles at the first load level, then the load is
increased and the specimen is again tested for 50,000 cycles.
The procedure is repeated until the instant when the specimen
fails. Then using Figure 1, the magnitudes of

$$\sum \left(\frac{n_i}{N_i}\right)_C , \quad \sum \left(\frac{h_i}{N_i}\right)_B , \quad \sum \left(\frac{n_i}{N_i}\right)_A$$

are determined. With these three parameters (or, if
necessary, a greater number of points) and knowing the
corresponding stresses, we can find the coordinates of the
points which result in the curve shown in Figure 2. Now, if
according to an accepted hypothesis fatigue strength
corresponds to a definite value (for example, $\sum \frac{h_i}{N_i} = 1$),
we can easily determine the fatigue strength magnitude
(Figure 2).
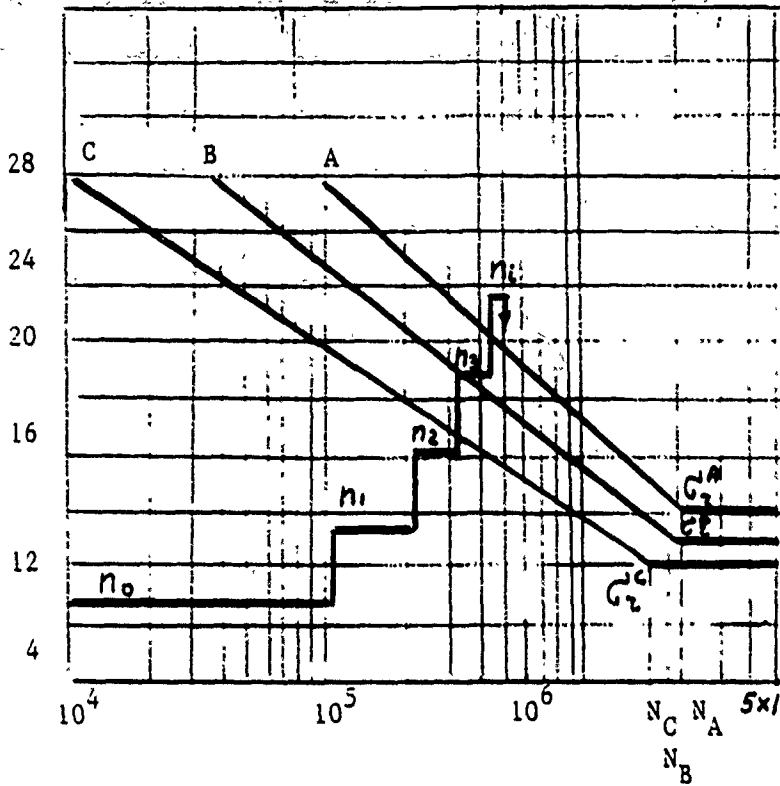
Fig 1. Loading Program
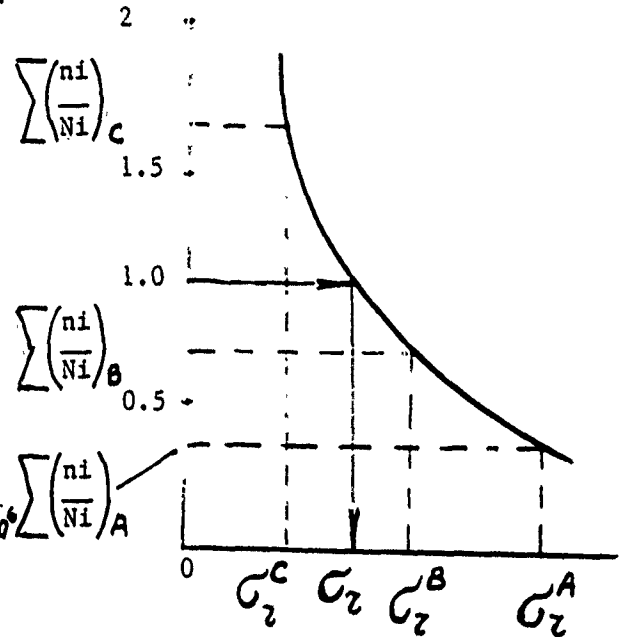
Fig 2. Graphical Determination
of Fatigue Strength

2. Nomenclature and symbols

Parameter	Symbol	Unit	Reference
Initial stress	σ_o	Pa	Figure 1
Stress Increment	$\Delta\sigma$	Pa	Figure 1
Initial number of cycles	n_o	Cycle	Figure 1
Number of cycles at ith level of stress	n_i	Cycle	Figure 1
Rate of stress increment	α	Pa/Cycle	

Rate of stress increment	α	Pa/Cycle	
Stress at failure	σ_F	Pa	Figure 1
Fatigue strength determinated by acceler. method	σ_{ACC}	Pa	Figure 2
Number of cycles at ith level of stress determined during conventional fatigue test	N_i	Cycle	
Sum of relative lives	$\sum n_i/N_i$		
Fatigue strength corresponding to curve A, B, C respectively:	$\sigma_F^A;\ \sigma_F^B;\ \sigma_F^C,\ P_a$		
Expected fatigue strength	$\sigma_{EXP}\ ,$	Pa	
Knee points (if any) on curves A, B, respectively	N_0^A		Figure 1
	$N_0^B = N_{EXP}$		
	N_0^C		
Slopes of curves A, B, C respectively	K_A		
	K_B	$K = \dfrac{\sigma_{m+1} - \sigma_m}{\log N_m - \log N_{m+1}}$	
	K_C		

3. Loading Regimes

Intervals of the initial load level are assigned from the inequality :

$$\sigma_{EXP} < \sigma_0 < 1.2\,\sigma_{EXP}$$

Magnitude or σ_{EXP}; N_{EXP}, and K_A, K_B, K_C are assigned

after the completion of the conventional test. Optimal value α

of the stress increment rate depends on the selected ratio σ_0/σ_{EXP}

and can be (as in the first try) selected from 0.7 kpsi

X 10^{-5} to 3.5 kpsi X 10^{-5} per cycle.

Optimal intervals for the stress increment are found from the

expression $0.05\, \sigma^2_{EXP} < \Delta \sigma^2 < 2.0\, \sigma^2_{EXP}$

Value of n_i is determined from $n_i = \Delta \sigma^2 / \alpha$

In case when n_i appears to be greater than 3×10^5, it is necessary to decrease $\Delta \sigma^2$.

Most of the recommendations are based on test experience with metal and corrections are quite possible.

The specimens, parameters and regimes selected for the preliminary accelerated test are given below.

SPECIMENS

Specimen dimensions and loading diagram are given in Figures 3 and 4.
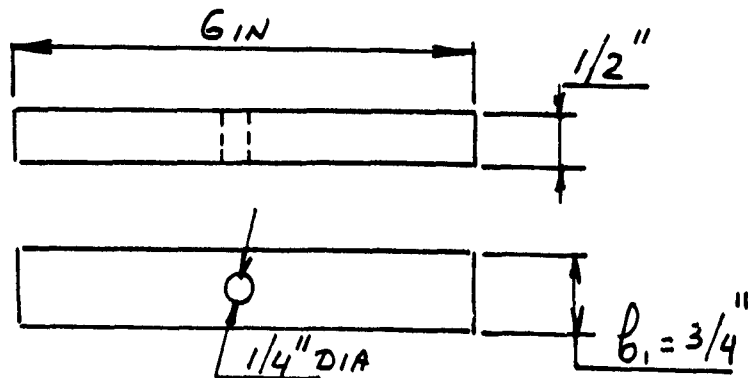
Figure 3. Specimen Used in Study
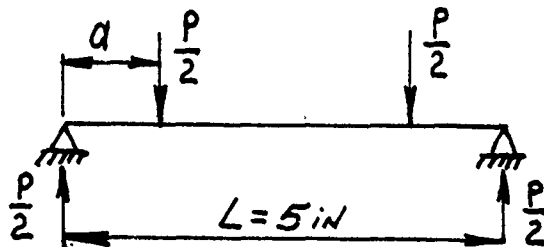
Figure 4. Loading Diagram

REGIMES OF THE CONVENTIONAL TEST

Loads and stresses assigned for the conventional test are given in Table 1.

Table 1. Conventional Test Regimes

Max. Load on the Machine Gauge, Lc	Min. Load on the Machine Gauge Lb	Amplitude Load Applied to the Specimen, Lb	*Nominal Amplitude Stress Applied to the Specimen, psi	Frequency, HZ	Rmin Rmax
1500	300	600	7200	3	0.2
1080	200	440	5280	4	0.2
720	144	288	3456	6	0.2
360	72	144	1728	12	0.2
180	50	65	780	15	0.28

Distance "a" (see Figure 4) is a constant for all regimes.

$$a = \frac{\sigma_{MAX}(bh)^2}{3P}$$

*Actual magnitude of amplitude stress calculated taking into account actual specimen dimensions.

$$a = \frac{6000(1/2)(3/4)^2}{3(750)} = 0.75$$

$$a = \frac{4320(1/2)(3/4)^2}{3(540)} = 0.75$$

$$a = \frac{2880\ (1/2)(3/4)^2}{3(360)} = 0.75$$

$$a = \frac{1440(1/2)(3.4)^2}{3(180)} = 0.75$$

REGIMES OF THE ACCELERATED TEST

50% probability fatigue curve "B"(Fig. 5)15 built using regression analysis of the data gained in (1) for the specimens with stress concentrators.

AMPLITUDE STRESS, PSI

log N = 9.63 - 1.24 log σ

50% Probability Fatigue Curve

A

B

C

Number of Cycles to Failure
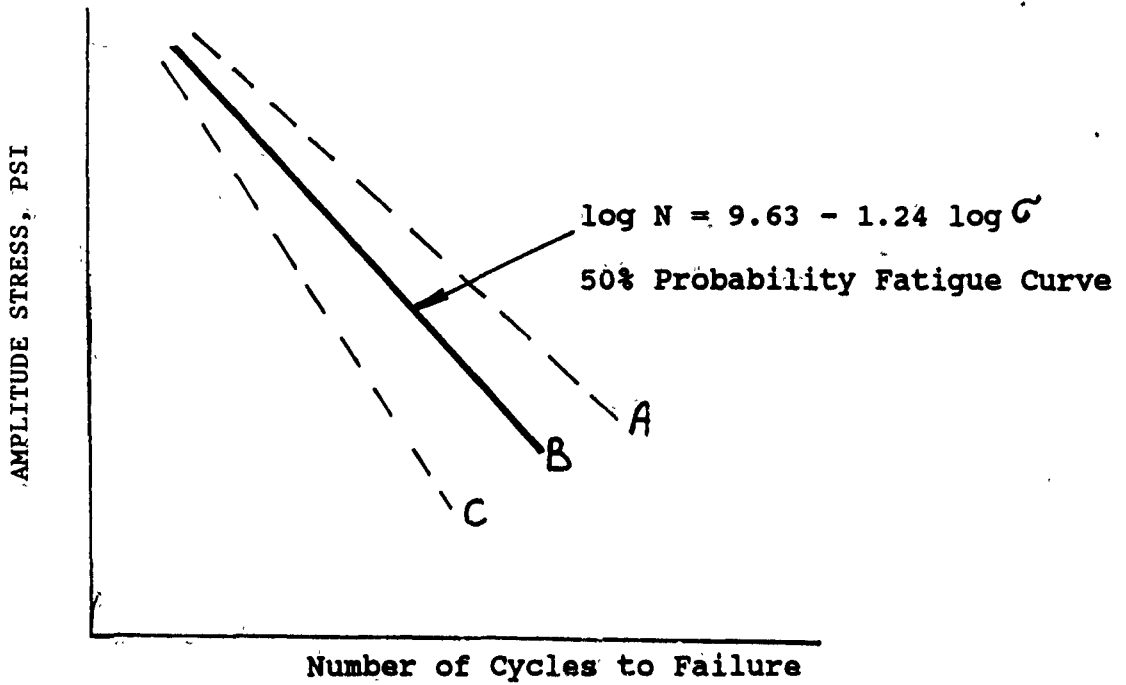
Figure 5. 5%, 50%, 95% probability of failure fatigue curves.

Slope of curve "B" $K_B \approx 5$

$$K_B = \frac{\sigma_{h+1} - \sigma_h}{\log N - \log N}$$

Expected fatigue strength for the life $N_0^B = 10^6$ cycles is $\sigma_R^B = 1440$ psi (selected from [1].

PARAMETERS FOR CURVES "A" and "C"

$$N_0^C = 0.8 \ N_0$$

$$N_0^A = 1.2 \ N_0^B \quad \text{or} \quad N_0^A = N_0^B = N_0^C$$

$$K_C = K_B - 3$$

$$\text{or} \quad K_A = K_B = K_C$$

$$K_A = K_B + 3$$

$$\sigma_R^C = 0.9 \ \sigma_R^B$$

$$\sigma_R^A = 1.1 \ \sigma_R^B$$

INITIAL LOAD, LOAD INCREMENT, NUMBER OF CYCLES AT ONE LEVEL

Initial Load:

Select initial amplitude stress = 1500 psi hence amplitude load on the machine gauge is

$$P_A = \frac{\sigma bh^2}{3a} = \frac{1500(1/2)(3/4)^2}{3(3/4)} = 18 \ 7.5$$

from equations

$$P_A = \frac{Pmax - Pmin}{2}$$

and

$$\frac{Pmin}{Pmax} = 0.2$$

$$Pmax = 500 \text{ lb}$$

$$Pmin = 100 \text{ lb}$$

Load increment:

$$0.05 \ (P_R^\beta) \ < \ P \ < \ 0.20 \ (P_R^\beta)$$

$$P_R^\beta = 450 \text{ lb (Table 1)}$$

$$0.05 \ (450) \ < \ \Delta P \ < \ 0.20 \ (450)$$

$$22.5 \ < \ \Delta P \ < \ 90$$

We assume that $\Delta P = 100$ lb

Number of cycles ni at one level

$$ni = \frac{\Delta \sigma_{MAX}}{\lambda}$$

$\Delta \sigma$ max corresponding to $\Delta P_{MAX} = 100$ lb:

$$\Delta \sigma_{MAX} = \frac{(0.75)(3)(100)}{(1/2)(3/4)^2} = 800 \text{ psi}$$

If we take $\quad \lambda = 1.6 \times 10^{-5}$ (see page 10)

$$ni = \frac{0.8}{1.6 \times 10^{-5}} = 50000 \text{ cycles.}$$

IV TEST RESULTS

The MTS testing machine was calibrated just before the test. The installed load magnitudes were red out from the oscilloscope and additionally (for verification purpose) controlled with digital multi-meter. 50% of the test was completed and results are given in Table 2.

Test Results

Table 2. Comparison of 1989 and 1990 Tests

Amplitude Stress, psi	Number of cycles to complete failure, 1990 test	Number of cycles to crack initiation, 1990 test	Number of* cycles to complete failure, 1989 test
7200	1200	--	
	1000	--	
5280	4800	--	110000
3456	12800	8500	
	16600	11000	180000
	23500	18000	
	17000	10500	
1728	134300	85000	310000
	156000	90000	
960	1000000 (No failure)	--	450000

*The numbers were taken from Figure 4 [1].

From the results gained it can be assumed (take note that the test has not been completed yet) that the 1989 material had greater fatigue strength than the 1990 polycarbonate sheet. But the scatter is significantly less in the 1990 test. The possible difference in strength of two polycarbonate sheets of the same type shows one more time that the tight and quick control of fatigue characteristics is necessary to verify the product quality. In the same time, relatively consistant repetition of the test results can be used as a confirmation of the possibility to develop a reliable accelerated test procedure.

V RECOMMENDATIONS

1. Continue the investigation to develop an accelerated fatigue test procedure for the structural polycarbonate of the F-16 canopy composite material. With the help of mini-grant and some additional funds the work can be completed in 1991.

2. Apply the accelerated testing technique to the coupons with and without stress concentrators.

3. Investigate yhe dependency of fatigue characteristics on thickness of the polycarbonate sheet. If the dependence exists, interesting proposals to improve design can be done.

4. Prepare a procedure for the fatigue test of the canopy composite material. Preliminary hypothesis of failure can be developed on the basis of finite element analysis.

5. Prepare a procedure for the environmental fatigue investigations of the canopy material.

6. Fatigue investigations are very time consuming and expensive. In the meantime, it is clear that Aircrew Enclosures Group of Vehicle Subsystems Division needs to provide different fatigue investigations in the near future. Taking this into consideration. It is proposed to establish a mutual research laboratory with the installation of a new modern fatigue machine at Purdue in Hammond. The testing machine which vehicle Aircrew Enclosures Group is planning to buy can be partially paid from the Purdue funds. Purdue University Calumet Engineering Department has space for installation and experienced personnel to provide investigation, maintenance and calibration of the machine. Engineering senior students will handle the tests while developing their senior design projects. (Graduate students can be also attracted.) For the first five years the machine will be used for the Aircrew Enclosures Group projects only. After five years the machine can be returned to Wright-Patterson or continue to be installed at Purdue with the availability for other projects also. The economical benefits of the proposal are obvious (the expenses are

divided between two institutions working on the same projects, the time of tests is decreased because of the number of students available for the work, the student salary is much less than for the regular personnel). The educational benefits are also very significant because the students will get the opportunity to work on very serious, interesting and practical projects. The principal investigator has already in general discussed this proposal with the department head and met his understanding.

REFERENCES

1. Yulian Kin., "Fatigue characteristics of F-16 composite transparency material determined by long-term and accelerated methods," Final Report, Contract No. FY9620-88-C-0053/SB5881-0378, Universal Energy Systems, Inc., Dayton, Ohio, 1990.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

<u>FINAL REPORT</u>

<u>Study of Fracture Behavior of Cord-Rubber Composites for</u>

<u>Lab Prediction of Structural Durability of Aircraft Tires</u>

Prepared by:	Byung-Lip ("Les") Lee
Academic Rank:	Associate Professor
Department and	Engineering Science and Mechanics .
University:	Pennsylvania State University, Univ. Park, PA
Research Location:	Wright-Patterson Air Force Base
	Wright Research and Development Center
	Flight Dynamics Laboratory
	Vehicle Subsystems Division
	Aircraft Launch and Recovery Branch
USAF Researchers:	Gary J. Migut, John P. Medzorian
Date:	30 September, 1990
Contract No:	F49620-88-C-0053

Study of Fracture Behavior of Cord-Rubber Composites for Lab Prediction of Structural Durability of Aircraft Tires

by

Byung-Lip ("Les") Lee

ABSTRACT

An aircraft tire durability study is underway to investigate the deformation and fracture behavior of cord-rubber composites. This study will identify the important parameters responsible for the structural failure of aircraft tires by the use of analytical and laboratory prediction methods. These methods will also identify the interaction between material property degradation and damage accumulation in cord-rubber composites. Preliminary results using coupon specimens of tire carcass have revealed that prolonged static and cyclic loading sequences produce extensive interply shear deformation at the free edges resulting in cord-matrix debonding followed by delamination type failure. These loading sequences represent the circumferential tension in the footprint region of aircraft tires. It was also determined experimentally that a *fatigue endurance limit* can be established for cord-rubber composites. Analytical methods using finite element models of coupon specimens have demonstrated reasonable accuracy in predicting load-displacement response and interply shear strain variations. Future plans will include the correlation between the fatigue resistance data of composite specimens and dynamometer test results of actual tires.

Acknowledgements

I. INTRODUCTION

Compared with other types of pneumatic tires, aircraft tires are subjected to unusual combinations of speed and load (*1*). For example, the baseline tire used in this study is a 49X17/26PR bias construction and is rated at a speed of 358 km/hr (224 mph) and a load of 173 500 N (39 000 lbf). Extreme combinations of speed and load result in high cyclic frequencies, large deflections and significant heat generation due to hysteretic loss in aircraft tires. As confirmed by field experience, these conditions cause damage in critical sub-regions of tires such as *shoulder*, *bead*, *lower sidewall* or *tread* which eventually develops into catastrophic failures of whole tires. The failure of the carcass ply in the shoulder area, often called a ply separation, occurs in the form of *delamination* which involves crack propagation mainly in the rubber matrix and cord-matrix interface (Figure 1). So-called bead area cracking and lower sidewall break also involve crack propagation in the rubber matrix of carcass ply, but there are strong indications of *fiber fracture* as well. The processes of damage accumulation and structural failure of tire carcass in the shoulder, bead and lower sidewall areas are attributed to a combination of mechanical overloading and heat generation along with the resultant deterioration of constituent materials.

On the other hand, tread failure of aircraft tires is attributed to the presence of strong centrifugal force resulting from unusually high speed and high load requirements. Tread failure which starts in the form of circumferential cracks at the base of tread groove is caused by repeated groove flexing motion when so-called *standing wave* is present. Although the term derives from their stationary appearance to the observer, standing waves occur when the natural circular frequency equals a critical frequency (i.e. critical speed) with which a section of tire passes the ground contact position (Figure 2). The repeated deformation due to standing waves produces a significant heat build-up and thereby deteriorates material properties. When circumferential groove cracks caused by standing waves expose cord reinforcement of underlying carcass ply, they develop into a complete tread-carcass separation rather quickly leading to catastrophic failure of tires. Past research works (*1-3*) showed that rotational speeds of tires approaching critical speed should be avoided to prevent premature tread failure.

As reviewed so far, basic understanding of tire failure mechanisms has been established at least in qualitative sense. In the case of tread failure, its major cause, the occurrence of

standing waves can be avoided by a proper design of tires (1-3). However, as far as the carcass ply is concerned, the tasks of identifyng critical operating conditions (speed, load, underinflation and tire deflection) responsible for the failure and predicting the useful life expectancy of an aircraft tire are still difficult at best. Accelerated testing based on dynamometers provides a valuable means of evaluating the structural durability and life expectancy of aircraft tires. But the results of these dynamometer tests reflect merely the sensitivity of each specific tire design and construction to a specific combination of test conditions, unless underlying mechanisms of property degradation and structural failure are identified.

Past approaches (4-6) to characterize the property degradation of the aircraft tires have utilized coupon specimens cut from the carcass region. Clark et al (4) conducted a study which assessed the degradation of adhesion strength between the carcass plies of conditioned tires through the peel test of coupon specimens. His study also examined the time-temperature dependence of peel strength. Another study (5) is currently attempting to develop experimental procedures which affect the degradation of adhesion strength of the tire coupon specimens by subjecting them to preload, flexural fatigue and temperature conditioning. The durability analysis of Ref 4-6 provides useful information on failures caused by the deterioration of material properties such as ply-to-ply adhesion strength. Some estimates assume that tire failure occurs at a 50 percent reduction of carcass ply adhesion strength. However, these works did not show how the deterioration of material properties leads to damage accumulation and eventual structural failure of tire carcass. The research program described in this paper has begun and plans to achieve the laboratory prediction of the durability of aircraft tires based on the study of deformation, damage accumulation and fracture behavior of cord-rubber composites.

II. OBJECTIVES OF THE RESEARCH EFFORT

The study will investigate the mechanisms of carcass ply failure of aircraft tires with an emphasis on failure modes in the *shoulder* area. Appropriate test methodologies will be chosen to simulate ply delamination using cord-rubber composite specimens. By assessing interply shear deformation behavior of composites under static and fatigue loading, the study will identify the important parameters responsible for the damage accumulation and structural failure. This assessment will also define the interaction of damage accumulation and material property degradation. Analytical modeling efforts will be performed for the

prediction of interply shear strain in tire carcass composite specimens. The study will correlate the laboratory data of fatigue resistance of cord-rubber composites with the durability test results of actual aircraft tires.

III. EXPERIMENTAL STUDY OF COMPOSITES

In developing test methodologies for the simulation of *carcass ply delamination* type failures, the study used two kinds of lab coupon specimens press-molded from the calendered ply stocks. The first type of specimen consists of a cord-rubber composite simulating typical *bias aircraft tire carcass* using a +/- 38 deg reinforcement angle, 1260/2 nylon cord and a proprietary rubber compound based matrix. The second type, which is a *model composite*, uses a +/- 19 deg cord angle, steel wire cables of circular cross-section and a natural rubber-based compound matrix. The model composite coupon specimen was included in the study to allow a better determination of the failure modes, because the large cord diameter and the cord angle chosen maximize the interply shear strain, a major contributing factor in composite failure (7).

To avoid tension-bending coupling, the specimens were constructed with a symmetric ply lay-up. The end tabs were added to the specimens to prevent failure near the gripping region. All other specifications of composite coupon specimens are listed in Tables 1 and 2. To represent the circumferential tension of an aircraft tire in the footprint region, coupon specimens were subjected to both static and cyclic uniaxial tensile loading. Cyclic testing was performed under a broad range of load amplitude and at three different levels of frequency (speed) with a close monitoring of heat generation. In both static and cyclic testing, local strain was estimated by measuring the displacement of line markings drawn on the specimen edge as shown in Figures 3 and 4.

Under both static and cyclic tension, the composite specimens were found to exhibit a large *interply shear strain* (Figures 3, 5 and 6). As analyzed by numerous studies in the past (7-9), interply shear strain develops in angle-ply laminates when the constituent plies exhibit in-plane shear deformation of opposite direction but the action is prevented by mutual constraint due to interply bonding. Compared with the case of fiber-reinforced plastic composites, cord-rubber composites exhibit unusually high level of interply shear strain which results from the load-induced change of reinforcement angle allowed by extreme compliance of rubber matrix. The results of the relative displacement for two adjacent plies

show that, at an axial strain of 10 percent, an interply shear strain of 70 percent develops in the nylon cord-reinforced composites and 120 percent in the model composites as shown in Figures 5 and 6.

Above a critical value of interply shear (static) strain corresponding to approximately one-third of ultimate strength of the composites, localized failure is induced in the form of *cord-matrix debonding*. Debonding is started around the cut ends of reinforcing cords at the edge of the finite width coupons as shown in Figures 3 and 4. This phenomenon is justified since the maximum interply shear strain occurs at the edge of the specimen. The role of cut ends of cords as built-in defects in the process of damage initiation confirms detailed observations made by Breidenbach and Lake (*10,11*). The model composites, shown in Figure 4, reveal the progressive development of cord-matrix debonding into *matrix cracking* as the strain increases. At higher strain, the axial stress-strain curves for both nylon cord-reinforced and model composites exhibit strain hardening type response (Figures 7 and 8).

A new finding was that the above-described critical load for the onset of cord-matrix debonding constitutes a threshold level for semi-infinite fatigue life, i.e. *fatigue endurance limit*, of the composites. As shown in Figure 9, the S-N curve (maximum cyclic stress vs the number of cycles to failure) for the model composites has a distinct endurance limit. Preliminary results also indicate the same behavior for nylon cord-reinforced composites. The physical description of an endurance limit is that, with cyclic stresses lower than the limit, cord-matrix debonding and matrix cracking are never initiated nor developed. Under cyclic stresses exceeding the endurance limit, debonding and matrix cracking are progressively widened and developed into the delamination leading to gross fracture of the composites (Figure 4).

The observed process of damage accumulation was accompanied by a continuous increase of temperature and cyclic strain as shown in Figure 10. The *stabilization of temperature or strain* (zero rate of increase) was determined to be a good indication of infinite fatigue life (Figure 11). Either by degrading cord-matrix adhesion strength or by lowering the modulus of rubber matrix, the increase of temperature may contribute to the increase of cyclic strain. The increase of cyclic strain at a given stress amplitude will in turn produce more hysteretic heating which can further reduce fatigue lifetime of tire carcass composites. However, the possibility of this type of interaction involving the changes of cord-matrix adhesion strength or matrix modulus seems to be rather remote, since the measured rates of

temperature and strain increase are almost linear (Figure 10). Additional fatigue experiments of cord-rubber composites at elevated temperatures or after heat aging are planned to resolve the question.

IV. COMPUTER MODELING

Experimental study has shown that cord-rubber composite coupon specimens representing the aircraft tire carcass exhibit extensive interply shear deformation between constituent plies under circumferential tension, which eventually leads to delamination type failures. In order to investigate the variation of interply shear strain within composite specimens, a finite element model (FEM) using the ADINA computer code was created by co-investigator J. P. Medzorian (WRDC Flight Dynamics Laboratory, Vehicle Subsystems Div.). Linear elastic-orthotropic material elements were used to model the constitutive behavior of the cord-reinforced plies, while linear elastic-isotropic material elements were used for modeling of rubber layers. Since shear strains vary linearly throughout the solid finite elements, the model is restricted to only predicting linear variations of interply shear strain through the coupon thickness.

Under an uniaxial tensile load of 3 491 N, the axial displacement of the *model composite* coupon was analytically predicted to be 4.06 mm which was within 14 perecent error from the experimental value of 3.55 mm. The analytically predicted value of interply shear strain of 80 percent compares well with experimentally measured values of 65 to 85 percent. In the case of nylon cord-reinforced composite that simulates a *bias aircraft tire carcass*, the axial displacement under a uniaxial tensile load of 792 N was predicted to be 21.26 mm, which is very close to an experimental value of 21.59 mm. The agreement between the analytical predictions and experimental results was less satisfactory for the variation of interply shear strain. Under the load of 827 N, the predicted value of maximum interply shear strain was 13 perecent which is much less than experimentally observed values of 25 to 35 perecent. One cause of error could be mesh refinement. Since solid finite elements are limited to predict only linear variation of interply shear strain, many more elements are required in each layer to accurately predict non-linear shear strain variations.

V. CORRELATION WITH TIRE TEST RESULTS

To correlate the laboratory data of interply shear fracture resistance of cord-rubber composites with durability test results of actual aircraft tires, an extensive tire test plan has been developed jointly with co-investigators P. M. Wagner, G. J. Migut and J. P. Medzorian (WRDC Flight Dynamics Laboratory, Vehicle Subsystems Div.). The test program will use 49X17/26PR KC-135 bias aircraft tires. The durability tests will be accomplished on a 120 inch diameter dynamometer located in the Landing Gear Development Facility at Wright-Patterson Air Force Base. The tests will be performed with a systematic variation of *footprint load, inflation pressure* and *speed*. For each variation of conditions, carcass ply delamination as a dominant failure mode of tires will be confirmed by fractography. The study will examine the effects of footprint load, inflation pressure and speed on the *mileage to failure* (lifetime), *deflection* and *temperature history*. An establishment of a threshold load for damage initiation is planned to determine an *endurance limit* for the tires.

These results of tire durability will be examined in comparison with laboratory fatigue resistance data of carcass composite specimens. The composite data will include the effects of *stress amplitude, ratio of maximum to minimum stress* and *cyclic frequency* on the *number of cycles to failure, strain increase rate, heat-up rate* and *fatigue endurance limit*. By comparing how deformation and fracture properties are dependent on the same controlling parameters, the dynamometer test results of actual tires will be correlated with the fatigue resistance data of composite specimens. In addition to the parameter correlation study, coupon specimens will be extracted from the tires with various load histories and then subjected to static and fatigue tests to characterize the degradation of material properties.

VI. CONCLUDING REMARKS

The described research effort aims at the laboratory prediction of aircraft tire durability based on the study of deformation, damage accumulation and fracture behavior of cord-rubber composites. As a major step toward the goal, the study plans to identify critical operating parameters in terms of stress, strain and temperature history responsible for structural failure of tire carcass. The study will also define the mechanisms of damage accumulation and its interaction with material property degradation. Emphasis is placed on carcass ply *delamination* in the shoulder region of aircraft tires.

Experimental methods have been developed to simulate delamination process in cord-rubber composite specimens. Under static or cyclic loading which represents circumferential tension in the footprint region of tires, angle-ply composite coupon specimens were found to exhibit extensive interply shear deformation which eventually leads to delamination type failures. The study determined that a *fatigue endurance limit* can be established. The process of damage accumulation was always accompanied by a steady increase of temperature and local strain.

The analytical model developed by co-investigator J. P. Medzorian was able to predict the deformation behavior of cord-rubber composite specimens. The finite element model based on the computer code ADINA accurately predicted axial load-displacement relationship and the variation of interply shear strain through the thickness.

Finally a plan was developed jointly with co-investigators P. M. Wagner, G. J. Migut and J. P. Medzorian to correlate the laboratory fatigue resistance data of cord-rubber composites with dynamometer test results of actual aircraft tires. The dynamometer tests will assess the effects of footprint load, inflation pressure and speed on the lifetime, deflection and temperature history of the tires. The results will be compared with the data showing the dependence of corresponding composite properties on the same controlling parameters.

VII. RECOMMENDATIONS

The following recommendations can be made for the future study on fatigue fracture behavior of cord-rubber composites and laboratory prediction of aircraft tire durability:

(1) Using developed experimental and analytical methods, identify the critical parameters of stress, strain and temperature history that are responsible for damage accumulation and delamination failures of aircraft tire carcass composite specimens under fatigue loading conditions. Some stress history parameters include the mean stress, stress amplitude and frequency.

(2) Determine the importance of *hysteretic heating*, *dynamic creep* and *degradation of constituent material properties* in controlling the processes of damage accumulation and its development into delamination failures of tire carcass composite specimens.

(3) In order to evaluate the degree of damage accumulation, it is necessary to perform a detailed assessment of the fatigue-induced degradation of static strength and stiffness for the tire carcass composite specimens. A study is needed to determine if the fatigue lifetime of cord-rubber composites is dependent on the rate of damage accumulation but independent of damage history.

(4) Based on these experimental results, assess the applicability and limitations of classic approaches in the theoretical prediction of the fatigue lifetime of cord-rubber composites under the various combinations of load amplitude and mean load (Goodman's) and complex loading sequence (Miner's).

(5) Develop experimental methods to accurately measure the stress, strain and temperature history in cord-rubber composites as well as aircraft tires during simulated operations. Specifically, the measurement of interply shear strain and stress is needed.

(6) Develop test methodologies for the lab simulation of other types of failures processes besides shoulder delamination of aircraft tires. Other common types of failure processes include *bead* and *lower sidewall area failure*.

(7) Examine the effects of footprint load, inflation pressure and speed on the mileage to failure, deflection and temperature history of aircraft tires and assess the failure modes of tires based on fractography. Correlate these results with the fatigue resistance data of cord-rubber composites.

References

(1) J. H. Champion and P. M. Wagner, "A Critical Speed Study for Aircraft Bias Ply Tires", AFWAL-TR-88-3006 (1988).

(2) J. H. Champion, S. K. Clark and M. K. Hilb, "A Study of Vibrational Modes in Rolling Aircraft Tires", WRDC-TR-89-3092 (1989).

(3) J. Padovan, "On Standing Waves in Tires", *Tire Science and Technology*, Vol. 5, No. 2, p.83 (1977).

(4) S. K. Clark, "Loss of Adhesion in Cord-Rubber Composites", Report on Contract No. 83-P-80607, U.S. Department of Transportation, Transportation Systems Center, Cambridge, MA (1984).

(5) North Carolina A&T State University, "Study of the Behavior of Aircraft Tire Coupons under Various Loading Conditions", Report on Contract No. F33615-87-C-3413, U.S. Department of Air Force, Air Force Systems Command, Wright-Patterson AFB, OH (1990).

(6) Personal communications with the researchers in tire industry.

(7) J. D. Walter, "Cord-Reinforced Rubber" in Mechanics of Pneumatic Tires edited by S. K. Clark, U.S. Department of Transportation, Washington D.C. (1982)

(8) Interlaminar Response of Composite Materials edited by N. J. Pagano, Composite Materials Series Vol. 5, Elsevier (1989).

(9) J. L. Ford, H. P. Patel and J. L. Turner, "Interlaminar Shear Effects in Cord-Rubber Composites", *Fiber Science and Technology*, Vol. 17, p.255 (1982).

(10) R. F. Breidenbach and G. J. Lake, "Mechanics of Fracture in Two-Ply Laminates", *Rubber Chemistry and Technology*, Vol. 52, p.96 (1979).

(11) R. F. Breidenbach and G. J. Lake, "Application of Fracture Mechanics to Rubber Articles Including Tyres", *Philosophical Trans. Royal Soc. London*, Vol. A299, p.189 (1981).

TABLE 1

Specifications of Aircraft Tire Carcass Composites
(Cord: 1260/2 Nylon cord)
(Matrix: Proprietary rubber compound)

Cord Angle	-38, -38, +38, -38 deg
Cord Modulus	2.07 GPa (300X10^3 psi)
Matrix Modulus	5.51 MPa (800 psi)
Cross-Sectional Area of Cord	0.342 mm^2 (5.3X10^{-4} inch^2)
Total Width of Specimen	25.4 mm (1 inch)
Total Thickness of Specimen	6.35 mm (0.25 inch)
Specimen Gage Length	101.6 mm (4 inch)

TABLE 2

Specifications of Model Angle-Ply Composites
(Cord: Steel wire cable)
(Matrix: Natural rubber based compound)

Cord Angle	-19, -19, +19, -19 deg
Cord Modulus	207 GPa (30X10^6 psi)
Matrix Modulus	5.51 MPa (800 psi)
Cross-Sectional Area of Cord	2.06 mm^2 (3.2X10^{-3} inch^2)
Total Width of Specimen	25.4 mm (1 inch)
Total Thickness of Specimen	9.39 mm (0.37 inch)
Specimen Gage Length	101.6 mm (4 inch)
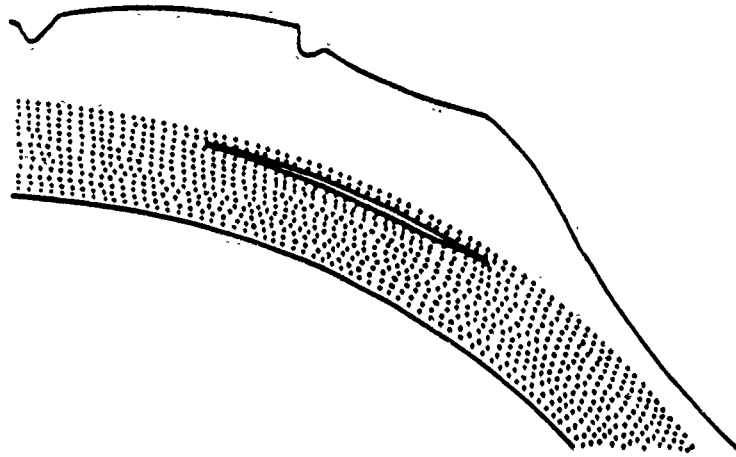
Figure 1

Schematic Drawing of Carcass Ply Separation in Shoulder Area of Bias Aircraft Tires

Figure 2

Standing Waves in Aircraft Tires (*Courtesy of M. H. Champion at Flight Dynamics Lab*)
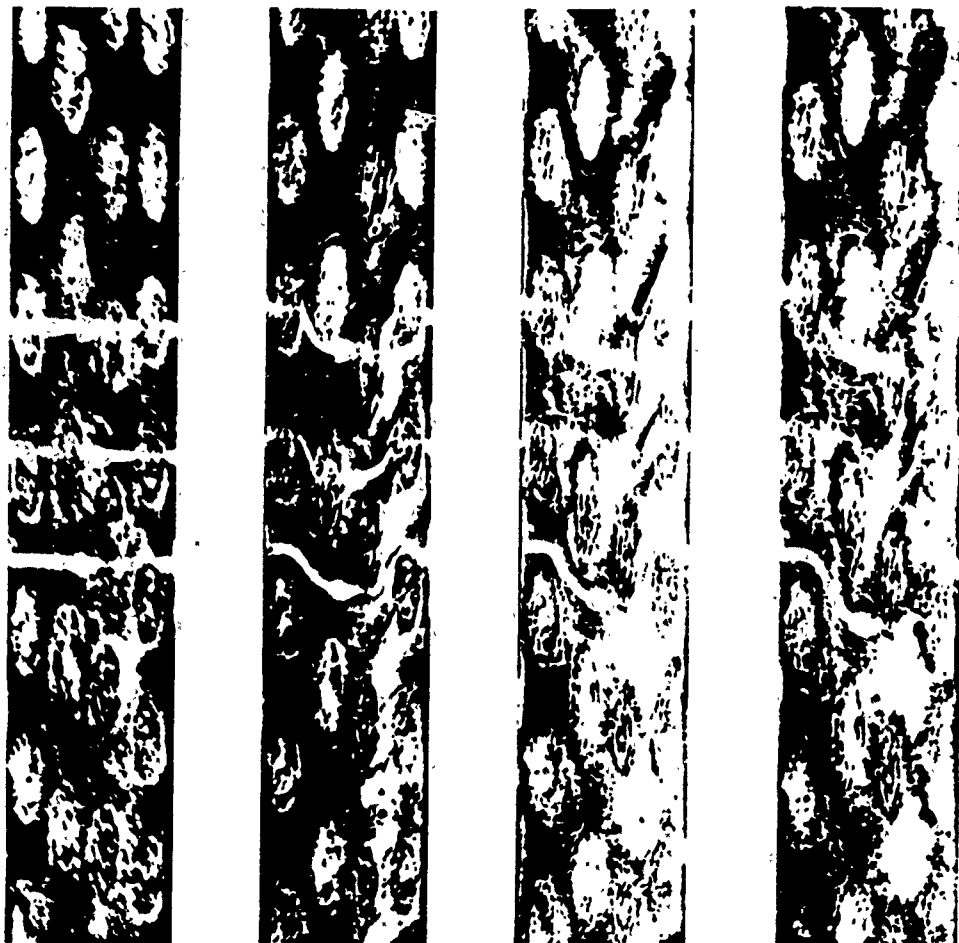
<u>Figure 3</u>

Cord-Matrix Debonding of Angle-Ply Composites Due to Interply Shear Deformation
(Model composites under tensile fatigue loading with 23 MPa maximum and 7 MPa
minimum stress: a- Before loading; b- Initial static loading; c- After 40 000 cycles; d-
After 460 000 cycles at 5 Hz frequency)

Figure 4

Other Failure Modes of Angle-Ply Composites Due to Interply Shear Deformation (Model composites: *a*- Cord-matrix debonding into matrix cracking; *b*- Matrix cracking into partial delamination)
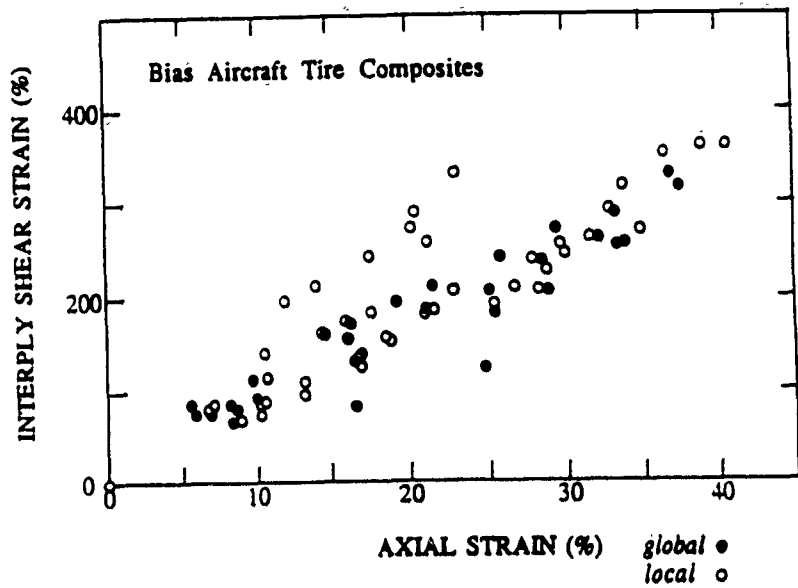
Figure 5

Interply Shear Strain vs Axial Strain for Aircraft Tire Carcass Composites
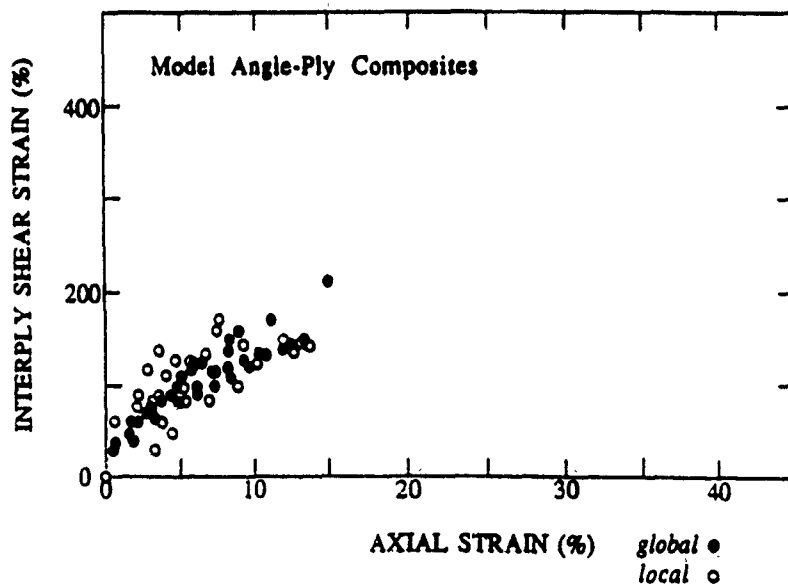
Figure 6

Interply Shear Strain vs Axial Strain for Model Angle-Ply Composites
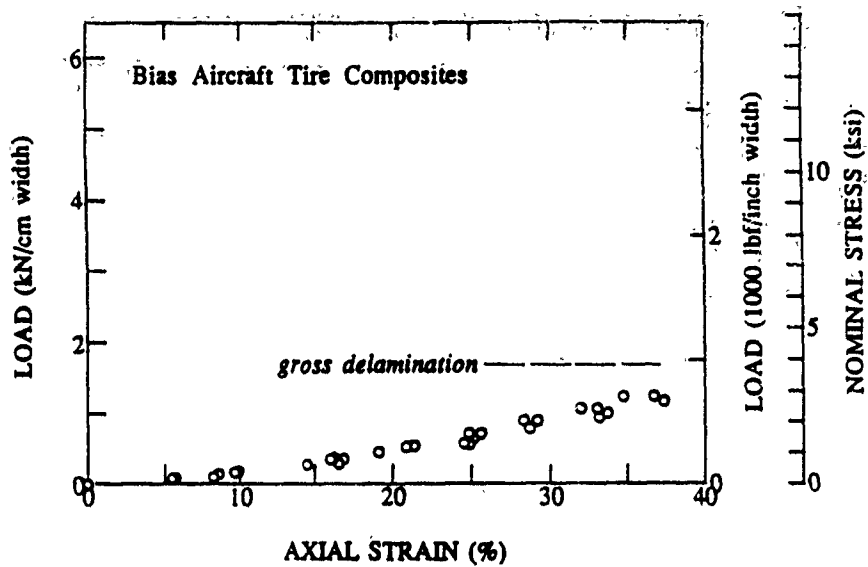
Figure 7

Load *vs* Axial Strain for Aircraft Tire Carcass Composites (deflection rate = 5 mm/min)
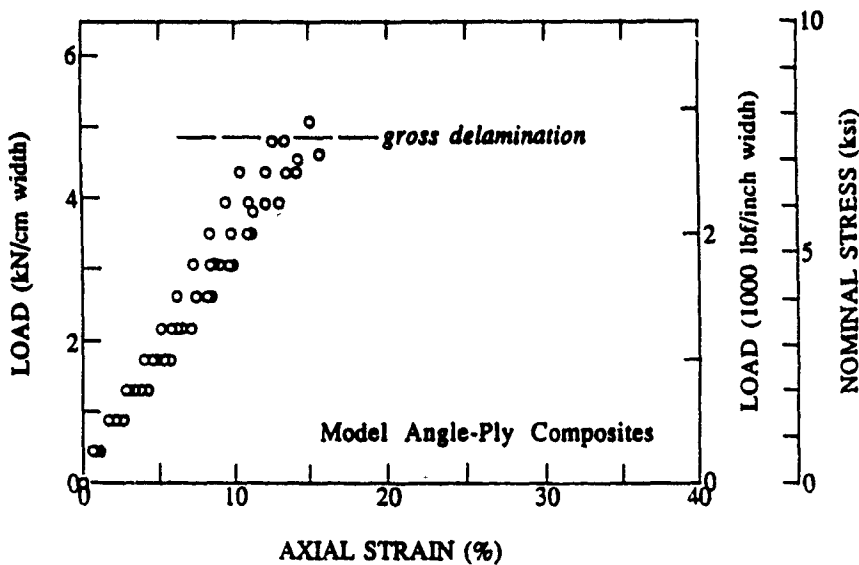
Figure 8

Load *vs* Axial Strain for Model Angle-Ply Composites (deflection rate = 1.25 mm/min)
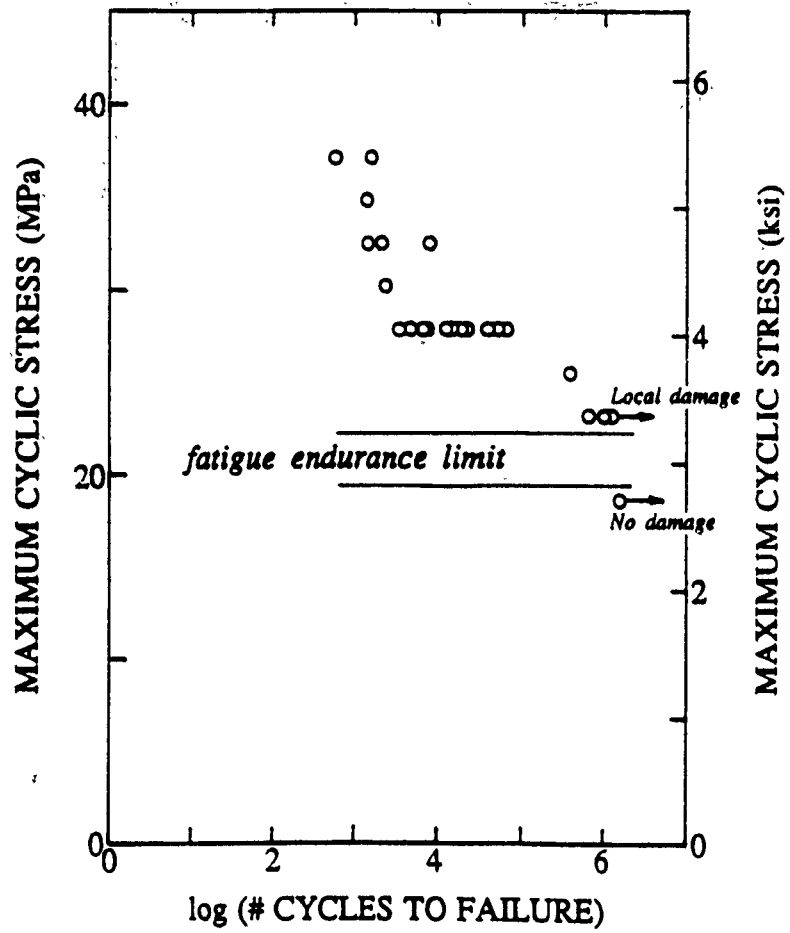
Figure 9

S-N Curve for Model Angle-Ply Composites (5 Hz freq.)
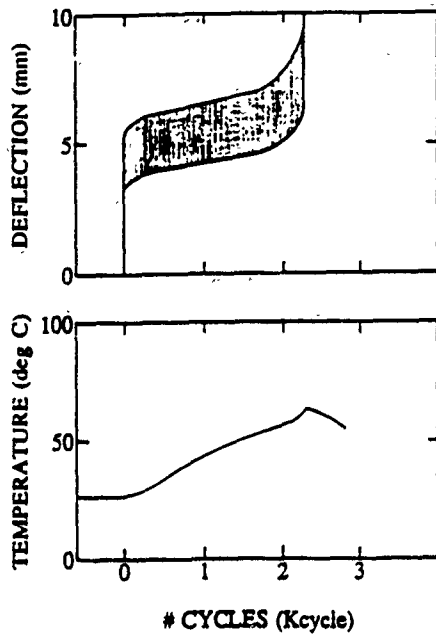
Figure 10

Continuous Increase of Deflection and Temperature under Cyclic Loading for Model Angle-Ply Composites (28 MPa max. stress, 7 MPa min. stress, 5 Hz freq.)
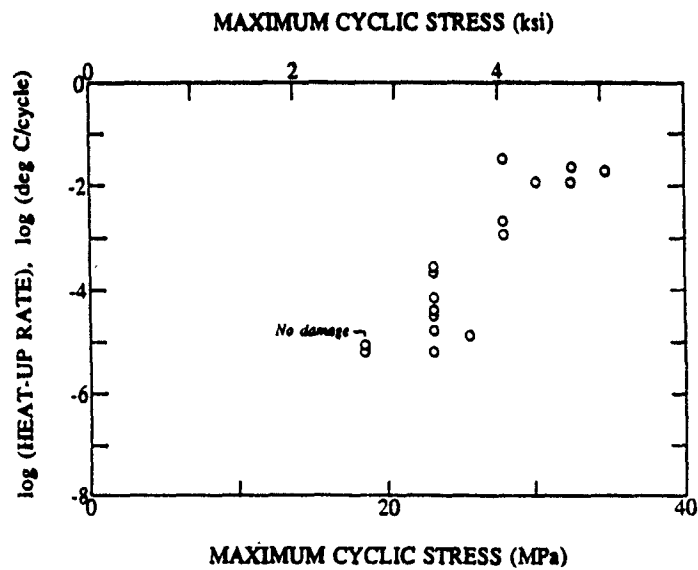
Figure 11

Rate of Temperature Increase *vs* Maximum Cyclic Stress for Model Angle-Ply Composites

Prof. Vernon Matzen

Technical Report submitted to Laboratory

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

DELAMINATION OF LAMINATED COMPOSITES

Prepared by: William E. Wolfe

Academic Rank: Associate Professor

Department and   Civil Engineering Department

University: The Ohio State University

Research Location: WRDC/FIBCA
                   Wright Patterson AFB
                   OH 45433

USAF Researcher: Dr. Raghbir S. Sandhu

Date: 30 Sep 90

Contract No: F49620-88-C-0053

DELAMINATION OF LAMINATED COMPOSITES

by

William E. Wolfe

ABSTRACT

In previous summer faculty research program appointments we have looked at the initiation of damage in laminated composites subjected to low velocity impact. A review of the literature as well as an analysis of our own tests showed that a sign,ficant mode of failure resulting from the impact event is delamination. A prediction of the extent of delamination requires an evaluation of interlaminar stresses and the material properties governing delamination.

The research performed during this summer's appointment followed two different lines. In the first effort, the theoretical studies begun in a 1989 mini-grant to determine the state of stress at each interface in a laminated composite plate subjected to a dynamic load were continued. In the second line of study, an analytical and experimental investigation of the tendency for delamination as predicted by the delamination moment coefficient originally defined by Sandhu was performed.

Acknowledgements

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of the Summer Faculty Research Program. Universal Energy Systems Inc. administered the project. Their help throughout the summer is appreciated.

My summer experiences with the Air Force research community continue to be rewarding because everyone in the Flight Dynamics Laboratory at WRDC is interested in making it so. I am grateful to all who have helped me during this past summer. I wish to give particular thanks to Mr George Holderby, group leader of FIBCA for most of the past four summers, who made the work environment so conducive to success and to Dr. Jim Olsen, the Chief Scientist of the Flight Dynamics Lab, who has somehow found the time to take a personal interest in this research.

I am particularly grateful to Dr. Raghbir S. Sandhu who has been my technical monitor in this and three previous summer programs. Dr. Sandhu so enjoys sharing his huge supply of research ideas that I am always disappointed to see the summer come to an end.

I. INTRODUCTION

The behavior of laminated composites that have been impacted at low velocities is of great concern to designers of high performance aircraft. These impacts which may result from a maintenance accident or debris impact during takeoff or landing can significantly degrade aircraft performance and even cause structural failure without detectible surface evidence of the event. A large number of researchers have studied the delamination problem both analytically and experimentally. In spite of considerable effort, current predictive capabilities are limited. The research conducted during this summer faculty appointment was a continuation of an ongoing effort, begun with a 1987 Summer Faculty appointment, to develop a predictive model for the response of laminated graphite epoxy plates subjected to low velocity impact (5,6). Theoretical studies begun by the author with a 1989 mini-grant (7), to determine the state of stress at each interface in a laminated composite plate that results from the application of a dynamic load were continued. In addition, an analytical investigation of the tendency for delamination using the delamination moment coefficient as originally defined by Sandhu (2) as the delamination criterion was conducted.

II. OBJECTIVES OF THE RESEARCH EFFORT

The work described in this report is a continuation of previous summer faculty research program appointments (5,6) in which the initiation of damage in laminated composites subjected to low velocity impact was studied. Those efforts showed that, during the impact event, a significant cause of failure is ply delamination. A prediction of the extent of delamination requires an evaluation of both the interlaminar stresses and the material properties governing delamination. The main objectives of this summer's research program resulted in two separate efforts. The first task was to continue the theoretical studies begun in a 1989 mini-grant to determine the state of stress at each interface in a laminated composite plate that results from the application of a dynamic load. The goal was to extend this work by deriving a self-adjoint form of the governing equations. This would enable solution schemes based on finite element methods to be developed. The second task consisted of a numerical investigation of the tendency for delamination as predicted by the delamination moment coefficient originally defined by Sandhu (2). The objective of this second effort was to identify ply orientations in which delamination was the sole mode of failure so that test specimens suitable for use in determining interlaminar strength properties could be designed.

III. RESULTS OF WORK EFFORT

3.1 Theoretical Studies:

There has been a considerable amount of research conducted over the past three decades on the behavior of composite materials. Much of the work has been in the area of laminated plate theory which has, in recent years, focussed on ply-by-ply analyses. Reviews of previous work have been presented by several investigators including the author and his students (Wolfe and Schoeppner (7), Schoeppner(4)), and the SFRP technical monitor of this project (Sandhu (2), Sandhu and Sendeckyj (3)). Wolfe and Schoeppner discussed the advantages of a stress based theory for describing the behavior of laminated composites. In Pagano's stress based theory (1), a linear variation of in-plane stresses is assumed. The transverse stresses are obtained by integrating the equilibrium equations. This theory has been shown to accurately model the high stress gradients near the free edge of a laminated plate while maintaining stress continuity at the ply interfaces. However, the stress based formulation as presented by Pagano does not allow for inertia effects. The extension of this method to include dynamic behavior has been the subject of the present study since such a method would be a valuable tool in the determination of stresses in laminated composites subjected to transient loading. In a recent report to UES, Wolfe and Schoeppner (7) described the initial steps in the development of an extension of Pagano's method to

include the effects of the inertia terms in the expression for the ply-by-ply stress field. During the SFRP an energy formulation was used to derive the governing equations. A self-adjoint system of the governing equations was obtained consisting of twenty-three field equations. The derived formulation is presented in more detail in Schoeppner (4).

3.2 Preliminary Numerical Studies:

Sandhu (2) in a review of a number of theoretical and numerical programs, reported that most investigators have identified interlaminar normal stresses as the primary cause of ply delamination. He observed that a number of attempts had been made to evaluate the magnitude of the interlaminar normal stresses due to inplane loads. In several of these studies expressions for delamination tendency were based on an assumed distribution of interlaminar normal stresses. In his work, Sandhu observed that these assumed distributions of interlaminar normal stresses were not necessary for determining the ply orientations most prone to delaminate. Sandhu defined a delaminating moment coefficient (DMC), and showed that delamination could be expected when the DMC reached a certain critical value. This delaminating moment coefficient given in terms of the average axial stress is presented as equation 1.

$$DMC_s = \frac{DM}{\sigma_o} = m(4m+n)\left(\frac{t^2}{2}\right)\frac{\sigma^\theta_y}{\sigma_o} \qquad \text{Eq. 1}$$

where:

DM = delaminating moment at the mid-surface of the laminate

m = number of 90^0 plies

n = number of θ^0 plies

t = ply thickness

σ^θ_y = transverse stress in the θ plies

σ_o = average axial stress

Subsequent work by Sandhu and Sendeckyj (3) identified the critical value of the DMC as being a function of material type, plate thickness and ply orientation. They observed that laminated graphite/epoxy composites with DMCs greater than 10×10^{-5} in^3 delaminated, whereas plates with ply orientations resulting in DMCs less than 8×10^{-5} in^3 did not delaminate. They speculated that the DMC could be considered to be a material parameter capable of defining the delamination tendency of composites of different material systems.

An analysis of the effects of different ply orientations on the value of Sandhu's delamination moment coefficient was performed for plates of several thicknesses with off-axis ply orientations varying from 0 to 90 degrees. Figure 1 shows the results of these calculations. It is apparent that, for the graphite/epoxy system studied, a critical angle for the off-

axis plies between 40 and 50 degrees can be defined. It is also clear from Figure 1 that the DMC and therefore delamination tendencies increase with increasing laminate thickness. Sandhu and Sendeckyj have shown that although coupons made with the angle of the off-axis plies slightly greater than 40 degrees delaminated, they also experienced axial strains greater than 4000 uin/in prior to the onset of delamination. Since the failure strain of the epoxy matrix is approximately 4000 uin/in, any delamination in these specimens would be associated with matrix cracking and not just a simple opening mode failure. Therefore it was determined that for the finite element studies to be conducted in this program, only laminates with the off-axis plies oriented at 40 degrees to the long axis of the test coupon would be examined.

3.3 Finite Element Studies:

Using the results of the above preliminary numerical studies, a finite element analysis of the specimens with the greatest tendency to delaminate without experiencing matrix cracking during in-plane loading was undertaken. In this section, we present the results of these studies which were conducted to identify the state of stress and thereby the tendency for delamination in graphite/epoxy plates of six different thicknesses. As was done in the preliminary studies, each plate was modelled as having two 90 degree plies located at the midplane of the specimen. The outer plies all consisted of fibers oriented at ±40 degrees. The number of pairs of ±40

degree plies placed symmetrically about the 90 degree plies ranged from one, $[\pm45/90]_s$, to six, $[(\pm45)_6/90]_s$. The results of the computer simulations of the tensile test for the thinnest coupon studied $[\pm45/90]_s$, and the thickest, $[(\pm45)_6/90]_s$, are shown in Figures 2 and 3. Note the clear tendency to peeling as well as stretching in each of the figures. The stress contours plotted for each of the coupons studied shows that the interlaminar normal stresses are considerably larger and the gradients steeper for the thicker specimens than for the thinner ones, thus verifying the finding of Section 3.2 that the likelihood of delamination increases with the thickness of the panel. The distribution of interlaminar normal stress normalized by the mean axial stress for each of the coupon thicknesses tested is presented in Figure 4.
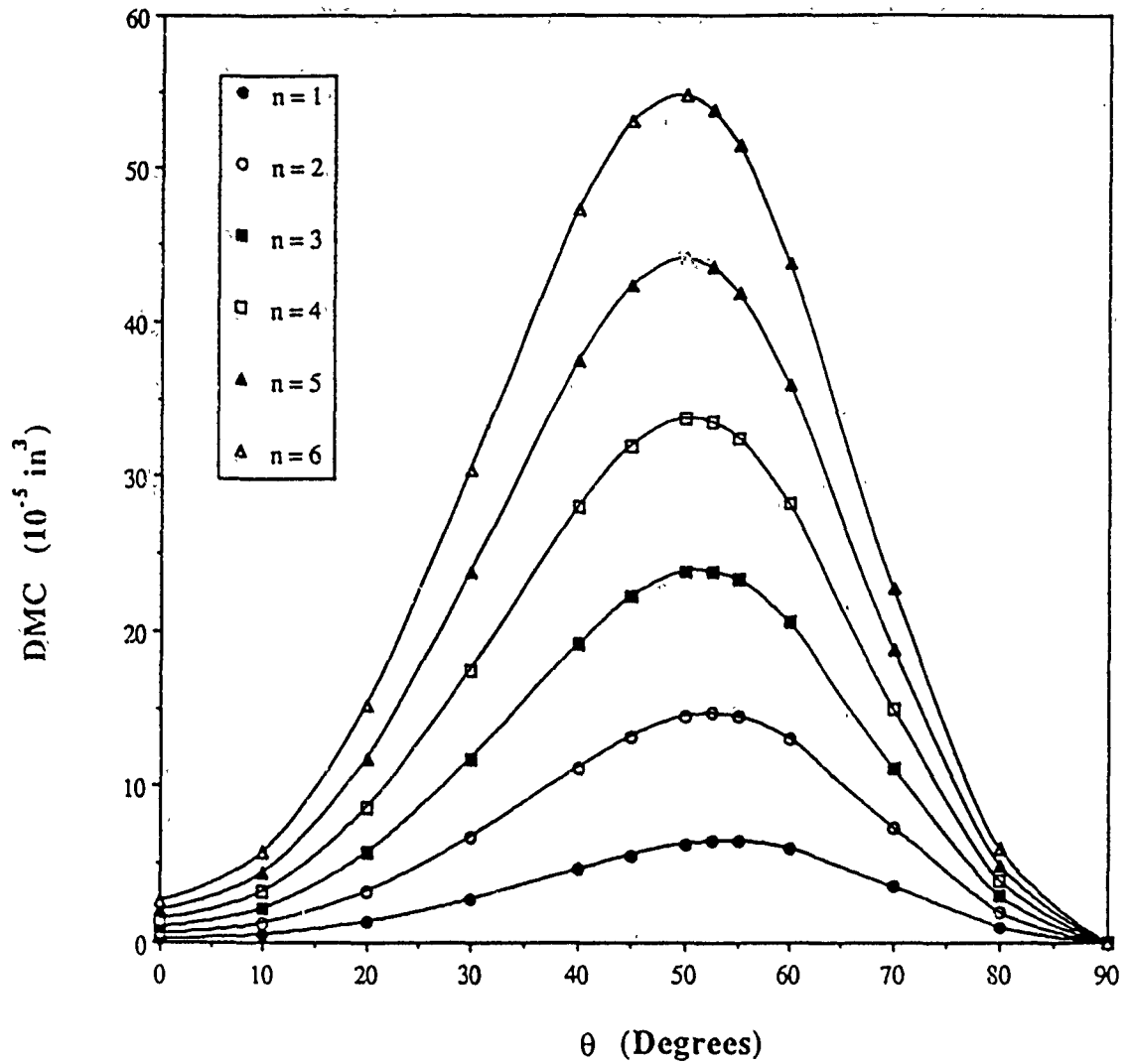
Figure 1: Variations of Delamination Moment Coefficient (DMC) in $[(\pm\theta)_n / 90]_s$ Laminates
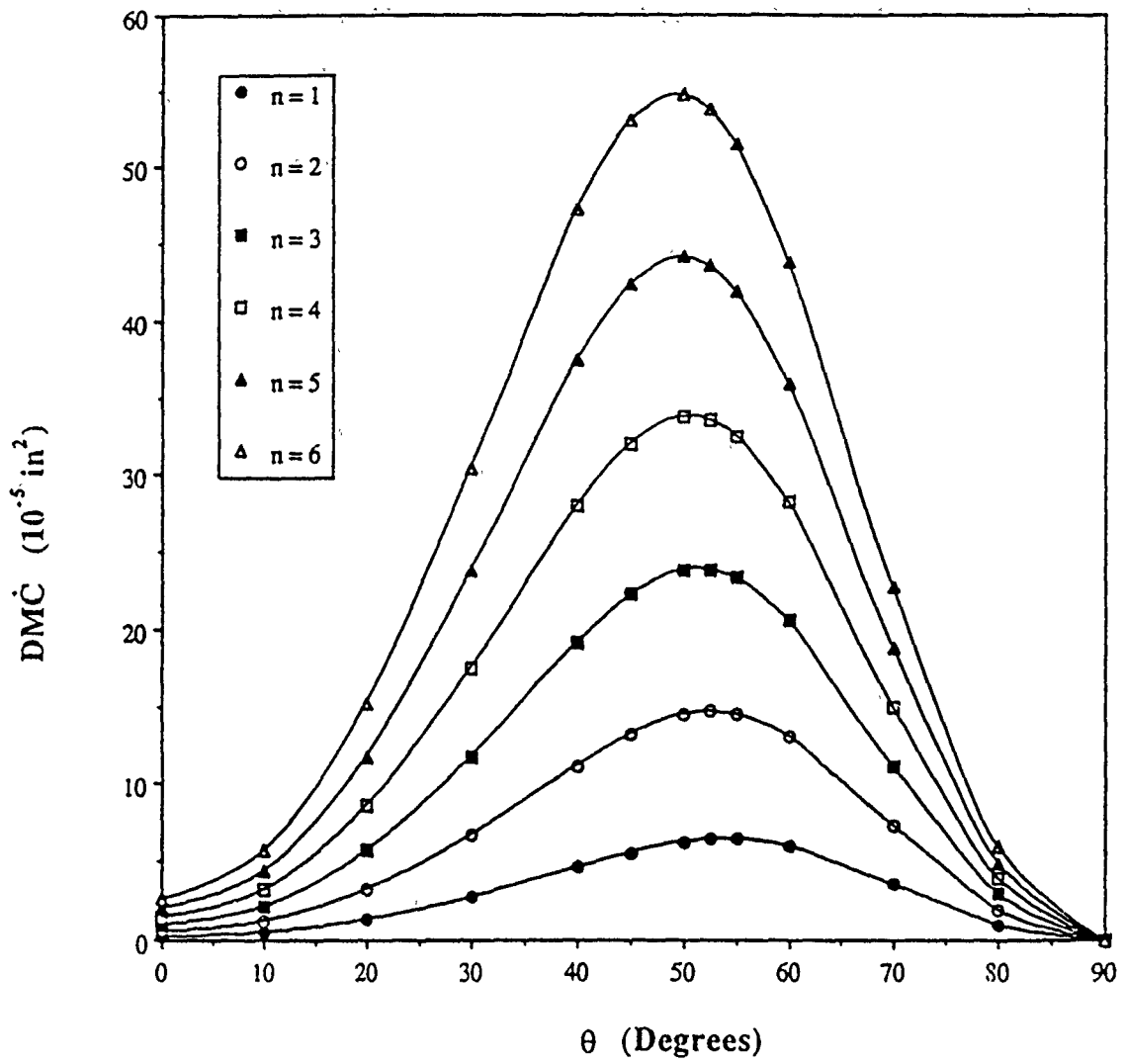
Figure 1: Variations of Delamination Moment Coefficient (DMC) in
$[(\pm\theta)_n / 90]_s$ Laminates
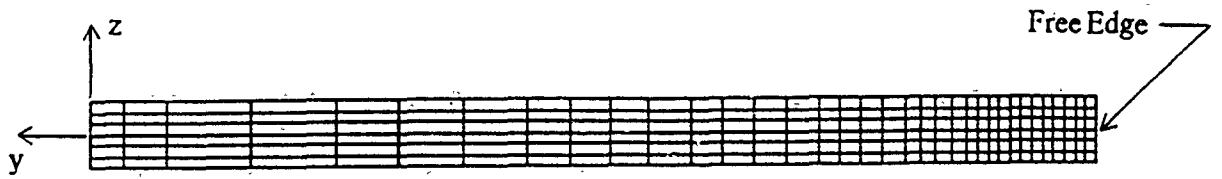
107-12

a. Undeformed Finite Element Mesh.

b. Deformed Finite Element Mesh under 100 psi Axial Stress.
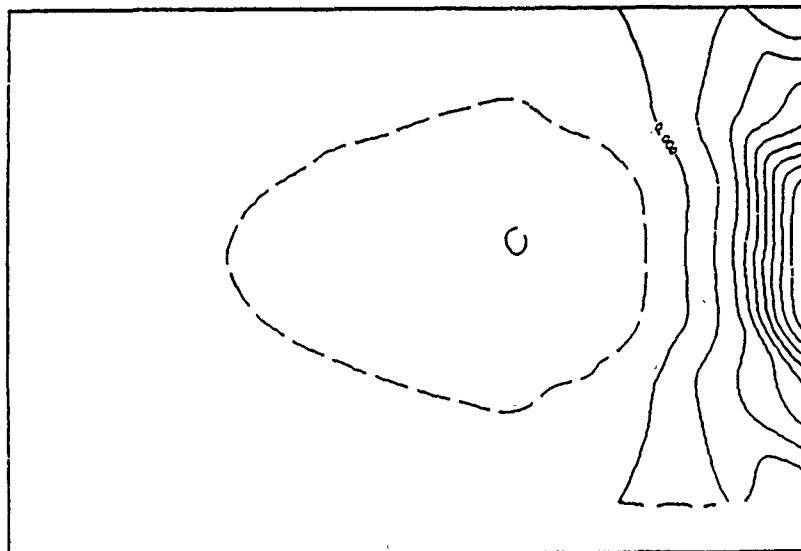
c. Interlaminar Normal Stress Distribution, σ_z, at Free Edge.

Figure 2: 6 Ply Laminate, $[\pm 40 / 90]_s$
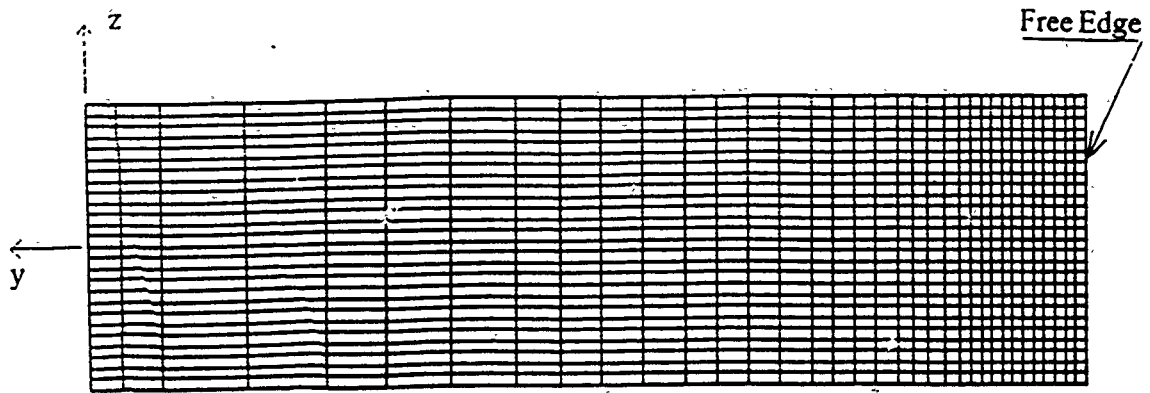
a. Undeformed Finite Element Mesh.
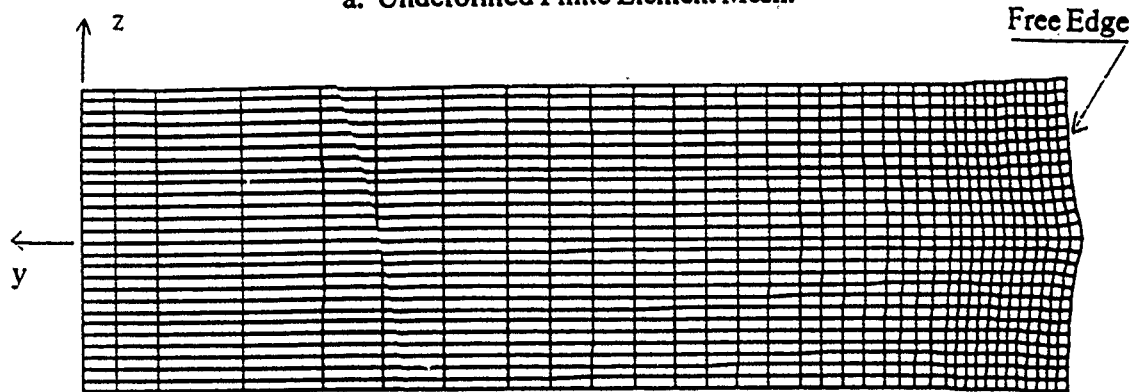
b. Deformed Finite Element Mesh under 100 psi Axial Stress.
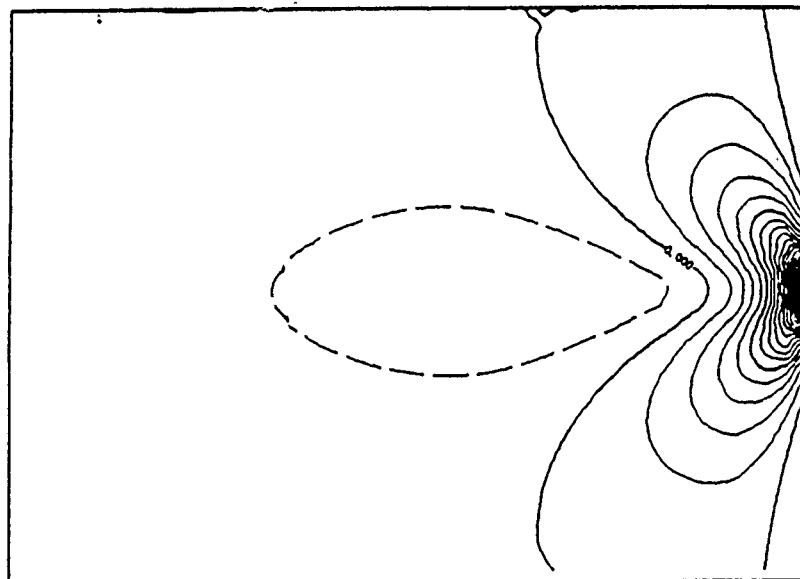
c. Interlaminar Normal Stress Distribution, σ_z, at Free Edge of Specimen.

Figure 3: 26 Ply Laminate, $[(\pm 40)_6 / 90]_s$
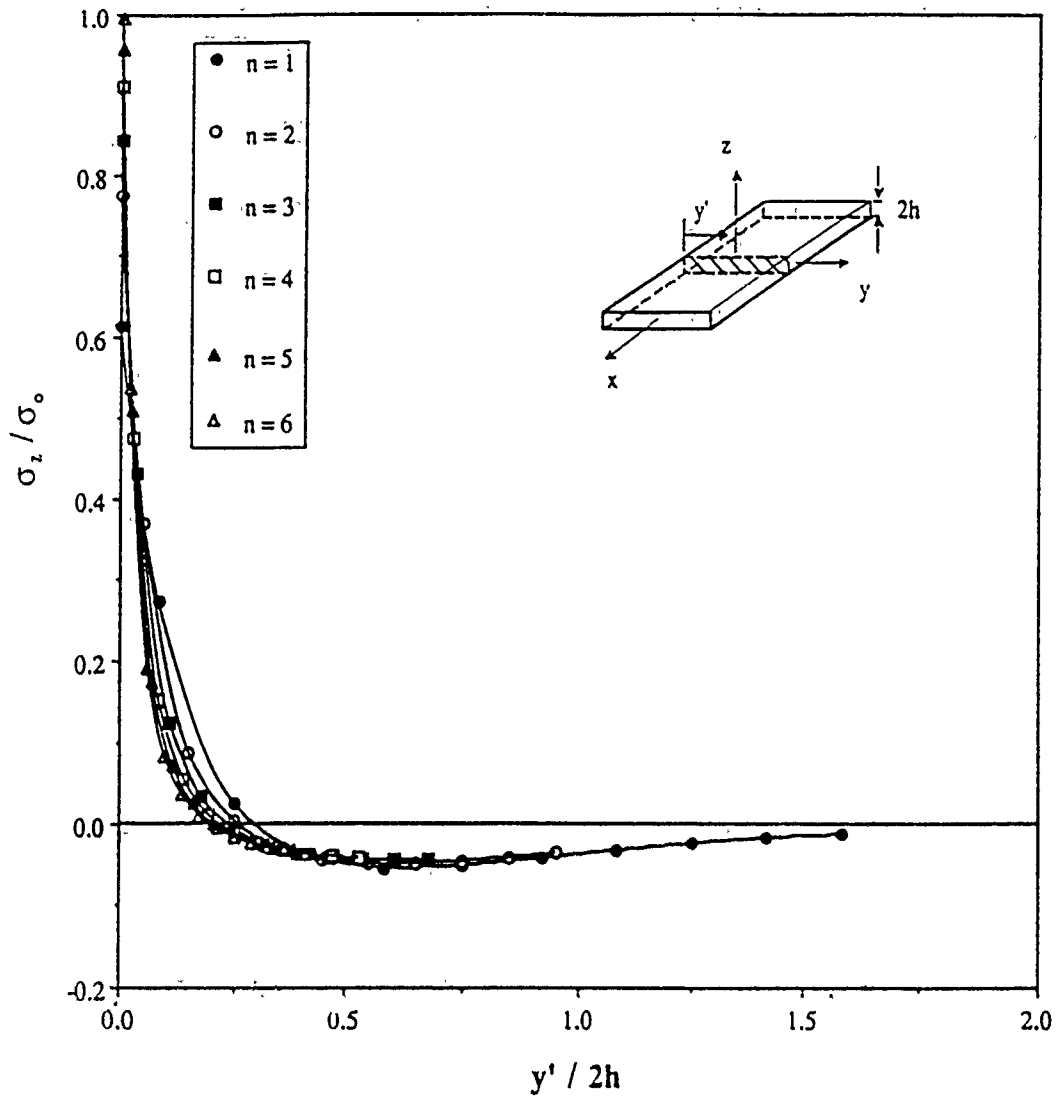
Figure 4: Interlaminar Normal Stress Distribution along Midplane of
$[(\pm 40)_n / 90]_s$ Laminates

IV. RECOMMENDATIONS

4.1 Theoretical Studies:

The development to date of extensions to the stress based theory clearly shows that this approach can be a significant improvement over currently available alternatives for describing the stress field in laminated composites subjected to time dependent loading.

The work presently being conducted should be continued in order to include:

1) the development of a numerical procedure (finite elements) capable of solving practical problems;

2) continued extension of the method to accommodate a wider range of problems including material failure and post impact behavior.

4.2 Experimental Studies:

The numerical studies have shown that an evaluation of a composite laminate's tendency to delaminate may be made by identifying a proposed material property, the delaminating moment coefficient (DMC). By designing laminates with the maximum DMC, different composite systems can be made to delaminate without simultaneously undergoing complicating additional failure modes.

Proposed follow-on studies should include:

1) an experimental program designed to verify the utility of the DMC in predicting the failure mode. For the graphite /epoxy system we studied in this effort, we suggest specimens of several thicknesses each with off axis plies oriented at $\pm 40^0$ be fabricated and tested;

2) additional numerical investigations of different materials to identify the effects of other system properties on the calculated value of the DMC.

REFERENCES

1. Pagano, N.J., "Stress Fields in Composite Laminates," *International Journal of Solids and Structures*, Vol. 14, 1978, pp. 385-400.

2. Sandhu, R.S., "Analytical-Experimental Correlation of the Behavior of 0^0, $\pm 45^0$, 90^0 Family of AS/3501-5 Graphite Epoxy Composite Laminates under Uniaxial Tensile Loading," *AFFDL-TR-79-3064*, Wright-Patterson Air Force Base, Ohio, May, 1979.

3. Sandhu, R.S. and Sendeckyj, G.P., "On Delamination of $(\pm\theta_m/90_{n/2})_s$ Laminates Subjected to Tensile Loading, "*AFWAL-TR-87-3058*, Wright-Patterson Air Force Base, Ohio, July, 1987.

4. Schoeppner, G.A., *Composite Laminate Stress Fields During Dynamic Loading*, Final Report, USAF-UES Summer Graduate Student Research Program, September, 1990.

5. Wolfe, W.E., *Low Velocity Impact of Graphite/Epoxy Plates*, Final Report, USAF-UES Summer Faculty Research Program, September, 1987.

6. Wolfe, W.E., *Damage in Graphite/Epoxy Plates Subjected to Low Velocity Impact*, Final Report, USAF-UES Summer Faculty Research Program, September, 1988.

7. Wolfe, W.E. and Schoeppner, G.A., *Low Velocity Impact of Composite Materials*, Final Report, Grant Nos. S-760-7MG-102 and S-210-9MG-082, Universal Energy Systems, Dayton, Ohio, February, 1990.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Experimental Identification of Internally Resonant Nonlinear Systems Possessing Quadratic Nonlinearity

Prepared by:	Lawrence D. Zavodney, Ph.D.
Academic Rank:	Assistant Professor
Department and	Engineering Mechanics
University:	The Ohio State University
Research Location:	Wright Research and Development Center
	Flight Dynamics Laboratory
	Structures Division
	Structural Dynamics Branch
	Sonic and Fatigue Group (FIBGD)
	Wright-Patterson Air Force Base
	WPAFB, Ohio 45433-6553

USAF Researcher:	Kenneth Wentz
Date:	8 September 1990
Contract No:	F49620-88-C-0053

Experimental Identification of Internally Resonant Nonlinear Systems Possessing Quadratic Nonlinearity

by

Lawrence D. Zavodney

ABSTRACT

The identification of MDOF nonlinear systems possessing internal resonance is discussed and possible solution strategies are proposed. It is shown that it is possible for nonlinear coupling between two internally resonant modes to go undetected during a modal analysis using broad-band random-excitation. If this type of nonlinearity is not identified, it is possible that the response to harmonic excitation may be many times larger than that predicted by the random response. Due to the combined presence of quadratic coupling and an internal resonance, it is possible for subharmonic and Hopf bifurcations, combination resonances, and subharmonic resonances to occur. Nonlinear coupling terms can also cause excited modes to become saturated. Chaotic responses were observed and documented. In this report, the results of experiments conducted at WRDC using conventional and state-of-the-art means for system identification are summarized.

ACKNOWLEDGMENTS

1. INTRODUCTION

The identification of nonlinear systems has been receiving increasing attention because most real structures have some degree of nonlinear behavior associated with them. Problems occur with traditional modal testing methods because most commercial analyzers run software that assumes linear behavior, resulting in a best-fit linear model. In general, a linear model is the least precise because it can always be improved with nonlinear terms to accomplish three things: 1) to improve the accuracy of the model, 2) to extend the range of useful results by permitting different amplitude ranges, and 3) to account for behavior unique to nonlinear systems that has no counterpart in linear theory. This third category is significant because many nonlinear identification techniques proposed today are based upon measuring deviation from linear behavior, hence there is an underlying link to linear behavior from the outset. These methods will be innacurate for nonlinear systems exhibiting uniquely nonlinear behavior. For work on the identification of systems exhibiting uniquely nonlinear behavior, see Nayfeh[1], Zavodney[2], and Zavodney and Shihada[3].

The significance of an internal resonance is that, when it is accompanied with a quadratic coupling term, it causes a parametric excitation to one mode when the other mode is externally excited (by driving the structure near resonance). Since the parametric excitation to one mode is caused by another mode in the structure, it is called an autoparametric resonance. This nonlinear resonance can cause subharmonic bifurcations, the saturation phenomenon, a subcritical instability, or amplitude- and phase-modulated motions (Hopf bifurcation). These behaviors are uniquely nonlinear phenomena; hence, one would not anticipate that linear based modal analysis techniques would give meaningful results. For a complete discussion of parametric resonances in mechanical systems, the reader is referred to Zavodney[4].

Quadratic coupling terms can appear in governing equations whenever there is a quadratic type material nonlinearity, geometrical asymmetry (the test model used in this report), or a symmetrical structure that is bent, buckled (beams and plates with in-plane edge loads), or curved.

In this report, conventional vibration analysis techniques are performed on a "simple" two-degree-of-freedom (2DOF) lumped-mass model that can be tuned to possess an internal resonance ($f_2 = 2f_1$). It is shown that randomly excited nonlinear structures can give seemingly accurate (linear) results but err significantly in predicting the response to a stationary harmonic excitation.

2. TEST MODEL

The experiments were performed on the structure shown in Figure 1; the length of the lower beam was five inches, its width was 0.5 inches, and its thickness was 0.030 inch. This structure can model a wing carrying a store or a space structure with an attachment such as a photovoltaic array. This structure was chosen for the following reasons:

1) the asymmetrical geometry causes the structure to have nonlinearly coupled modes, giving rise to quadratic coupling terms,

2) the ease of tuning the structure (by adjusting the length of the lower beam and the position of the second mass on the upright beam) to change the natural frequencies of the modes,

3) the possibility of seeing the nonlinear responses without the aid of a strobe light or dependance upon instrumentation because the amplitude of motion is large and the natural frequencies are low.

Two different configurations of the structure were considered; they were achieved by changing the location of the mass on the vertical beam. The two different configurations are:

1) the detuned case where the natural frequencies were 7.69 Hz. and 18.46 Hz. and

2) the tuned case where the freuencies were 5.20 Hz. and 10.60 Hz.

Tuning in this case refers to the natural frequencies being commensurable, i.e., in nearly a 2:1 ratio. This structure, subject to harmonic excitation, has been analyzed by several researchers including Haddow et al[5], Nayfeh and Zavodney[6], and Nayfeh and Balachandran[7]. The mathematical models developed show that the structure is dominated by quadratic nonlinearity that couples the modes. In the absence of an internal resonance, the structure can be analyzed as a linear system using conventional modal analysis techniques. However, when the system natural frequencies are commensurable, the system response is highly nonlinear and complicated; the two modes are coupled together even though the first two natural frequencies are well separated.

The response of the structure was measured from strain gages attached near the base of each beam. All displacements (and amplitudes) in the figures correspond to the displacement w_1 of mass m_1. For the amplifier gains chosen, one volt corresponded to one inch peak-to-peak displacement. Additionally, strain rosettes were cemented to the beams to measure shear strain to indicate if torsional motion was occurring. During the experiments, no out-of-plane motion was observed, verifying that even if the torsional mode was close to or coupled to one of the in-plane modes, it would not affect the dynamic response or be excited through any of the nonlinear coupling terms. The structure was clamped to a table excited by a 12,000-pound force electrodynamic shaker; the data was collected by a Zonic 7000 analyzer.

3. RANDOM EXCITATION

The most common modal analysis excitation method uses random noise. This method excites all of the structural modes simultaneously and hence reduces the test time. If the analyst suspects that his structure is nonlinear, he will often excite the structure twice: once with a low-level excitation and once with a high-level excitation. This dual level

excitation is often used as an indicator of nonlinearity because linear systems have two characteristics: the response amplitude is proportional to the excitation amplitude (the principle of proportionality) and the response to simultaneous multiple inputs is the sum of the system responses to the individual inputs (the principle of superposition). The principle of proportionality is verified by testing at two different levels of excitation; if the transfer function does not change with increasing levels of excitation, then the system is said to be proportional. Furthermore, if the system is superpositional, then there would be no difference in the Frequency-Response Function (FRF) obtained with a random excitation compared to a sinusoidal excitation.

The procedure to check for proportionality was applied to the structure. The response of the tuned and detuned structure to low-level and to high-level random excitation had very similar characteristics; the response of the tuned structure is shown in Figure 2. There is a very small change in the FRF, but not the type that is often encountered--a shift in and a broadening of the resonant peak when the excitation level is increased. In this case, there were neither frequency shifts nor broadening effects, so both the tuned and detuned structures "passed" the linearity test.

4. HARMONIC EXCITATION

Harmonic excitation methods require an exciter to impart harmonic forces or displacements to the structure; an electrodynamic shaker is often used for this purpose. An oscillator is used to drive the shaker's power amplifier. In the experiments that follow, harmonic excitation at both resonances, at a frequency near the sum of both resonances, and at a frequency near twice the second resonance were performed.

4.1 Detuned Structure

The frequency response of the detuned structure to a stationary base-displacement harmonic excitation near the first and second resonance is shown in Figures 3 and 4. These figures were obtained by sweeping the frequency up and then down through resonance while keeping the table acceleration constant. There is no significance in the different excitation levels for the modes. A 50 milli-g excitation at 5 Hertz corresponds to a peak-to-peak displacement of about 0.04 inch, approximately 0.8% of the length of the horizontal beam. Hence, we are dealing with extremely low-level excitation. The amplitude-response curves to an excitation of the second mode are shown in Figure 5; this curve was obtained by fixing the excitation frequency and slowly varying the amplitude up and then down. Ideal linear systems would exhibit a straight line indicating that the system response was proportional; the slight curvature is present because the data are from a real structure.

Another method to detect nonlinearity is to harmonically excite the structure at resonance and look for harmonics of the driving frequency in the response. If, for example, there is quadratic nonlinearity present, then resonances at twice the excitation frequency will appear; if there is cubic nonlinearity present, then resonances at three times the excitation frequency will appear in the response spectrum. Quadratic nonlinearity was detected in the response caused by harmonic excitation at resonance; there was a large harmonic at two times the excitation frequency. The strain gage mounted on the vertical beam was the most sensitive to this harmonic.

4.2 Tuned Structure

The frequency response of the second mode to a harmonic excitation is shown in Figure 6. As the frequency was slowly increased from below, the second-mode amplitude grew while the first mode remained trivial (because it was not excited). However, as the driving frequency approached the

108-8

resonance, the second mode coupled to and excited the first, resulting in larger steady-state amplitudes. As the frequency was increased further (still approaching the second resonance) the amplitudes of the modes actually decreased! Due to the slight detuning of the internal resonance, there was no jump phenomenon (turning point, saddle-node or flip bifurcation) on the left side of the resonance region as there was on the right side. However, there was a Hopf bifurcation that lead to amplitude- and phase-modulated steady-state responses for a small region of excitation frequencies.

The frequency response of the tuned structure at the first resonance is shown in Figure 7. It has behavior similar to that of the second resonance shown in Figure 6.

The amplitude response of a direct excitation to the second mode, as shown in Figure 8, reveals three distinct types of responses: the "linear" region, the nonlinear saturated response characterized by a constant-amplitude second mode and parametrically excited first mode (growing with the excitation), and modulated amplitudes and phases. The bounds on the mudulation are shown as a maximum and minimum value for the amplitude. The steady-state response in the modulation region can be easily visualized in a 3-D spectral history plot as shown in Figure 9. The Eigensystem Realization Algorithm (ERA)[8] can track the energy exchange that occurs between modes in the response by using a sliding window. Figure 10 shows the ERA sliding window results for the modulated response to a harmonic excitation. Once the nonlinear behavior is detected, it is clear that a linear model is incapable of predicting the response for various levels of excitation.

Although the frequency-response curves (Figures 6 and 7) and the amplitude-response curve (Figure 8) show the steady-state amplitude responses, whenever any system parameter, such as the amplitude or frequency of excitation, is varied slightly such that a bifurcation

boundary is crossed, the system does not immediately jump to the new steady
state as shown on the plots. Instead, the response grows with each cycle
of motion towards the steady-state amplitude. Figure 8 shows the
theoretical linear solution as dashed lines; if the system is in "steady-
state" below the critical value of excitation causing saturation and the
amplitude of excitation is slowly increased to a value above the critical
value, the system will behave as shown in Figure 11. The second mode will
initially grow to the unstable linear response, but as time increases, the
first mode (which is parametrically excited) begins its exponential growth.
As it grows, it takes energy from the second mode which is clearly seen.

4.3. Combination and Subharmonic Resonances

One of the unique properties of nonlinear systems is that they
possess resonances not found in linear systems. In this case, the
nonlinear coupling between the two modes allows additional resonances to be
present in the structural response to harmonic excitation that are not
excited with broad-band random excitation. In particular, one of the
quadratic coupling terms is responsible for a combination resonance
whenever the excitation frequency is close to the sum of the frequencies of
the two modes; another quadratic term is responsible for a one-half
subharmonic bifurcation which gives rise to a subharmonic resonance. When
the system simultaneously possesses an internal resonance, the modal
interactions observed in the direct excitation of either mode can also be
observed in the indirectly excited responses.

For example, when the structure is excited at a frequency two times
the second natural frequency, the response will consist of the forced
response (particular solution of the governing differential equation at the
driving frequency) and a subharmonic of the excitation frequency, which
happens to coincide with the second resonance. This behavior is shown in
Figure 12(a). It is possible to couple the first mode when an internal

resonance exists just by changing the excitation frequency or amplitude slightly, as shown in Figure 12(b). Furthermore, it is also possible for the coupled modes to modulate as previously seen, as shown in Figure 12(c). By zooming in on the resonance region of the spectrum, the subharmonics can be more easily seen; this is shown in Figure 13.

When the structure is excited by a frequency near the sum of the first and second resonant frequencies (a combination resonance of the sum type), the two modes will be simultaneously be excited. For certain ranges of system parameters, modulated responses can again be observed; Figure 14 shows the spectral history observed when the structure was excited with a stationary constant-amplitude acceleration. Using the ERA sliding window, the modal amplitudes are observed to modulate while the particular solution remains essentially constant, as shown in Figure 15.

5. IMPULSE EXCITATION

Modal analysis using impulse excitation is quite appealing for field work because it does not require a shaker, power amplifier, or signal generator. With an instrumented hammer, a complete modal analysis can be performed with a two-channel analyzer. If the structure is linear, the random excitation, impulse excitation, and swept sine excitation will give identical results. Hence, if the system is nonlinear, one would expect the different methods to produce different results depending upon the type of nonlinearity present in the structure.

For this experiment, m_1 was hit with a small hammer and the time history of the response was recorded. Modal interaction was present in the free response of the tuned structure; however, it would not be obvious to the untrained analyst. It is not uncommon to observe a beat phenomenon in a free response of a MDOF structure; typically all of the modes are excited, and as they decay at their damped natural frequencies, they will combine to form a composite response that may beat. However, unless the

individual modal amplitudes are extracted from the composite response
signal, the interaction would most likely be obscured. A transient
response of the structure to an impulse is shown in Figure 16; the modal
decomposition (extracted with the aid of the ERA sliding window) is shown
in Figure 17. This Figure shows the energy being exchanged from one mode
to another while both are simultaneously decaying. These results are
similar to those obtained by Zavodney and Pappa[9] for the free response of a
similar structure. If there was no internal resonance, each mode would
exponentially decay from its original amplitude.

6. TIME SERIES ANALYSIS

The method of delays was applied to the time series to construct
pseudo-phase planes. This method uses the time series representing the
displacement and constructs the "pseudo" velocity by using the position at
a later time. In this case, a delay of 0.025 seconds was used to construct
the phase plane. The response, of the tuned structure when the excitation
frequency was near the second resonance was analyzed because it had three
distinct responses (as shown in Figure 8). The phase portraits are shown
in Figure 18; the data show typical scatter about the equilibrium
attractor. Poincaré maps were obtained from the phase planes and are shown
in Figure 19. Only a portion of the modulated response is shown in the
phase plane because the complete cycle would completely fill in the bounded
region; the Poincaré map is constructed from the entire set of 60,000
points. To construct the Poincaré maps, the excitation signal was filtered
with an 8th-order low-pass ARMA filter before it was used to sample the
phase plane. The modulated response appears to be chaotic.

7. SUMMARY

1. Random excitation may give results that appear linear, making it possible for nonlinearity to go undetected. One reason why the random excitation would tend to obscure the autoparametric resonance is that the modal coupling is parametric in nature; parametric resonances feed on themselves and grow exponentially. Consequently, there is insufficient time for a parametric resonance to grow when the excitation is random.

2. Multi-level random excitation will not reveal certain types of nonlinearity because the nonlinearity in this case is not associated with the elastic stiffness; instead the nonlinearity is primarily from the geometry. Hence, increasing the excitation would not cause changes in the frequency response because the elastic stiffness is not strongly dependant upon the amplitude of response.

3. Low-level harmonic excitation can cause linear responses or nonlinear responses with constant or nonconstant amplitudes that modulate in time.

4. Subharmonic resonances and combination resonances may exist that could produce large responses.

5. Impulse responses may have modulated responses that are embedded and may not be obvious; they would need to be extracted to be identified.

6. CONCLUSIONS

1. System identification should include sufficient checks for nonlinear resonances caused by material properties, geometrical asymmetry, and stiffness nonlinearity, even when the data obtained from low and high level random excitation do not suggest any nonlinear behavior.

2. System identification should include all possible types of excitation a structure could possibly be exposed to; this would insure that undetected nonlinear resonances using one type of excitation would not appear later using a different type of excitation.

3. Slight structural modifications after a design has been qualified could cause nonlinear resonances to appear. These slight modifications could cause, for example, an internal resonance to occur, and provide the mechanism for the observed nonlinear behavior documented in this report.

4. If a structure demonstrates autoparametric coupling in response to a harmonic excitation, random excitation will effectively prevent the nonlinear coupling from occuring.

5. An identification technique that is uniquely suited to nonlinear behavior needs to be implemented for the more difficult structures. This technique would use the nonlinear resonances and jumps in the

system response to determine qualitatively and quantitatively the best

mathematical model for the structure.

8. REFERENCES

1. Nayfeh A.H., "Random Motion and Dynamic Response-Parametric Identification of Nonlinear Dynamic Systems," Computers and Structures, Vol. 20, No. 1, pp. 487-493, 1985.

2. Zavodney L.D., "Can the Modal Analyst Afford to be Ignorant of Nonlinear Vibration Phenomena," Proceedings of the Fifth International Modal Analysis Conference, 1987.

3. Zavodney L.D. and Shihada S.M., "The Identification of Nonlinearity in Structural Systems: Theory and Experiment," The Third Conference on Nonlinear Vibrations, Stability, and Dynamics of Structures and Mechanisms, Blacksburg Virginia, June 1990.

4. Zavodney L.D., "A Theoretical and Experimental Investigation of Parametrically Excited Nonlinear Mechanical Systems," Ph D. Dissertation, Virginia Polytechnic Institute and State University, 1987.

5. Haddow A.G., Barr A.D.S., and Mook D.T., "Theoretical and Experimental Study of Modal Interaction in a Two-Degree-of-Freedom Structure," Journal of Sound and Vibration, Vol. 97, pp. 451-473, 1984.

6. Nayfeh A.H. and Zavodney L.D., "Experimental Observation of Amplitude- and Phase-Modulated Response of Two Internally Coupled Oscillators to a Harmonic Excitation," Journal of Applied Mechanics, Vol. 55, pp. 706-710, 1988.

7. Nayfeh A.H. and Balachandran B., "Experimental Investigation of Resonantly Forced Oscillations of a Two-Degree-Of-Freedom Structure," International Journal of Nonlinear Mechanics, Vol. 25, pp. 199-209, 1990.

8. Juang J-N. and Pappa R.S., "An Eigensystem Realization Algorithm for Modal Parameter Identification and Model Reduction, Journal of Guidance, Control and Dynamics, Vol 8, pp. 620-627, 1985.

9. Zavodney L.D. and Pappa R.S., Unpublished ERA results obtained during a NASA Langley/ASEE Summer Faculty Research Program in 1988.
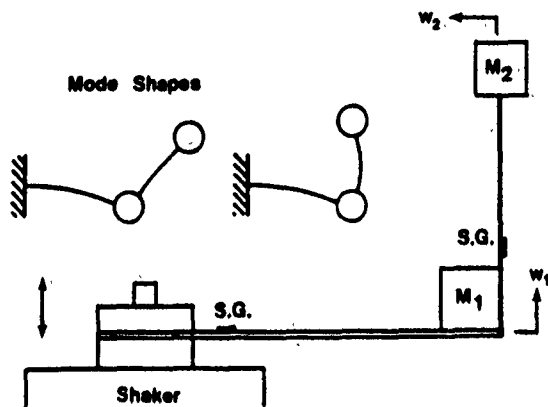
Figure 1. Two-degree-of-freedom model possessing quadratic nonlinearity; by adjusting the lengths of the beams, the structure can be tuned for a 2:1 internal resonance. The insets show the linear mode shapes.

Figure 3. Frequency-Response Function of the detuned structure for stationary sinusoidal excitation at the first resonance. The excitation level is 50 milli-g's, f_1 = 7.69 Hz, f_2 = 18.52 Hz.
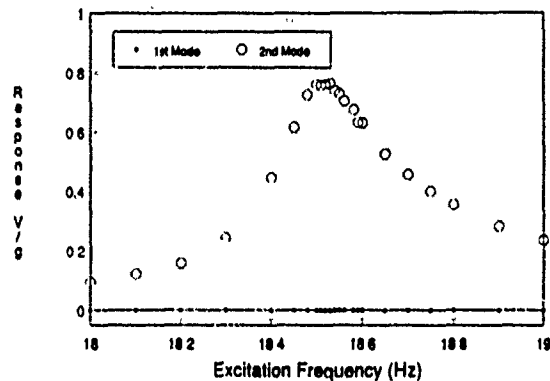
Figure 4. Frequency-Response Function of the detuned structure for stationary sinusiodal excitation at the second resonance. The excitation level is 21.2 mili-g's, f_1 = 7.69 Hz, f_2 = 18.52 Hz..

Figure 2. Frequency-Response Function of the tuned structure for random excitation at two different levels. The FRF of the detuned structure at two different input levels also correspond to each other.

Figure 5. Amplitude response of the detuned structure for harmonic excitation of the second mode showing typical linear behavior. An ideal structure's response would be perfectly straight.
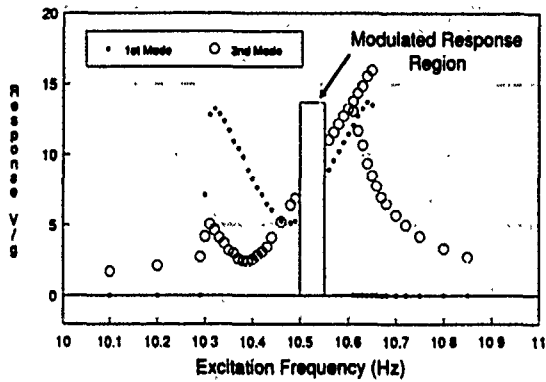
Figure 6. Frequency response of the tuned structure for a harmonic excitation near the second resonance. Note that the first mode is nonlinearly coupled to the second mode; for some frequencies there is no steady-state response. The excitation level is 21.2 milli-g's, $f_1 = 5.2$ Hz., $f_2 = 10.6$ Hz.
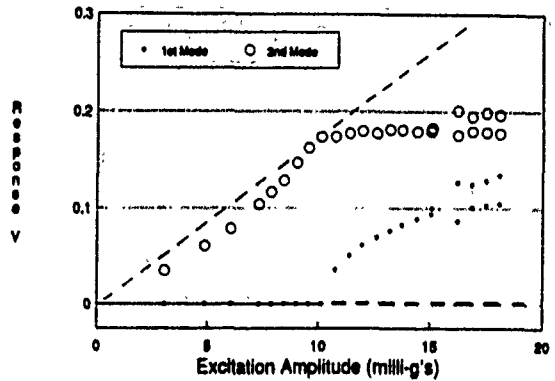
Figure 8. Amplitude response of the tuned structure for harmonic excitation of the second mode. Note that for low levels of excitation the response is similar to the linear detuned case; at a critical level of excitation the second mode is saturated and the first mode appears in the response. This behavior is similar to parametric resonances of nonlinear systems. At the second critical level of excitation a Hopf bifurcation occurs and the steady-state response consists of .modulated amplitudes and phases.
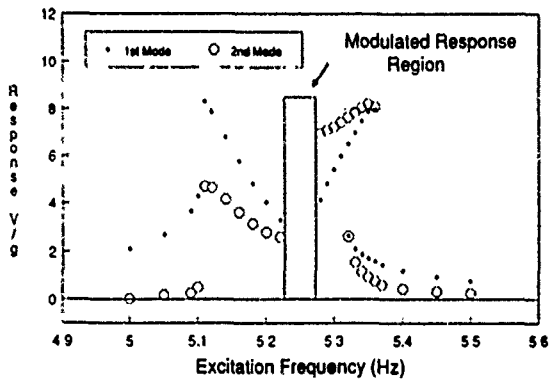
Figure 7. Frequency response of the tuned structure for a harmonic excitation near the first resonance. Note that the first mode is nonlinearly coupled to the second mode; for some frequencies there is no steady-state response. The excitation level is 50 milli-g's, $f_1 = 5.2$ Hz., $f_2 = 10.6$ Hz.
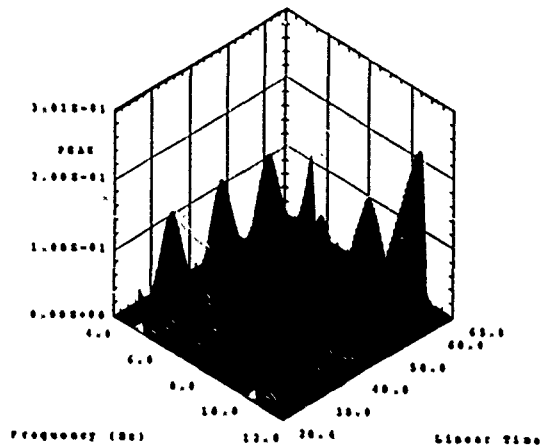
Figure 9. Waterfall plot showing the spectral history of the modulated response of the structure to a harmonic excitation at the second resonant frequency.
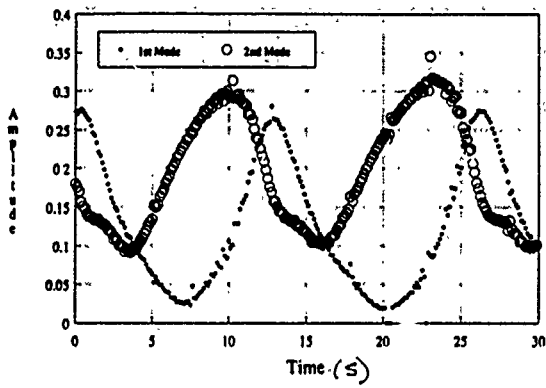
Figure 10. Modal amplitude modulation extracted from the time series data using the ERA sliding window. See Figure 9 for the associated spectral histories.
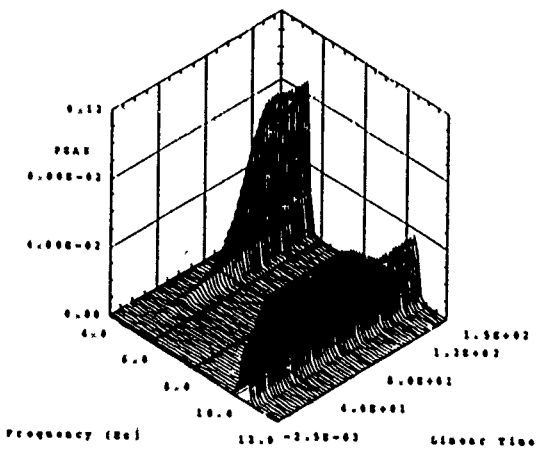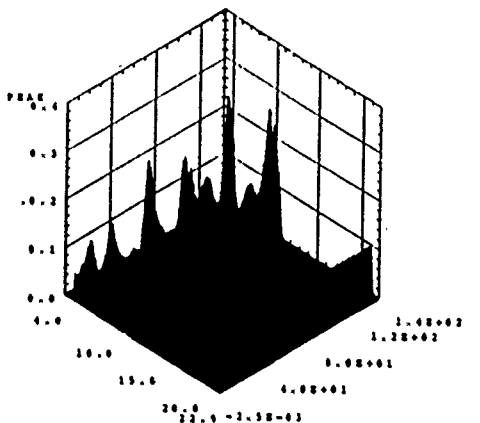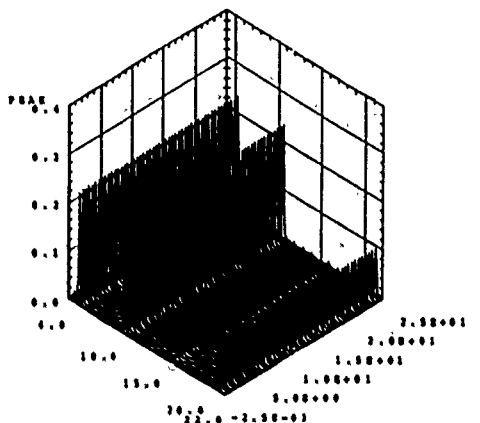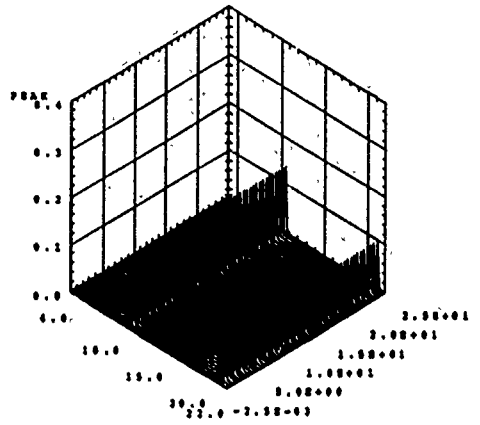
Figure 11. Spectral history of the system response to a direct excitation of the second mode. Initially, the system is below the critical value causing saturation. When the amplitude of excitation is increased above the critical value, saturation occurs. The response initially goes to the unstable linear solution but after sufficient time the stable nonlinear solution appears.

Figure 12. Waterfall plot showing the spectral history of the subharmonic resonance of the tuned structure when (a) only the second mode is excited, (b) the first mode is coupled to the second mode and achieves steady-state amplitudes, and (c) the modes interact and exchange energy.

Figure 15. Modal amplitude variation with time showing a modulated combination resonance; this data was obtained by an ERA sliding window. The diamonds represent the particular solution.
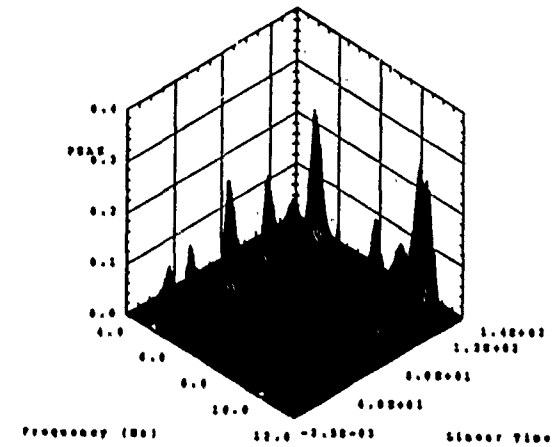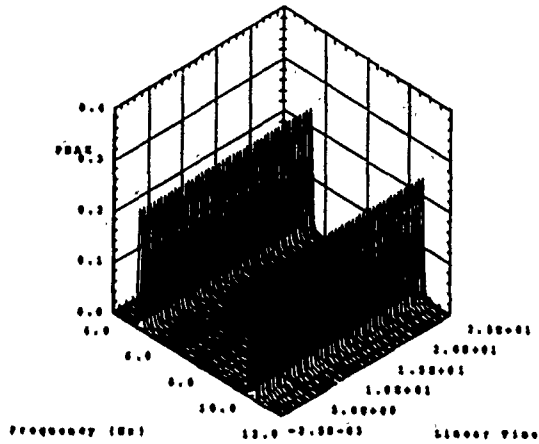
Figure 13. Waterfall plot showing expanded frequency scales for: (a) Figure 12(b), (b) Figure 12(c).
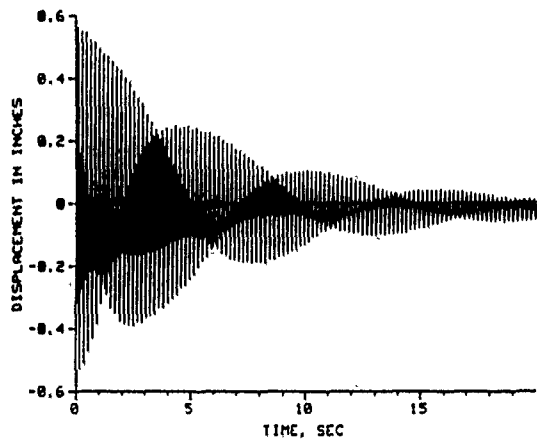
Figure 16. Impulse response of the tuned structure. See Figure 17 for the amplitudes of each mode extracted from this data.

Figure 14. Waterfall plot showing the spectral history of the modulated subharmonic response of the structure to a harmonic excitation at a frequency twice the second resonant frequency.
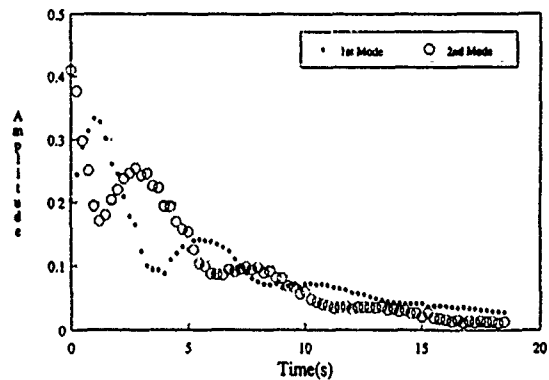
Figure 17. Modal amplitude response extracted from the time history shown in Figure 16 using the ERA sliding window.
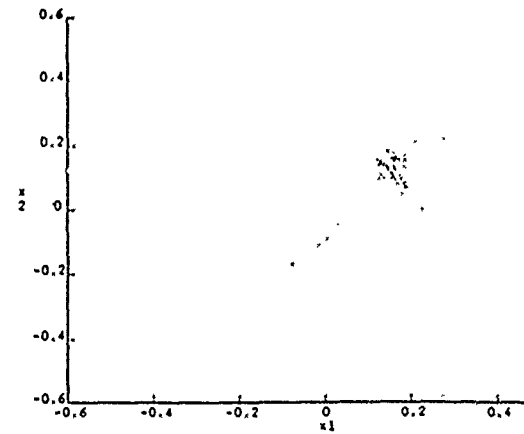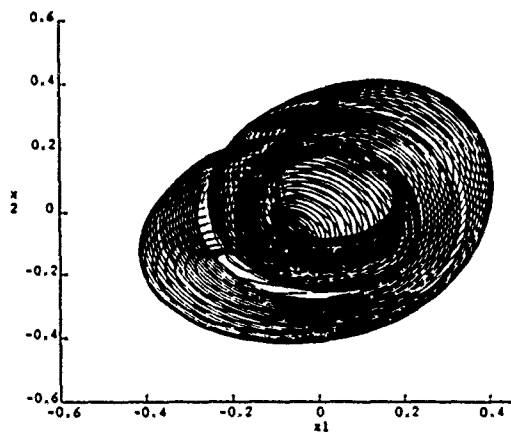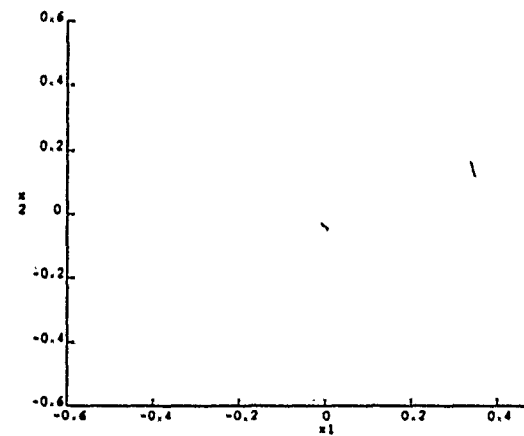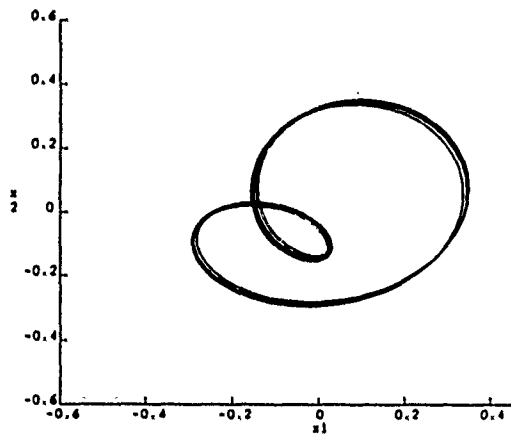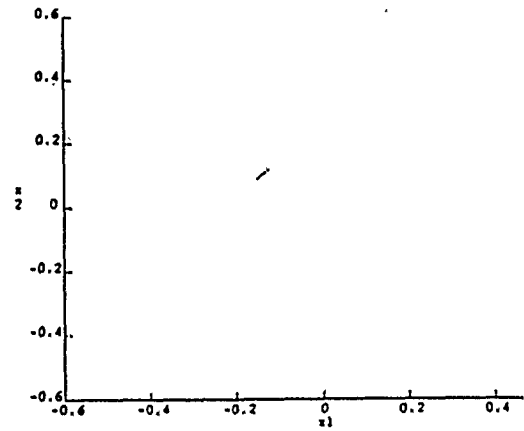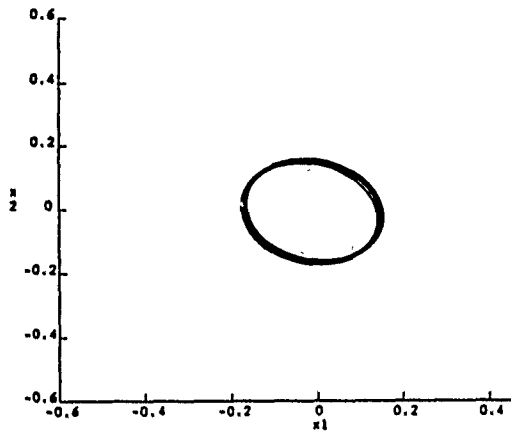
Figure 18. Pseudo phase portraits of a direct excitation of the second mode (of the tuned structure) constructed from the displacement using the method of delays: (a) harmonic linear response, (b) subharmonic bifurcation, and (c) chaotic modulated response. See Figures 6 through 10.

Figure 19. Poincare maps of the phase portraits shown in Figure 18: (a) single point showing periodic response with the excitation, (b) bifurcated response showing a period 2 response, and (c) a strange attractor showing chaotic behavior.

108-20

1990 USAF-UES SUMMER FACLTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

The In-situ Laser Deposition of High Tc Superconducting Thin Film.

Prepared by: Donald D. W. Chung, Ph.D

Academic Rank: Associate Professor

Department and   Materials Science

 University: San Jose State University

Research Location: USAF WRDC/MLPO

USAF Researcher: Patrick M. Hemenger

Date: 27 Aug 90

Contract No.:   F49620-88-C-0053

The In-situ Laser Deposition of High Tc Superconducting Thin Film

by

Donald D. W. Chung

ABSTRACT

ArF excimer laser ablation of an $YBa_2Cu_3O_{7-x}$ target pellet in 100 mTorr of O_2 ambient was used to deposit thin superconducting films on $SrTiO_3$ and MgO substrates at 650 - 780 °C. The as-deposited 0.6 - 0.9 μm thick films were superconducting, without further high-temperature annealing. Cooled to ambient temperature in-situ for 1.5 hours in flowing oxygen gas, the films showed complete diamagnetism and zero resistance up to 89 K with a critical current density of 5×10^5 A/cm^2 in zero magnetic field at 81 K. Low angle X-ray diffraction analysis showed that all the films were highly oriented with the C-axis perpendicular to their surface. Smooth surface morphology was observed in all films.

Acknowledgements

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. Universal Energy Systems must be mentioned for their concern and help to me all administrative and directional aspects of this program.

My experience was rewarding and enriching because of many different influences. Dr. Patrick Hemenger provided me with support, encouragement, and a truly enjoyable working atmosphere. The help of Tim Peterson was invaluable overcoming many technical roadblocks. Dr. Iman Maartense's interest in every phase of this project truly served as source of stimulation. The encouragement and help of Dr. Terry Murray were deeply appreciated.

I. Introduction:

Considerable progress has been made in the field of thin film preparation of high critical temperature superconducting materials both for fundamental research and possible application in electronic devices. Such film preparation is being attempted by many film deposition techniques. Among them, one promising method has been by mean of laser ablation of a bulk target superconducting materials in vacuum or under the low oxygen pressure to deposit the film. Since the low temperature processing will be essential for making sophisticated electronic devices containing the superconducting films the laser deposition method for such film preparation are especially attractive in the potential for producing films with required heat treatment in oxygen at only modest temperatures. Among the various types of laser processes short pulsed laser deposition has provided excellent epitaxial films with shallow resistance transition widths and high critical currents. Various researchers have obtained superconducting films without high temperature post-annealing using a background O_2 ambient in the vacuum chamber.

The elimination of the post-annealing process is extremely important in terms of growing multiple layers of different materials for device applications. Recently the in-situ processing of superconducting films at substrate temperatures of 650 - 750°C has been reported by a variety of techniques. The introduction of oxygen jet into the laser, ablated plasma was found to be a key for in-situ processing of superconducting thin films at 650°C. Further

reduction in deposition temperature have been reported by the incorporation of an oxygen plasma. So far, epitaxial film growth of $YBa_2Cu_3O_{7-x}$ on $SrTiO_3$ has yielded films with the highest current densities, the sharpest resistance transitions, and in general, the best overall properties.

Superconducting thin film processes and control programs are particularly interested in space application for detectors and sensor devices in Materials Laboratory of MLPO of the USAF. A number of sensor devices have been demonstrated with high Tc superconducting thin films; namely, superconducting quantum interference devices (squids), tunnel junction, and a fast nonlinear switch for noise discrimination in digital circuits. Special attention is directed to the optical and infrared detection using these new materials.

My research interest have been in the area of application of variety of thin film deposition techniques in the investigation of electrical, magnetic, and structural properties as function of processing parameters. My work on the deposition of $YBa_2Cu_3O_{7-x}$ superconducting film by RF and DC magnetron sputtering in which the principle of deposition is complementary to the laser ablation deposition, which is in principle, the same physical vapor deposition, and characterization of the film in terms of electrical, magnetic and structural properties contributed to my assignment to the superconductor group; Materials Laboratory, MLPO division of the USAF.

II. OBJECTIVES OF THE RESEARCH EFFORT

The thin film processes of high Tc superconducting $YBa_2Cu_3O_{7-x}$ compound can be divided into what require a high temperature annealing subsequent to deposition, and what require no further annealing, or low temperature annealing in oxygen to achieve optimum superconductivity. Low temperature deposition techniques which do not require further annealing are, by far, the most promising technique for producing films for applications because a better film morphology is usually obtained. The pulsed laser deposition is on of the most successful methods of achieving low temperature films in which it is capable of processing very smooth films with high transition temperature and critical current densities. However, further investigation is called for in this process in order for these to be a definitive evaluation of this technique with respect to applicability of the grown film properties as practical device materials.

During my 1989 SFRP several film processing data was generated and collected. The important deposition conditions were determined since the film processing data could not be analyzed in time. Research in this project resumed at San Jose State during January of this year under the sponsorship of an AFOSR Mini Grant. The significant processing conditions had been established prior to the 1990 SFRP at Wright-Patterson Air Force Base. Preparation of high Tc superconducting thin films are now routinely performed.

My assignment as a participant in the 1990 Summer Faculty Research Program (SFRP) was to continue to study and determine the process and the properties for depositing high quality temperature superconducting films and assess parameters of the laser deposition conditions for optimum properties which would be required for suitable device structure applications such as electronic detectors and sensor materials.

III.

High quality thin $YBa_2Cu_3O_{7-x}$ film has been produced by laser ablation technique. In many cases an amorphous film is deposited and a post annealing step up to $900^\circ C$ is necessary to form the crystalline superconducting phase. However, a good orientation of these films, a minimization of grain boundaries, and the prevention of inter diffusion are important points in order to achieve high critical current densities. All of the above is easier to obtain with an in-situ growth process, as has recently been demonstrated by several research groups.

For our experiments, we used a ArF excimer laser with 193 nm wavelength, 20 Hz, and 80 mJ per pulse. The laser beam was focused on a rotating, sintered $YBa_2Cu_3O_{7-x}$ pellet in a vacuum chamber. The target material which was evaporated perpendicularly to the substrate surface was then deposited on the substrate at a distance of about 60 mm. The substrate was heated up to $850^\circ C$ for cleaning and then cooled to the deposition temperature. Oxygen can be added

both the deposition and cooling processes. The chamber was evacuated by a turbomolecular pump; the base pressure was less than 10^{-6} Torr. Single crystalline $SrTiO_3$ and MgO with $< 100 >$ orientation were used as substrate materials. The polished substrates were cleaned in an ultrasonic bath in trichloroethylene, acetone, methanol, and de-ionized water, in that order and heat treated at 850°C for half an hour in-situ before the deposition. The deposition was carried out in about 30 minutes with a typical deposition rate of 4 Å/sec, and the resulted in film thicknesses ranged from 0.6μm to 0.9μm. Immediately after the deposition, pure oxygen was introduced into the deposition chamber. The films were then cooled down in flowing oxygen gas to ambient temperature in 1.5 hours.

X-ray diffraction analysis was performed with a slow scan diffractometer. The films were further examined by four-probe resistivity and A.C. susceptibility measurements. For the measurements of resistivity, the films were patterned by sputtering bridge paths 5mm length and 100 μm width. Four silver contacts were evaporated onto the sample and bonded with gold wires.

To determine the desired film processes, we first investigated the film properties as a function of several deposition parameters such as the substrate temperature, oxygen pressure, and target-substrate distance. The film properties were then evaluated by A.C. magnetic susceptibility measurement and X-ray diffraction analysis. The in-situ film growth behavior was depending sensitively on the

substrate temperature and the oxygen pressure. Films deposited at a substrate temperature lower than 600°C were semi-crystalline and had to be annealed to 900°C to form the desired crystal structure. At a substrate temperature higher than 650°C the film had grown in-situ crystalline. For the epitaxial growth with the C axis perpendicular to the substrate plane, oxygen pressure of 100 mTorr and a substrate temperature between 730 - 780°C were optimal. The X-ray diffraction pattern indeed showed a highly < 001 > oriented film grow expitaxially on both $SrTiO_3$ and MgO.

The quality of the C-axis orientation was also documented by observing the (00n) peaks from the x-ray diffraction pattern. It has been reported in the literature that the occupation of the oxygen sites in the lattice is strongly correlated with the C-axis parameter and the critical temperature. Films quenched after the deposition showed superconductivity but with low Tc components. Samples cooled slowly, generally showed high quality superconductivity. Although most of the films were exposed to the substrate temperature above 750°C, little interdiffusion between the substrate and the film had taken place. For the majority of the in-situ film the resistivity behavior was metallic. We have observed critical transport current density values Jc above 10 A/cm^2 at 81 K.

In conclusion, we have succeeded in preparing epitaxial $YBa_2Cu_3O_{7-x}$ films with high critical current densities by laser ablation in an easily reproducible one step process. No additional oxygen plasma source was needed for the in-situ crystalline growth.

IV. RECOMMENDATIONS:

The in-situ deposition of high critical temperature superconducting films ($YBa_2Cu_3O_{7-x}$) on $SrTiO_3$ < 100 > substrate by an ArF excimer laser have been successfully accomplished without additional system modification. However, the evaluation of data showed a room for improvement in the future deposition condition by implementing the existing system modification in the following areas:

a. The actual substrate (sample) temperature measurement must be established by installing a new substrate holder stage (including heat source) perhaps with a silver block which may be rotated during deposition.

b. The laser power and pulse rate during deposition must be controlled and optimized.

c. The experimental condition for low temperature deposition must be established for device applications.

d. The oxygen content of the film after the deposition and cooling process must be monitored in junction with the film property change.

e. The high dielectric constant and loss tangent of $SrTiO_3$ substrate would limit its practical utility, particularly in high-frequency microelectronic applications. Some other types of substrate such as NdGaO, LaGaO, or LaAlO which have good dielectric constants at high frequencies must be established for laser deposition conditions.

REFERENCES

Bauerle, D., Laser-Induced Formation and Surface Processing of High-Temperature Superconductors, Appl. Phys., 1989, Vol. 48, pp. 527-542.

Koren, G., Gupta, A., and Baseman, R.J., Role of Atomic Oxygen in Low-Temperature Growth of $YBa_2Cu_3O_{7-x}$ Thin Films by Laser Ablation Deposition, Appl. Phys. Lett., 1989, Vol. 54(19), pp. 1920-1922.

Mogro-Campero, A., et, al., Epitaxial Growth and Critical Current Density of Thin Films of $YBa_2Cu_3O_{7-x}$ on $LaAlO_3$ Substrate, Appl. Phys. Lett., 1989, Vol. 54(26), pp. 2719-2721.

Singh, R.K., et. al., In-situ Processing of Epitaxial Y-Ba-Cu-O High Tc Superconducting Films on (100) SrTiO and (100) YS-ZrO Substrates at 500-650°C. Appl. Phys. Lett., 1989, Vol. 54(22), pp. 2271-2273.

Witanachchi, S., Kwok, H.S., Wang, X.W., and Shaw, D.T., Deposition of Superconducting Y-Ba-Cu-O Films at 400°C without Post Annealing, Appl. Phys. Lett., 1988, Vol. 53, pp.234-236.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICER OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

<u>FINAL REPORT</u>

AM1 CALCULATIONS ON RIGID ROD POLYMER MODEL COMPOUNDS

Prepared by:	John W. Connolly
Academic Rank:	Professor
Department and	Chemistry
University:	University of Missouri-Kansas City
Research Location:	WRDC/MLBP Wright Patterson Air Force Base, Ohio 45433-6533
USAF Researcher:	Douglas S. Dudis
Date:	11 Aug. 90
Contract.	F49620-88-C-0053

AM1 CALCULATIONS ON RIGID ROD POLYMER MODEL COMPOUNDS

by

John W. Connolly

ABSTRACT

Using AM1 semi-empirical Molecular Orbital calculations, conformational energies were obtained for structures designed to model the rigid rod polymers, poly(p-phenylenebenzobisoxazole), PBO, poly(p-phenylenebenzobisimidazole), PBI, and poly(p-phenylenebenzobisthiazole), PBT, including examples in which the phenylene group is mono and dimethlylated. Minimum energy torsional angles and barriers to rotation can be understood in terms of steric factors and disruption of pi-electron delocalization. The model system used shows that when adjacent segments of the polymer chain are mutually perpendicular, the barrier to rotation is less than the thermally available energy at 300K.

Acknowledgements

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. The Assistance of the personnel of Universal Energy Systems, especially Mr. Danishek, with the every-day aspects of this project is appreciated.

The atmosphere in the Computational Chemistry Group of the Polymer Branch at the Materials Laboratory was made enjoyable by the enthusiasm and good cheer of Doug Dudis. The patience and help of the other members of the group, especially Todd Yeates and Jacque Henes is gratefully acknowledged.

I. INTRODUCTION:

The rigid rod polymers poly(p-phenylenebenzobisoxazole), PBO, and its sulfur analog, poly(p-phenylenebenzobisthiazole), PBT, are the focus of the U. S. Air Force's "Ordered Polymers" Program, which has been established to develop low-weight, high-performance materials for military and aerospace applications. These materials have been found to exhibit exceptional specific strength and , thermooxidative stability and environmental resistance when made into films and fibers (1).

The Computational Chemistry Group of the Polymer Branch of the Materials Laboratory at the Wright Research and Development Corporation is investigating the properties of PBO and PBT model compounds, primarily through the use of semi-empirical molecular orbital calculations. A wide range of polymer properties, including electronic, optical and mechanical, can be modelled using the appropriate calculational technique.

My research experience in the use of mainframe computers both for NMR spectra simulation and for semi-empirical molecular orbital calculations on organometallic compounds contributed to my assignment to the Computational Chemistry Group.

II. OBJECTIVES OF THE RESEARCH EFFORT:

There have been two recent AM1 molecular orbital (MO) calculations on the conformation energy of PBO and PBT model compounds, (2,3).The structures used in these calculations are shown in Figure 1. In both cases rotational barriers for PBO of 5.0 Kcal/Mol and 1.7

Kcal/Mol for PBT were reported. However a molecular modelling by Farmer, as yet unpublished, indicated greater polymer chain flexibility than is compatible with the above results. Consequently, it was decided to calculate conformational energies for PBO and PBT model compounds in which rotation about adjacent carbon-carbon bonds in the polymer backbone need not have the same barrier height.

A

B
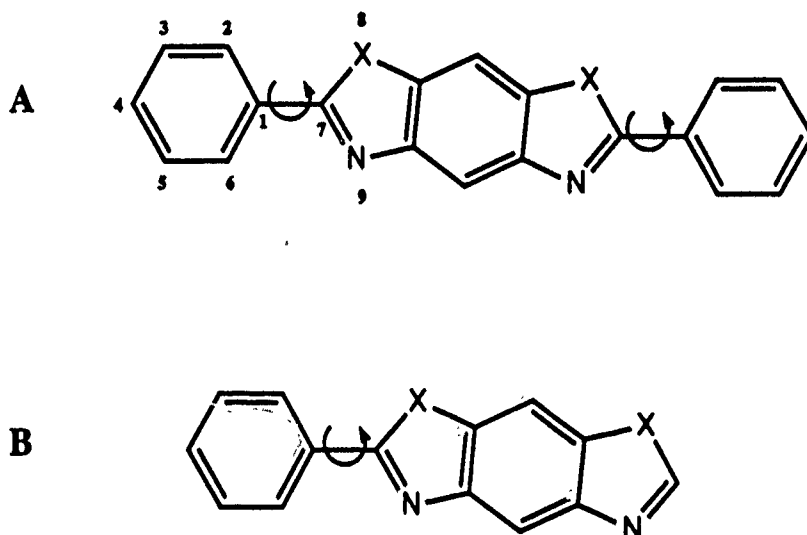
Figure 1. Model Compound Structures used in Previous AM1 Rigid Rod Studies

III

The calculations done here were performed using the AM1 (4) semi empirical method as implemented by the AMPAC 2.01 series of programs, which is available from the Quantum Chemistry Exchange Program,Indiana University, Bloomington, IN 47405 as QCPE #506. The newly introduced sulfur parameters were used in the calculations (5). The model compound used in our calculations is shown in Figure 2. The phenyl group was rotated paddle-wheel fashion about the carbon-carbon bonds indicated in the structure and heats of formation were calculated at every 10° rotation.
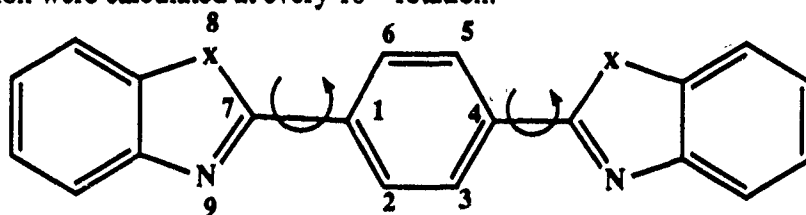
Figure 2. Rigid Rod Model Compound Structure Used in This Study

Table 1 shows a complete listing of structural types which were investigated; specific structures investigated have X=O, X=NH, and X=S. Thus, in addition to the all-planar structures (I, III, V, VII), we have investigated conformations in which the two heterocyclic groups in the structure are mutually perpendicular. Also we have investigated the effect of phenyl substitution on rotational barrier height (III-VIII).

TABLE 1

Structural Types on which AM1 Calculations were done

SYMBOL	STRUCTURE
I	
II	
III	
IV	
V	
VI	
VII	
VIII	

Figure 3 shows a typical result of the calculation described here. The barrier height is the difference between zero and the curve maximum on the Y axis. This figure shows the significant result that when the two ends of the structure are perpendicular the barrier height is less than 1 Kcal/Mol (curve C).
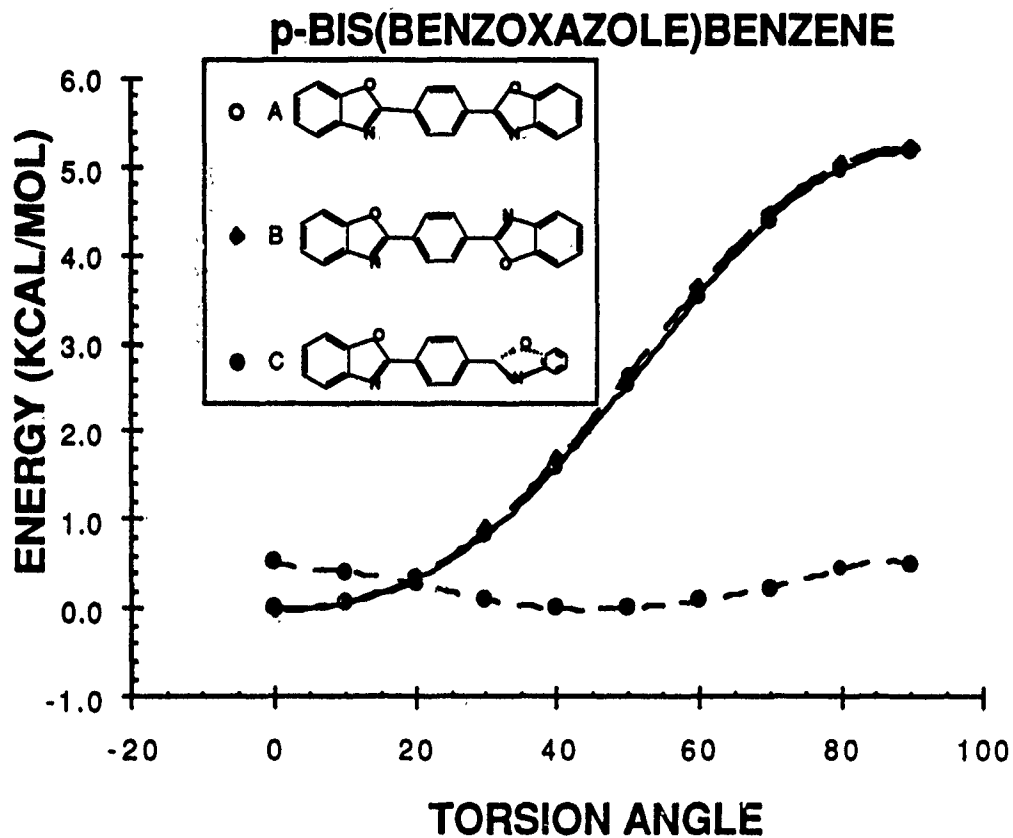
Figure 3. Conformational Energy Curves for PBO Model Compounds.

Table 2 shows a compete summary of all the calculations done during the summer 1990 period. In summary, we find that in all the model systems investigated here the calculated

TABLE 2

STRUCTURAL TYPE	TORSION ANGLE (DEG) MIN ENERGY			TORSION ANGLE (DEG) MAX ENERGY			BARRIER (Kcal/Mol)		
	X=O	X=NH	X=S	X-O	X=NH	X=S	X=O	X=NH	X=S
I	0	30	10	90	90	90	5.2	2.7	2.2
II	45	45	45	0,90	0,90	0,90	0.5	1.3	0.5
III	0	140	30	90	180	90	4	3.1	1.3
IV	20	150	60	90	0	0	1.7	4	1.6
V	40	50	70	90	0	0	1.7	7.1	1.7
VI	20	50	30	90	0	0	4,2	8.5	3.7
VII	80	50	90	90	0	0	2.5	4.7	0.2
VIII	140	130	45	0	180	0	1	3.1	1.

In summary, we find that in all the model systems investigated the calculated barrier height is drastically reduced in the structure where the two heterocyclic groups are mutually perpendicular. Since this represents a physically achievable conformation in the polymer chain, it suggests that the rigid rod polymers modeled in this study may be more flexible than had previously been thought.

The loss of resonance energy as the phenyl group rotates is a major factor in the magnitude of the rotational barrier in all the systems we investigated. Substitution on the central phenyl group causes steric interactions to become a factor as well. In nearly every case we examined methylation ⊂ the central phenyl group *decreases* the rotational barrier since it destabilizes the 0° torsion angle structure. In the case of structure type VI this

destabilization becomes so great that the 0° structure is the highest energy conformation and the rotational barrier is increased.

IV.

Bond Order: Output from AM1 calculations yields information about bond order and atomic charge in addition to the heats of formation described above. The change in bond order between C1 and C7 (Fig. 2), while quite small, is an indication of the change in pi-electron delocalization between the arylene ring and the rest of the molecule. We found that the C1-C7 bond order decreased monotonically by a total of about 3% as the torsion angle went from 0° to 90°. The calculated bond order is evidently not sensitive to steric interaction, but it has the advantage that it isolates the electronic aspect of the barrier height in every case.

Figure 4 shows the excellent correlation between bond order and barrier height for the o-methylated PBO species (structures III and IV in Table 2).
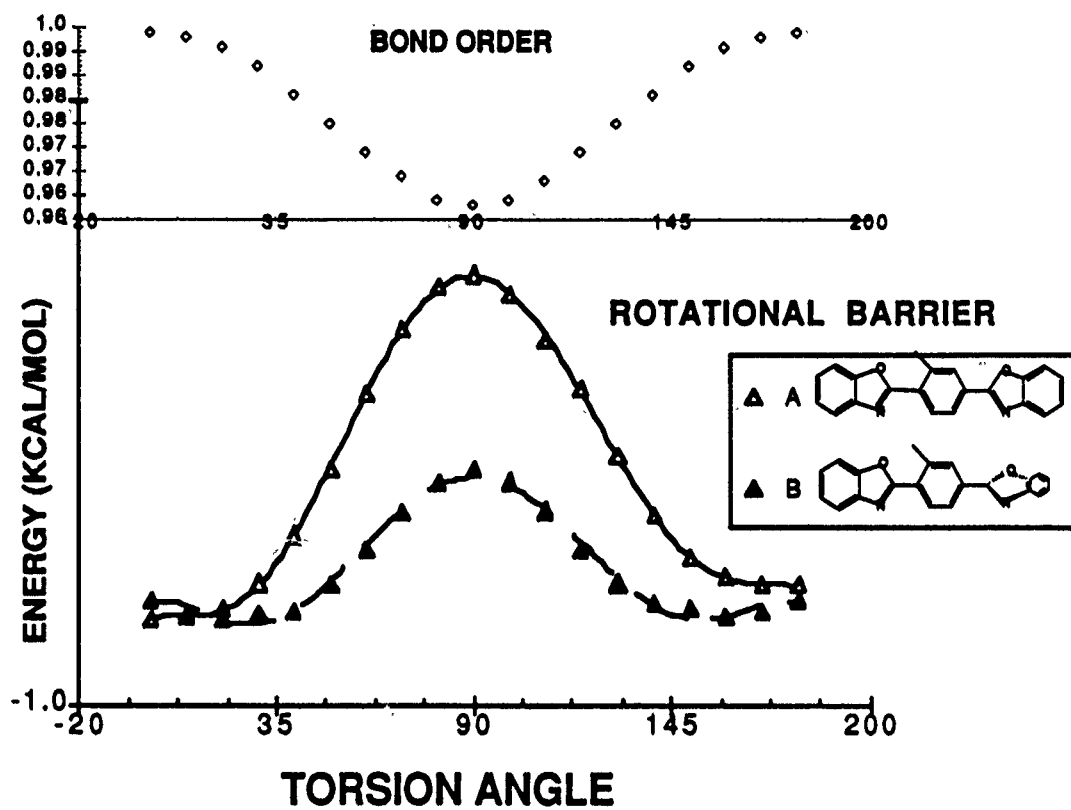
Figure 4. Bond Order-Torsion Energy Correlation for PBO Model Compounds.

On the the other hand, Figure 5 shows that the bond order-rotational barrier correlation is not simple for the corresponding imidazole species. Since the bond order-torsion angle curve is the same for all the species examined here it is clear that non electron-delocalization factors cause the more complex rotational barrier curves seen in most of the species studied here.
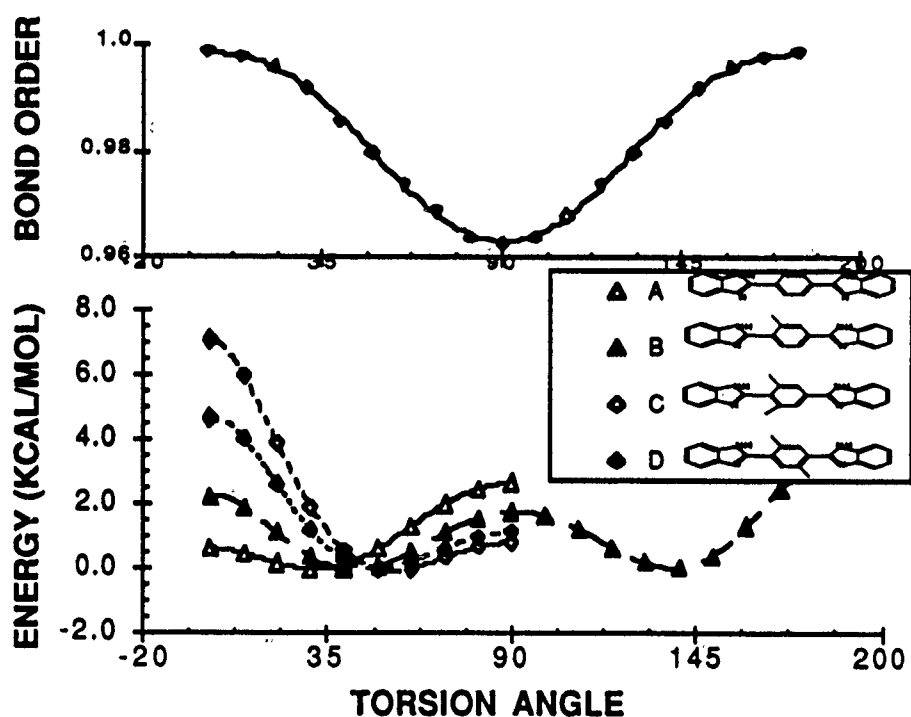
Figure 5. Bond Order-Rotation Barrier Correlation for PBI Model Compounds.

V.

Atomic Charge: The AM1 results obtained here indicate that minor charge redistribution occurs as the torsion angle changes. Charge flows away from the imine nitrogen as coplanarity is lost and charge flows away from the oxazole oxygen as coplanarity is lost in all cases. The charge on the amine nitrogen is unaffected by change in the torsion angle. In the thiazole series charge flows away from the sulfur as coplanarity is lost except for the sulfur adjacent to the two methyl groups in the ortho-dimethyl species, where charge flows toward sulfur as the torsion angle increases. As charge flows away from the heterocyclic atoms, C1 becomes more negative which is consistent the decrease in electron delocalization as coplanarity is lost. Overall atomic charge changes were less than .03 charge units in all cases. Charge changes do not correlate with barrier height in any simple way which is illustrated in Figure 6.
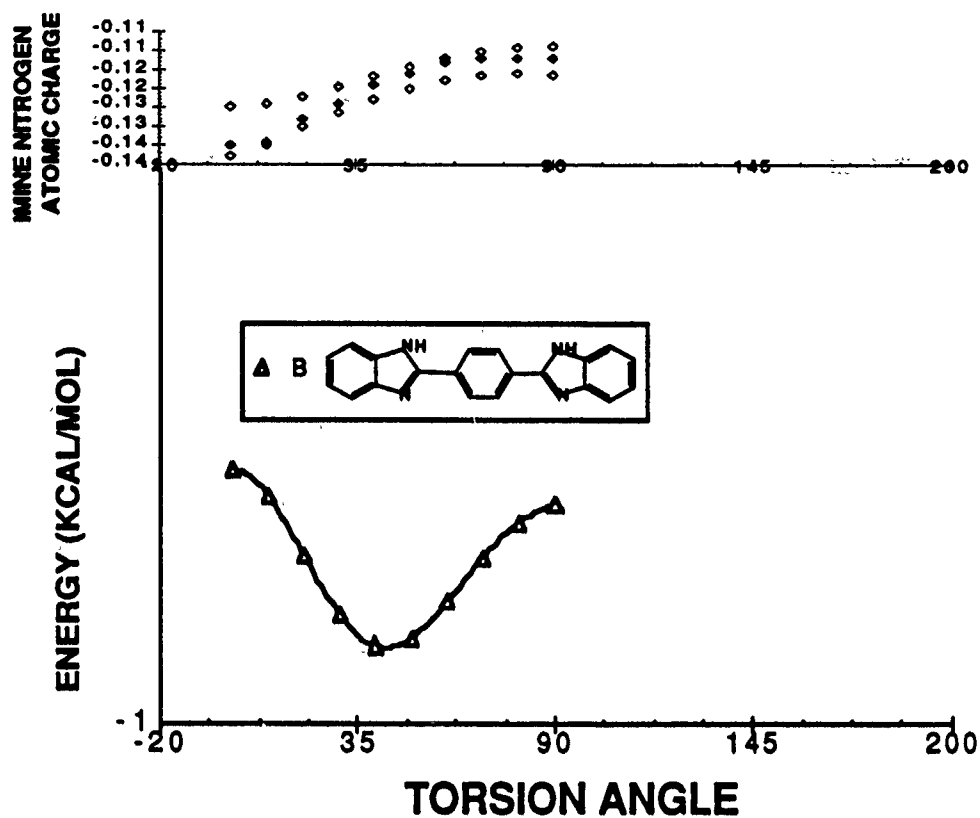
Figure 6. Atomic Charge-Rotational Barrier Correlation for PBI Model Compounds.

VI.

Structural Comparison: It has already been demonstrated that the energy-minimized AM1 structures are very similar to those determined experimentally for PBO and PBT model compounds (6). The major discrepancies are that in the AM1 structures the oxazole oxygen atoms and the thiazole sulfur atoms appear to be closer to sp^3 hybridization while in the experimental structures these atoms are closer to sp^2 hybridization.

The crystal structure of 1,4-bis(2-benzoxazolyl)-2,5-bis(2-benzimidazolyl)benzene, Figure 7, has been determined (7). While the steric interactions between adjacent heterocyclic rings in this compound are significant, the oxazole ring is only 5° from coplanarity with the central phenyl group while the torsion angle of the imidazolyl ring is 57°. This is consistent with our findings as well as those of Farmer (3) regarding PBI model compounds. In Table 5 we compare some experimental and calculated bond distances and angles. We chose the 0° torsion angle PBO model and the 60° torsion angle PBI calculated species for our comparison.
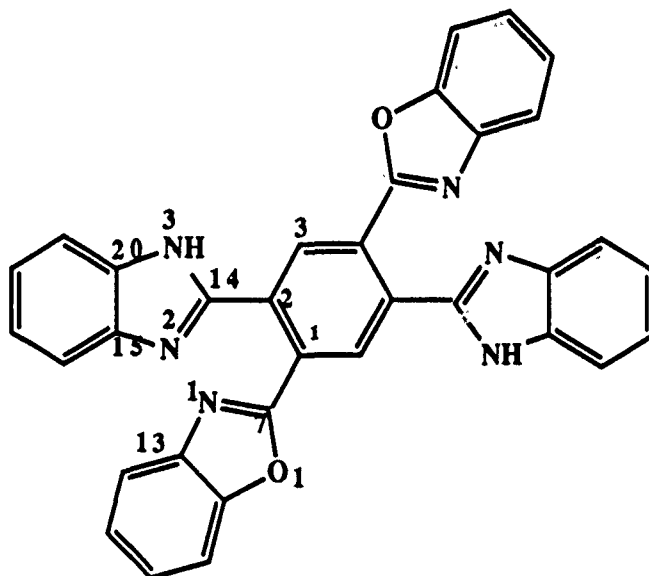
Figure 7. 1,4-Bis(2-benzoxazolyl)-2,5-bis(2-benzimidazolyl)benzene

TABLE 5

Selected Values of Observed and AM1 Calculated Bond Lengths and Bond Angles for 1,4-Bis(2-benzoxazolyl)-2,5-bis(2-benzimidazolyl)benzene and PBO and PBI Model Compounds

Bond Lengths

Bond*	Exp.	Calc.	Bond	Exp.	Calc.
C1-C7	1.466	1.453	C7-N1	1.329	1.332
C7-O1	1.347	1.433	C13-N1	1.392	1.412
C8-O1	1.385	1.396	C2-C14	1.490	1.467
C14-N3	1.393	1.400	C14-N2	1.319	1.351
C1-C2	1.412	1.400			

Bond Angles

Bond*	Exp.	Calc.
C2-C1-C7	125.4	119.9
C1-C7-N1	126.3	113.7
C1-C7-O1	119.5	115.7
C7-N1-C13	105.0	104.2
C1-C2-C14	125.7	119.2
C2-C14-N2	123.8	126.0
C14-N3-C20	104.3	106.5
C14-N2-C15	105.7	105.4

*The numbering system is as in Figure 7; analogous atoms in Figure 2 were used.

VII. RECOMMENDATIONS:

The above calculations have all been done on neutral, electron-paired species. Actual PBT fibers are processed at high acid strength (1) and the resulting dope is subjected to mechanical stress during the formation polymer films. Electron Spin Resonance, ESR, measurements made on unstressed PBT fibers show the presence of free radicals. The ESR signals of the fibers diminished on annealing and increased when the annealed fiber samples were stretched or crushed. A reasonable speculation is that the radicals in the unstressed fibers are ion radicals produced during strong acid workup, and the radicals induced mechanically are neutral radicals resulting from breaking the polymer backbone (8). The presence of an unpaired electron in the PBT polymer may well affect the stiffness appreciably.

It is possible to do AM1 MO calculations on radical species as well as electron-paired species. The complications with odd-electron species are that unrestricted Hartree-Fock (UHF) calculations must be done, which are CPU time intensive, and the energy minimization process may not converge. We recommend then that a thorough, systematic series of MO calculations be done on the cation radical, anion radical, and neutral radical species derived from the even-electron compounds already reported on here. Our primary interest is in the effect of unpaired spin density on the barrier to rotation about the backbone carbon-carbon bond as well as bond order changes within the polymer backbone. Besides having implications for mechanical properties such "defects" are also relevant to postulated electrical conduction mechanisms in these and similar organic materials.

REFERENCES

1) Evers R. C, Arnold, F. E., Helminiak, T. E., "Articulated All-Para Polymers with 2,6-Benzobisoxazole, 2,6-Benzobisthiazole, and 2, 6 Benzobisimidazole Units in the Backbone J. Am. Chem. Soc. (1981), 14, 925

2) Welsh W. J. and Yang, Yong AM1 MO Calculations of the Conformational Characteristics of the Rodlike Polymers PBO and PBT **Macromoleules,** (1990), **23,** 2410

3) Farmer, B. L., Wiershkke, S. G., and Adams, W. W., Study of the Conformations of Stiff Chain Rigid Rod and Substituted Rigid-rod Polymers, **Ibid.** in press

4) Dewar, M. J. S., et. al., AM1: A New General Purpose Quantum Mechanical Molecular Model, **J. Am.Chem. Soc.,** (1985), **107,** 3902

5) Dewar, M. J. S. and Yang, Y-C, **Inorganic Chemistry,** in press

6) Welsh, W. J. and Mark, J. E., "Applications of Quantum Mechanical Techniques to Rigid Rod Polymers, in "**Molecular Level Calculations of the Structure and Properties of Non Crystalline Polymers**", Bicerano, J., Ed. Wiley in press

7) Fratini, A. Private Communication

8) VanderHart, D. L., Wang, F. W., Eby, R. W., Fanconi, B. M. and DeVries, K. L., "**Exploration of Advanced Characterization Techniques for Molecular Composites**", AFWAL-TR-85-4137, February 1986

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Potentials of Mushy-State Forming of Composite Materials

Prepared by:	Sherif D. El Wakil, Ph.D.
Academic Rank:	Professor
Department and	Mechanical Engineering
University:	Southeastern Massachusetts University
Research Location:	WRDC/MLLM
	Wright-Patterson AFB
	Dayton OH 45433
USAF Researcher:	Mr James C. Malas
Date:	28 August 90
Contract No:	F49620-88-C-0053

Potentials of Mushy-State Forming of Composite Materials

by

Sherif D. El Wakil

ABSTRACT

Experimental work was carried out to investigate the problems involved in the mushy-state forming of dispersion-strengthened composites, as well as to assess the possible potentials of such processes. Billets, all having the same Al-Cu-Mn matrix but different alumina contents, were obtained by hot compaction of canned powder mixtures in a blind extrusion die. They were then homogenized for two hours before being extruded at the required temperature. Three temperatures were chosen, to yield different liquid fractions for the extrusion billets, namely 0.4, 0.2 and zero. Billets that had 0.2 liquid fraction were successfully extruded, resulting in sound, defect-free products. Also, metallographic examination of those mushy-state extruded bars revealed an excellent degree of homogenity. In addition, the density and chemical composition were found to be uniform along the length for those bars, indicating the absence of any sensible segregation.

Acknowledgements

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. Acknowledgements must also go to the Universal Energy Systems for their concern and great help to me in all aspects of this program.

My experience was extremely rewarding and enriching not only because of the outstanding facilities of the Materials Lab at WPAFB, but also because of the interaction with the highly qualified scientists in it. The encouragement of James Malas and the fruitful discussions I had with him clearly added to every aspect of this research. The help of Dr Venkat Seetharaman, Dr Young Kim and Carl Lombard was invaluable in overcoming many technical difficulties. Thanks are also due to the staff of the foundry lab, metallograpy lab, SEM lab, Chemical analysis Department, and Mark Dodds. Last, but not least, I would like to thank Jim Morgan and Bill O'Hara for providing a truly enjoyable working atmosphere.

I. INTRODUCTION:

Mushy-state forming is gaining increasing attention as a potential process for manufacturing complex shaped components using intermetallic and ceramic composite materials.

The Material Processing Group of the Metals and Ceramics Division of the Materials Laboratory at Wright-Patterson Air Force Base is particularly interested in developing this new method. The goal is the production of light-weight composites with superior mechanical properties that meet the demands of the aerospace industry in the nineties.

My research interests have been in the area of mushy-state alloys. My work on modeling of the plastic behavior of mushy-state alloys contributed to my assignment to the Materials Processing Group.

II. OBJECTIVES OF THE RESEARCH EFFORT:

Currently, dispertion-strengthened composites are manufactured by casting or powder metallurgy. Although these processes have some advantages, they still have some limitations and shortcomings, for instance segregation, nonhomoginity, and fragmentation of the hard particles. There is, therefore, a need to develop a new manufacturing method to eliminate the above-mentioned shortcomings.

My assignment as a participant in the 1990 Summer Faculty Research Program (SFRP) was to assess the feasibility of mushy-state forming processes for producing complex shaped components using intermetallic and ceramic composite materials.

Since time was a real constraint, it was decided to investigate primarily the processing parameters and to assess the problems associated with that process. An investigation into the effect of the process parameters on the mechanical properties and their correlation with the microstructure obtained, will be continued at my laboratory with funding from the Mini Grant Program.

III. EXPERIMENTAL WORK:

Experimental work was carried out to investigate the problems involved in the mushy-state extrusion of dispersion-strengthened composites. An attempt to employ the conventional powder metallurgy process to obtain suitable billets was not successful. The compressibility and compactibility of Al-Cu powder mixtures containing different percentages of alumina were extremely poor. Also, the compaction dies suffered from excessive wear. It was, therefore, decided to obtain the billets by hot compaction of canned powder mixtures in a blind die. The canned billets were then homogenized for two hours at the required temperature before extrusion.

In the first extrusion experiment, the billet had a temperature of 1140°F and a liquid fraction of about 0.4. The resulting extruded bar was bulged and torn. This mode of failure was attributed to the expansion of the liquid phase. It was, therefore, decided to carry out extrusion at 1040°F so that the metal matrix of each billet would contain only 20% liquid phase. The speed of the extrusion ram was kept at the minimum. Sound and defect-free products were obtained. Next, the density and chemical composition were determined at various points along the length of each extruded bar. They were found to be uniform, indicating the absence of any sensible segregation.

Finally, in order to have a bases for comparison, some billets were extruded at 900°F i.e. there was no liquid phase in any of those billets. Although the extrusions were seemingly good, a decisive factor that determines the appropriate extrusion temperature, is certainly the final mechanical properties of the extruded bars.

IV. RECOMMENDATIONS:

a. The feasibility of forming dispersion-strengthened composites have been basically proven. Nevertheless, the effect of the process parameters on the mechanical properties as well as the correlation between these properties and the microstructures must still be investigated.

b. The results of the above-mentioned investigation can be used to optimize the process of extruding composites in their mushy state.

c. Mathematical modeling of the behavior of mushy-state alloys is required, if those new processes are to be developed to their full potentials.

d. A yield function and a flow rule must be developed first. The constitutive equation can then be incorporated in a FE code, thus coupling the temperature distribution and the plastic flow in order to rationally design tool profiles and anticipate any defects that may occur in the final products.

e. An effort should be made to include the microstructure charateristics in the constitutive equation.

f. It is also recommended to incorporate an analysis of possible chemical reactions in the proposed study.

REFERENCES

El Wakil, S.D., A Model Study of Metal Forming in the Mushy State. <u>Advanced Technology of Plasticity</u>. 1984, Vol 1, pp. 45-49.

Kiuchi, M., et.al., <u>Annuals of the CIRP</u>. 1987, Vol 36-1, pp. 173.

Kiuchi, M. and S. Sugiyama, Application of Mushy-State Extrusion, <u>Proc. of Conf. on Extrusion</u>. 1990, pp. 1-25.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/
GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

STRUCTURAL ANALYSIS OF POLYMER PRECURSORS WITH
POTENTIAL NONLINEAR OPTICAL PROPERTIES

Prepared by: David A. Grossie, Ph.D.

Academic Rank: Assistant Professor

Department and   Department of Chemistry
    University: Wright State University

Research Location: AFWRDC/MLBP
                                      Wright-Patterson AFB
                                      Dayton, OH 45433

USAF Researcher: Douglas S. Dudis, Ph.D.

Date: 28 Sep 90

Contract No: F49620-88-C-0053

STRUCTURAL ANALYSIS OF POLYMER PRECURSORS WITH

POTENTIAL NONLINEAR OPTICAL PROPERTIES

by

David A. Grossie, Ph.D.

ABSTRACT

Single-crystal x-ray diffraction data was collected on two compounds having potential nonlinear optical (NLO) properties, $C_{23}H_{29}NOS$ and $C_{38}H_{50}N_2O_2S_3$. Both compounds crystallize in triclinic crystal lattices, the first having cell constants of a=10.340(2), b=11.632(1), c=8.894(3) Å, α=97.18(2), β=103.18(2), and γ=88.05(1)°. The second compound has cell constants of a=10.292(2), b=20.231(8), c=9.270(2) Å, α=102.48(1), β=98.67(2), and γ=88.66(1)°. The space group observed in each compound is PI. The structure of compound 1 was solved and refined, yielding a R-factor of 0.061. $C_{23}H_{29}NOS$ is planar with little distortion in the internal bond distances and angles.

The second compound, $C_{38}H_{50}N_2O_2S_3$, has not been completely solved, in spite of application of the most recent and capable direct methods programs.

ACKNOWLEDGEMENTS

I would like acknowledge the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. Additionally, acknowledgement of Universal Energy Systems must be given for their assistance in the administrative aspects of the program.

I wish to thank the personnel of the Polymer Branch of the Materials Laboratory and University of Dayton Research Institute for their assistance during the execution of this research project.

I. INTRODUCTION:

The Polymer Branch of the Materials Laboratory at the Wright Aeronautical Laboratory, Wright-Patterson Air Force Base is interested in the synthesis and characterization of polymeric materials. Basic research is also conducted in the structure of polymeric materials and the correlation of the structure and physical properties. The emphasis of this area is to predict the properties of a polymer prior to its synthesis. In this way, the synthesis problem can have greater direction and produce new and better materials with more efficiency. One of the techniques used in determining the structure of polymers is to examine by single-crystal x-ray diffraction methods compounds that may be used to form the backbone, pendents, or cross-links of the polymer. By knowing the structure of a small, repeating portion of the polymer, the polymer itself may be mathematically modeled, yielding the physical properties.

My research interests are in the structural analysis of small organic and inorganic molecules by single-crystal x-ray diffraction. My experience in the structural determinations of organic molecules with similar features to those of interest and my considerable familiarity with the available diffraction hardware and software contributed to my assignment to the Morphology Section of the Polymer Branch.

II. OBJECTIVES OF THE RESEARCH EFFORT:

A study of model compounds of polymeric materials that have potential nonlinear optical properties will be conducted. This study will involve the structural analysis of several compounds by single-crystal x-ray diffraction techniques, with the intent to amass data which may be used to correlate the observed structure and the magnitude of the nonlinear optical response. The primary structural information that is needed by the currently accepted theories is the centricity of the crystal lattice in which the compound of interest crystallizes and the extent of π-orbital conjugation. To this end, this study will examine the effects of steric bulk on the centricity observed in the crystal structure of an otherwise planar molecule.

III.

a. Crystalline samples of a series of compounds were examined using an optical microscope to determine the size and quality of the individual crystals. Two of the compounds examined showed promise of containing suitable crystals for diffraction analysis, whereas the remainder were of insufficient size to be analyzed. Single crystals of $C_{23}H_{29}NOS$ and $C_{38}H_{50}N_2O_2S_3$ were prepared for analysis by attaching them to a thin glass fiber and placing them at the center of an Enraf-Nonius CAD-4 automated diffractometer.

Preliminary x-ray analysis of each of the selected crystals was made. These results are summarized in Table 1, along with the parameters of the subsequent data collection.

b. Data collected on each of the crystalline samples was examined for the presence of space-group-determining systematic absences using the program LOOK (Chapius, 1984). For each sample, an appropriate space group was determined—P1 for both compounds.

c. Using the program GENTAN (Hall, 1989), one of the many crystallographic routines in the software package XTAL (Hall and Stewart, 1989; Grossie, 1990), the structure of $C_{23}H_{29}NOS$ was determined. The initial structure was refined using the full-matrix least-squares refinement routine SFLSX (Hall, Spadaccini, Olthof-Hazekamp and Dreissig, 1989) contained in XTAL. Metrical details of the refined structure are tabulated in Tables 2-7. The structure of $C_{38}H_{50}N_2O_2S_3$ was not completely solved, in spite of the application of several of the most recent and capable direct methods programs available in the laboratory. Work on the solution of this compound is continuing.

d. Figure 1 shows an ORTEP (Johnson, 1971; Davenport, Hall and Dreissig, 1989) drawing of the refined structure of $C_{23}H_{29}NOS$, and a summary of the interatomic distances and angles is presented in Table 2. The central ring system of the molecule is planar, with a maximum deviation from planarity of 0.044 Å. The degree of planarity of the central benzene ring, 0.009 Å and the benzothiazole ring system, 0.005 Å, are much better. This would indicate that the two ring systems are slightly twisted with respect to each other. This is also seen in the examination of the dihedral angles about the C2-C10

bond. The oxydecyl chain that is ortho to the benzothiazole ring is found to extend slightly below the plane of the central benzene ring, with deviations increasing from 0.014 Å for O11 to 1.370 Å for C25.

e. Figure 2 shows an ORTEP drawing of the unit cell contents of $C_{23}H_{29}NOS$. The packing of the molecules in the unit cell is in the herringbone pattern, a common feature in the structure of planar compounds.

IV.

From the completed analysis of $C_{23}H_{29}NOS$ and partial analysis of $C_{38}H_{50}N_2O_2S_3$, certain conclusions can be made regarding to the nonlinear optical properties of these compounds. Since both crystallize in lattices that are centrosymmetric, the second-order term of the polarizability equation must be zero. The response due to the third-order term for these compounds is less certain. The effect of bulky substituents on the centricity of the crystal lattice observed for a compound is still uncertain. The study of compounds with similar structural features may provide the additional information to begin to understand the influence of steric bulk in the arrangement of molecules into acentric lattices.

V. RECOMMENDATIONS:

There is a need to improve the synthesis of nonlinear optical materials and influence the formation of acentric lattices. At the current time, many compounds show great potential

for high NLO responses, however the tendency is for them to crystallize in centrosymmetric lattices. Synthetic modifications of these materials to favor the formation of noncentric lattices should be possible with the knowledge of the mimimum perturbation to the compound necessary to promote asymmetry. The minimum perturbation is needed so as to limit the effect of the change on the NLO response.

Currently, there are three factors that are assumed to produce the desired NLO response, with one factor being quantitative and a second based on a relative scale. These two factors are the centricity of the crystal lattice and the electron-donating and withdrawing effects of the commonly used functional groups. The third factor, the extent of conjugation within the molecule, is currently unquantified.

Molecular planarity is normally taken as the first clue that a non-fused, π-bonded ring system is conjugated. Since this information, like the centricity, is directly obtainable from the structural analysis of a crystalline compound, the extent of conjugation can be quantified by this process.

As a final step, the same compounds examined by x-ray diffraction need to be tested for a nonlinear optical response and the magnitude of that response. With the above pieces of data obtained and analyzed, the synthesis of nonlinear optical materials can be by rational design. This will allow the physical properties of the material to be optimized without compromising the desired nonlinear optical properties.

REFERENCES

Chapius, G. (1984) "LOOK. A FORTRAN Program for Generating Simulated Precession Photographs from Diffractometer Data," University of Lausanne.

Davenport, G., Hall, and Dreissig, W. (1989) "ORTEP" XTAL 2.6 User's Manual. Eds. S. R. Hall and J. M. Stewart. Universities of Western Australia and Maryland.

Grossie, D. A. (1990) "Desktop Crystallography: XTAL for the Personal Computer," International Union of Crystallography Congress, Bordeaux, France.

Hall, S. R. (1989) "GENTAN" XTAL 2.6 User's Manual. Eds. S. R. Hall and J. M. Stewart. Universities of Western Australia and Maryland.

Hall, S. R., Spadaccini, N., Olthof-Hazekamp, R., and Dreissig, W. (1989) "SFLSX" XTAL 2.6 User's Manual. Eds. S. R. Hall and J. M. Stewart. Universities of Western Australia and Maryland.

Hall, S. R. and Stewart, J. M. (1989) Eds. XTAL 2.6 User's Manual. Universities of Western Australia and Maryland.

Johnson, C.K. (1971) ORTEP II. Report ORNL-3794, revised. Oak Ridge National Laboratory.

Table 1. Experimental Details

	Compound 1	Compound 2
Formula	$C_{23}H_{29}NOS$	$C_{38}H_{50}N_2O_2S_3$
Formula Weight	367.55	663.03
F(000)	396	712
Crystal Dimensions	0.2 x 0.3 x 0.4 mm	
Radiation	MoKα	MoKα
Wavelength	0.71073	0.71073
Temperature	23°C	23°C
Crystal Form	Triclinic	Triclinic
Space Group	PI	PI
Cell Constants	a = 10.340(2) Å	a = 10.292(2) Å
	b = 11.632(1)	b = 20.231(8)
	c = 8.894(3)	c = 9.270(2)
	α = 97.18(2)°	α = 102.48(1)°
	β = 103.18(2)	β = 98.67(2)
	γ = 88.05(1)	γ = 88.66(1)
Volume	1033.4 Å³	1863.0
Z	2	2
Density	1.18 g/cm³	1.18 g/cm³
Absorption Coefficient	1.68 cm⁻¹	2.22 cm⁻¹
Scan Type	ω:2θ	ω:2θ
Scan Width	1.00 + 0.344 tan θ	0.85 + 0.344 tan θ
Maximum 2	54.0 °	60.0
Reflections Measured	4060	11492
Corrections	Numerical absorption (range of correction = 1.036 - 1.048) Reflection averaging (Rint = 0.69% for 481 duplicate reflections)	
Observations	1957	
Parameters	235	
R	0.061	
wR	0.086	
S	1.014	
Maximum Shift/Error	0.016	
Residual Density		
Maximum	0.7 e/Å³	
Minimum	-0.4 e/Å³	

Table 2. Atom Coordinate and Isotropic Thermal Parameters for $C_{23}H_{29}NOS$.

	x/a	y/b	z/c	U
S(1)	0.1231(1)	0.1319(1)	0.2004(2)	0.0756(6)
C(2)	0.1997(4)	0.0016(4)	0.1435(5)	0.069(2)
N(3)	0.3053(4)	0.0152(3)	0.0904(4)	0.073(2)
C(4)	0.3315(4)	0.1303(4)	0.0896(5)	0.072(2)
C(5)	0.4368(5)	0.1729(4)	0.0403(6)	0.086(3)
C(6)	0.4511(6)	0.2902(5)	0.0471(7)	0.096(3)
C(7)	0.3619(6)	0.3664(5)	0.1037(8)	0.100(3)
C(8)	0.2585(5)	0.3272(4)	0.1536(7)	0.092(3)
C(9)	0.2424(4)	0.2079(4)	0.1470(5)	0.074(2)
C(10)	0.1517(4)	-0.1130(4)	0.1531(5)	0.067(2)
C(11)	0.0400(4)	-0.1318(4)	0.2122(5)	0.072(2)
O(11)	-0.0225(3)	-0.0361(3)	0.2638(4)	0.082(2)
C(12)	0.0005(5)	-0.2436(4)	0.2173(6)	0.086(3)
C(13)	0.0679(6)	-0.3363(5)	0.1607(7)	0.095(3)
C(14)	0.1765(6)	-0.3196(5)	0.1030(8)	0.100(3)
C(15)	0.2169(5)	-0.2090(4)	0.0968(6)	0.085(3)
C(16)	-0.1420(4)	-0.0473(4)	0.3176(6)	0.074(2)
C(17)	-0.1938(5)	0.0721(4)	0.3524(6)	0.077(2)
C(18)	-0.3186(4)	0.0726(4)	0.4152(6)	0.077(2)
C(19)	-0.3656(5)	0.1935(4)	0.4605(6)	0.079(2)
C(20)	-0.4899(5)	0.2002(4)	0.5240(6)	0.080(3)
C(21)	-0.5337(5)	0.3232(5)	0.5624(7)	0.093(3)
C(22)	-0.6630(5)	0.3359(6)	0.6121(8)	0.105(3)
C(23)	-0.7110(7)	0.4674(6)	0.6287(11)	0.137(5)
C(24)	-0.8276(8)	0.4906(7)	0.6734(11)	0.136(5)
C(25)	-0.8727(7)	0.6129(6)	0.6758(10)	0.124(4)

Table 3. Hydrogen Coordinate and Isotropic Thermal Parameters for $C_{23}H_{29}NOS$.

	x/a	y/b	z/c	U
H(5)	0.5053(5)	0.1182(4)	0.0059(6)	0.075
H(6)	0.5248(6)	0.3217(5)	0.0067(7)	0.075
H(7)	0.3765(6)	0.4524(5)	0.1083(8)	0.075
H(8)	0.1947(5)	0.3838(4)	0.1942(7)	0.075
H(12)	-0.0799(5)	-0.2577(4)	0.2588(6)	0.075
H(13)	0.0416(6)	-0.4178(5)	0.1683(7)	0.075
H(14)	0.2240(6)	-0.3885(5)	0.0590(8)	0.075
H(15)	0.2961(5)	-0.1988(4)	0.0511(6)	0.075
H(16A)	-0.2118(4)	-0.0967(4)	0.2317(6)	0.075
H(16B)	-0.1206(4)	-0.0877(4)	0.4197(6)	0.075
H(17A)	-0.1193(5)	0.1222(4)	0.4307(6)	0.075
H(17B)	-0.2201(5)	0.1088(4)	0.2475(6)	0.075
H(18A)	-0.3948(4)	0.0265(4)	0.3339(6)	0.075
H(18B)	-0.2937(4)	0.0312(4)	0.5164(6)	0.075
H(19A)	-0.2897(5)	0.2382(4)	0.5439(6)	0.075
H(19B)	-0.3871(5)	0.2353(4)	0.3595(6)	0.075
H(20A)	-0.5659(5)	0.1540(4)	0.4427(6)	0.075
H(20B)	-0.4684(5)	0.1615(4)	0.6275(6)	0.075
H(21A)	-0.4593(5)	0.3663(5)	0.6497(7)	0.075
H(21B)	-0.5460(5)	0.3630(5)	0.4603(7)	0.075
H(22A)	-0.7375(5)	0.2894(6)	0.5289(8)	0.075
H(22B)	-0.6501(5)	0.3022(6)	0.7187(8)	0.075
H(23A)	-0.6395(7)	0.5092(6)	0.7208(11)	0.075
H(23B)	-0.7093(7)	0.5018(6)	0.5255(11)	0.075
H(24A)	-0.8991(8)	0.4439(7)	0.5863(11)	0.075
H(24B)	-0.8268(8)	0.4626(7)	0.7813(11)	0.075
H(25A)	-0.8012(7)	0.6596(6)	0.7628(10)	0.075
H(25B)	-0.8735(7)	0.6409(6)	0.5678(10)	0.075
H(25C)	-0.9656(7)	0.6350(6)	0.7005(10)	0.075

Table 4. Anisotropic Thermal Displacement Parameters for $C_{23}H_{29}NOS$.

	U11	U22	U33	U12	U13	U23
S(1)	0.0718(7)	0.0744(7)	0.0848(7)	0.0099(5)	0.0272(5)	0.0109(5)
C(2)	0.065(2)	0.076(3)	0.063(2)	0.008(2)	0.012(2)	0.009(2)
N(3)	0.072(2)	0.074(2)	0.076(2)	-0.001(2)	0.025(2)	0.004(2)
C(4)	0.073(3)	0.072(3)	0.070(3)	0.001(2)	0.016(2)	0.009(2)
C(5)	0.076(3)	0.084(3)	0.099(4)	-0.002(2)	0.027(3)	0.008(3)
C(6)	0.082(3)	0.091(3)	0.123(4)	-0.004(3)	0.034(3)	0.027(3)
C(7)	0.099(4)	0.080(3)	0.125(5)	-0.007(3)	0.025(3)	0.025(3)
C(8)	0.086(3)	0.076(3)	0.114(4)	0.009(2)	0.023(3)	0.015(3)
C(9)	0.067(3)	0.079(3)	0.077(3)	0.004(2)	0.015(2)	0.014(2)
C(10)	0.066(2)	0.071(2)	0.065(2)	-0.001(2)	0.015(2)	0.006(2)
C(11)	0.070(2)	0.072(3)	0.071(3)	0.001(2)	0.018(2)	0.004(2)
O(11)	0.082(2)	0.076(2)	0.100(2)	0.001(1)	0.044(2)	0.006(2)
C(12)	0.081(3)	0.083(3)	0.099(3)	-0.004(2)	0.032(3)	0.009(2)
C(13)	0.102(4)	0.075(3)	0.113(4)	-0.004(3)	0.037(3)	0.010(3)
C(14)	0.101(4)	0.074(3)	0.130(5)	0.005(3)	0.049(3)	-0.004(3)
C(15)	0.087(3)	0.080(3)	0.094(3)	0.005(2)	0.036(3)	0.004(2)
C(16)	0.060(2)	0.081(3)	0.082(3)	-0.004(2)	0.021(2)	0.008(2)
C(17)	0.070(3)	0.089(3)	0.078(3)	0.002(2)	0.027(2)	0.012(2)
C(18)	0.064(2)	0.081(3)	0.087(3)	-0.003(2)	0.019(2)	0.010(2)
C(19)	0.067(3)	0.091(3)	0.081(3)	0.003(2)	0.022(2)	0.013(2)
C(20)	0.067(3)	0.087(3)	0.087(3)	0.001(2)	0.022(2)	0.007(2)
C(21)	0.076(3)	0.098(4)	0.107(4)	0.002(3)	0.027(3)	0.003(3)
C(22)	0.077(3)	0.117(4)	0.117(4)	0.005(3)	0.026(3)	-0.015(3)
C(23)	0.110(5)	0.131(5)	0.170(7)	-0.032(4)	0.061(5)	-0.049(5)
C(24)	0.115(5)	0.128(6)	0.169(7)	-0.011(4)	0.042(5)	0.008(5)
C(25)	0.116(5)	0.098(4)	0.161(6)	0.008(4)	0.038(4)	0.013(4)

Table 5. Bond Distances (Å) and Angles (°) for $C_{23}H_{29}NOS$.

Atoms	Distance	Atoms	Angles
S(1)-C(2)	1.758(4)	C(2)-S(1)-C(9)	89.4(2)
S(1)-C(9)	1.724(5)	S(1)-C(2)-N(3)	114.3(3)
C(2)-N(3)	1.307(6)	S(1)-C(2)-C(10)	124.0(4)
C(2)-C(10)	1.457(6)	N(3)-C(2)-C(10)	121.7(4)
N(3)-C(4)	1.377(6)	C(2)-N(3)-C(4)	111.9(4)
C(4)-C(5)	1.390(8)	N(3)-C(4)-C(5)	125.7(4)
C(4)-C(9)	1.403(7)	N(3)-C(4)-C(9)	114.7(4)
C(5)-C(6)	1.370(8)	C(5)-C(4)-C(9)	119.6(4)
C(6)-C(7)	1.391(9)	C(4)-C(5)-C(6)	119.4(5)
C(7)-C(8)	1.360(9)	C(5)-C(6)-C(7)	120.6(6)
C(8)-C(9)	1.395(7)	C(6)-C(7)-C(8)	121.3(5)
C(10)-C(11)	1.408(7)	C(7)-C(8)-C(9)	118.8(5)
C(10)-C(15)	1.384(7)	S(1)-C(9)-C(4)	109.7(4)
C(11)-O(11)	1.352(5)	S(1)-C(9)-C(8)	130.0(4)
C(11)-C(12)	1.385(7)	C(4)-C(9)-C(8)	120.3(5)
O(11)-C(16)	1.437(6)	C(2)-C(10)-C(11)	123.7(4)
C(12)-C(13)	1.369(8)	C(2)-C(10)-C(15)	118.4(4)
C(13)-C(14)	1.365(10)	C(11)-C(10)-C(15)	117.9(4)
C(14)-C(15)	1.378(8)	C(10)-C(11)-O(11)	116.3(4)
C(16)-C(17)	1.493(7)	C(10)-C(11)-C(12)	120.3(4)
C(17)-C(18)	1.519(7)	O(11)-C(11)-C(12)	123.4(5)
C(18)-C(19)	1.510(7)	C(11)-O(11)-C(16)	119.9(4)
C(19)-C(20)	1.513(7)	C(11)-C(12)-C(13)	120.0(5)
C(20)-C(21)	1.508(7)	C(12)-C(13)-C(14)	120.5(5)
C(21)-C(22)	1.499(9)	C(13)-C(14)-C(15)	120.2(5)
C(22)-C(23)	1.591(10)	C(10)-C(15)-C(14)	121.1(5)
C(23)-C(24)	1.362(12)	O(11)-C(16)-C(17)	107.4(4)
C(24)-C(25)	1.481(10)	C(16)-C(17)-C(18)	112.7(4)
		C(17)-C(18)-C(19)	112.6(4)
		C(18)-C(19)-C(20)	115.3(4)
		C(19)-C(20)-C(21)	112.6(4)
		C(20)-C(21)-C(22)	115.3(5)
		C(21)-C(22)-C(23)	111.9(6)
		C(22)-C(23)-C(24)	117.7(7)
		C(23)-C(24)-C(25)	115.5(8)

112-14

Table 6. Dihedral Angles (°) for $C_{23}H_{29}NOS$.

Atoms	Angle	Atoms	Angle
C(9)-S(1)-C(2)-N(3)	-0.3(3)	C(2)-C(10)-C(11)-C(12)	-179.9(4)
C(9)-S(1)-C(2)-C(10)	179.6(4)	C(15)-C(10)-C(11)-O(11)	178.9(4)
C(2)-S(1)-C(9)-C(4)	-0.2(3)	C(15)-C(10)-C(11)-C(12)	-2.0(6)
C(2)-S(1)-C(9)-C(8)	180.0(5)	C(2)-C(10)-C(15)-C(14)	-179.8(4)
S(1)-C(2)-N(3)-C(4)	0.7(4)	C(11)-C(10)-C(15)-C(14)	2.1(7)
C(10)-C(2)-N(3)-C(4)	-179.2(4)	C(10)-C(11)-O(11)-C(16)	-176.4(4)
S(1)-C(2)-C(10)-C(11)	1.4(6)	C(12)-C(11)-O(11)-C(16)	4.6(6)
S(1)-C(2)-C(10)-C(15)	-176.6(3)	C(10)-C(11)-C(12)-C(13)	1.9(7)
N(3)-C(2)-C(10)-C(11)	-178.8(4)	O(11)-C(11)-C(12)-C(13)	-179.1(4)
N(3)-C(2)-C(10)-C(15)	3.3(6)	C(11)-O(11)-C(16)-C(17)	174.8(4)
C(2)-N(3)-C(4)-C(5)	-179.8(4)	C(11)-C(12)-C(13)-C(14)	-1.9(8)
C(2)-N(3)-C(4)-C(9)	-0.9(5)	C(12)-C(13)-C(14)-C(15)	2.0(9)
N(3)-C(4)-C(5)-C(6)	179.6(5)	C(13)-C(14)-C(15)-C(10)	-2.2(8)
C(9)-C(4)-C(5)-C(6)	0.7(7)	O(11)-C(16)-C(17)-C(18)	178.3(3)
N(3)-C(4)-C(9)-S(1)	0.6(5)	C(16)-C(17)-C(18)-C(19)	-176.1(4)
N(3)-C(4)-C(9)-C(8)	-179.5(4)	C(17)-C(18)-C(19)-C(20)	-180.0(4)
C(5)-C(4)-C(9)-S(1)	179.7(4)	C(18)-C(19)-C(20)-C(21)	178.3(4)
C(5)-C(4)-C(9)-C(8)	-0.5(7)	C(19)-C(20)-C(21)-C(22)	-174.7(5)
C(4)-C(5)-C(6)-C(7)	-0.4(8)	C(20)-C(21)-C(22)-C(23)	172.5(5)
C(5)-C(6)-C(7)-C(8)	0.0(9)	C(21)-C(22)-C(23)-C(24)	180.0(7)
C(6)-C(7)-C(8)-C(9)	0.2(9)	C(22)-C(23)-C(24)-C(25)	175.6(7)
C(7)-C(8)-C(9)-S(1)	179.9(4)		
C(7)-C(8)-C(9)-C(4)	0.0(7)		
C(2)-C(10)-C(11)-O(11)	1.0(6)		

Table 7. Least Squares Planes for $C_{23}H_{29}NOS$.

Plane 1 0.5960 X + -0.0066 Y + 0.8030 Z = 1.8501

Atom	Deviation	Atom	Deviation	Atom	Deviation	Atom	Deviation
C10	0.006(6)	C11	-0.007(6)	C12	0.007(7)	C13	-0.005(8)
C14	0.079(8)	C15	0.009(7)				

----- Other Atoms -----

Atom	Deviation	Atom	Deviation	Atom	Deviation	Atom	Deviation
O11	-0.014(7)	C16	-0.106(10)	C17	-0.219(12)	C18	-0.284(15)
C19	-0.304(18)	C20	-0.369(21)	C21	-0.427(24)	C22	-0.619(27)
C23	-0.861(30)	C24	-1.050(33)	C25	-1.370(35)		

Plane 2 0.5730 X + 0.0248 Y + 0.8192 Z = 1.8128

Atom	Deviation	Atom	Deviation	Atom	Deviation	Atom	Deviation
S1	0.000(2)	C2	-0.002(5)	N3	0.004(5)	C4	-0.005(5)
C5	0.001(6)	C6	-0.001(7)	C7	0.000(8)	C8	0.003(7)
C9	0.001(5)						

Plane 3 0.5730 X + 0.0248 Y + 0.8192 Z = 1.8128

Atom	Deviation	Atom	Deviation	Atom	Deviation	Atom	Deviation
S1	-0.003(2)	C2	0.014(5)	N3	0.025(4)	C4	0.004(5)
C5	0.012(6)	C6	-0.002(7)	C7	-0.015(7)	C8	-0.015(7)
C9	-0.004(5)	C10	0.014(5)	C11	0.028(5)	C12	0.025(6)
C13	-0.029(7)	C14	-0.043(7)	C15	-0.044(6)		

Angles Between Planes

Plane	Plane	Angle	Plane	Plane	Angle
1	2	2.4(2)	1	3	1.8(1)
2	3	0.6(1)			
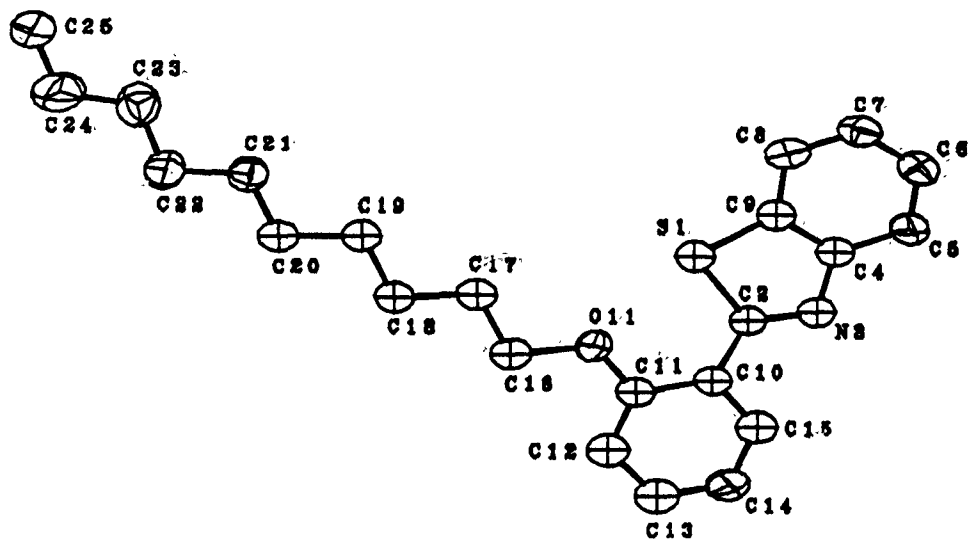
Figure 1. ORTEP drawing of the refined structure of $C_{23}H_{29}NOS$. Hydrogen atoms have been omitted for clarity.
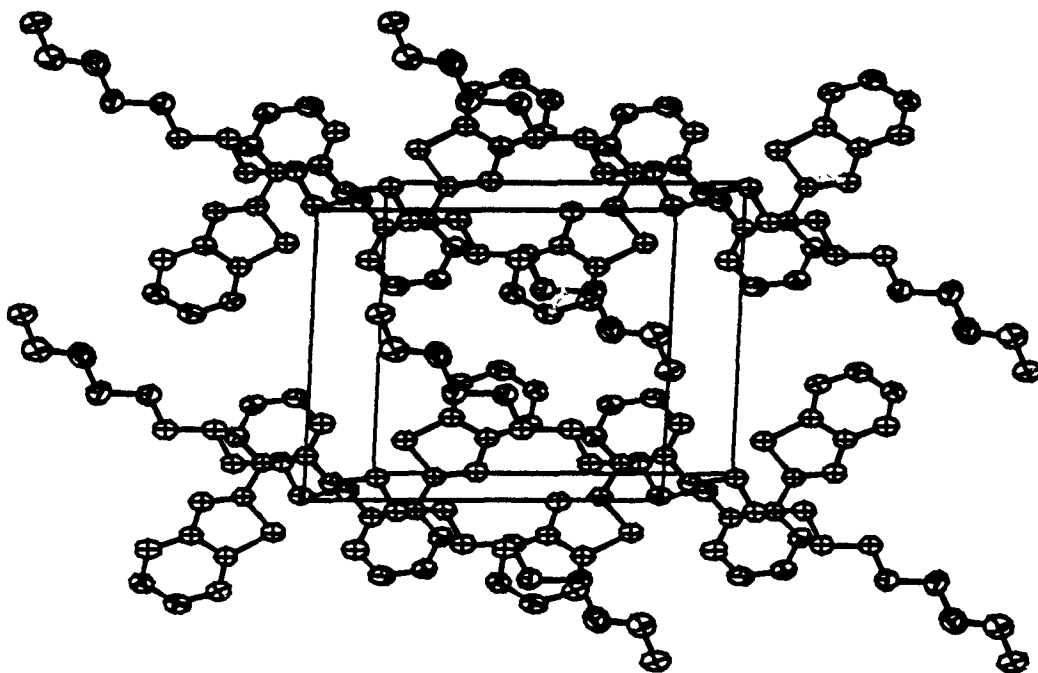
Figure 2. ORTEP drawing of the unit cell contents for $C_{23}H_{29}NCS$.

1717s

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/GRADUATE

STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the Universal Energy Systems, Inc.

FINAL REPORT

Eddy Current Testing in Nondestructive Evaluation

Prepared by:	Thomas J. Haas, B.S.E.E., B.S. Physics
	P.K. Kadaba, PhD.
Academic Rank:	Graduate Student (M.S.E.E.)
	Professor, University of Kentucky
Department and	Electrical Engineering
University:	University of Kentucky
Research Location:	WRDC/MLLP
	WPAFB
	Dayton, OH 45433
USAF Researcher:	P.K. Bhagat
Date:	31 Aug 90
Contract No:	F49620-88-C-0053

Eddy Current Testing in Nondestructive Evaluation

by

T.J. Haas and P.K. Kadaba

ABSTRACT

Advantages and limitations of the eddy current technique for the purpose of nondestructive testing have been evaluated. A cursary study of some of the analytical models and actual test systems that have been developed by researchers over the years has been made.

Using the commercially available eddy current testers - the Nortec NDT-16 and the Hocking AV100SE - tests were made on samples of rubidium and stainless steel with standard machined notches of depths 0.2mm, 0.4mm, and 1.0mm. Also tested was an unknown sample with a barely visible crack. The HP4192A Impedance Analyzer was adapted to eddy current testing by incorporating a power amplifier and specially designed transmit and receive probes. A minute hole in a sample of aluminum was easily detected with this set-up. A pulse technique capable of detecting defects in nonmagnetic metals to a depth of 0.5cm or better was developed. This technique seems to have potential to detect second layer cracks.

ACKNOWLEDGEMENTS

We wish to thank the Air Force Systems Command as well as the Air Force Office of Scientific Research for the opportunity to work with quality people in a quality environment.

Specifically, we would like to make special mention of the efforts of Dr. P.K. Bhagat, Dr. Thomas J. Moran, and Mr. Mark Blodgett, all of whom guided us with their research expertise. The talents of Mr. Mark Ruddell and Mr. Ed Klosterman need to be recognized, as well as Mr. J.A. Fox and Ms. L.L. Mann. The efforts of Mrs. Nancy Lammers were also appreciated.

Of course Universal Energy Systems needs mentioning for organizing and administrating this opportunity.

I. INTRODUCTION:

Eddy current testing is used in the field of nondestructive evaluation to perform the inservice inspection of metal (conductor) products. The experienced operator can, in general, determine the presence of surface flaws (subsurface flaws in nonmagnetic conductors), layer thickness, and perform metal sorting (conductivity).

The method is used to perform these types of inspections because they can be carried out quickly with no contamination of the surface. Despite the simplicity of the technique, it should be noted that the inversion problem, that is, obtaining detailed characterization of surface and subsurface defects of a test object from the probe output signal is a complex problem. Many factors influence the output signal: frequency of excitation of the probe coil, electrical conductivity and magnetic permeability of the material, geometry of the test object and search coil, as well as discontinuities and inclusions in the material.

II. OBJECTIVES OF THE RESEARCH EFFORT:

In the brief period of the program, we have restricted our effort to the evaluation of the following procedures. These are (i) use of Nortec NDT-16 and Hocking AV100SE eddy current testers, (ii) use of the HP4192A Impedance Analyzer, (iii) a Hall Effect pulsed eddy current tester - Nortec-30 Eddyscan, and (iv) a modification of a pulse technique reported in the literature. The above decision was motivated by the availability of the in house commercial equipment in the first two cases. Evaluation of the Hall Effect device was undertaken as it has potential to detect subsurface cracks to about 40 mils in metallic conductors.

The pulse technique researched by us seems promising and should be applicable to investigation of in-depth defects and also for investigation of magnetic materials.

The report is organized as follows: after a brief account of basic principles and analytical models reported in the literature, the four experimental techniques proposed have been described in some detail. This is followed by the lab results obtained by us and the report concludes with a brief discussion of the results.

III. THEORETICAL CONSIDERATIONS

a.) An extensive look at the literature provided insight into the basic principles of the eddy current technique. Textbooks exist on the technique specifically (as well as on nondestructive evaluation (NDE) in general), oftentimes dedicating chapters to the subject at hand. A number of articles in the various journals of the NDE field were sought out as well. These are listed in the bibliography at the end of the report.

b.) The eddy current test employs some of the most basic principles of electricity and magnetism that have been known to us for some 150 plus years. When an object to be tested is placed in a varying magnetic field, varying currents are induced in the test object. The original magnetic field (or primary field) is produced by applying a varying current to a coil, the test probe of the eddy current instrument. Those induced currents in turn produce a field of their own (secondary field) which opposes the primary field. The "size" of this opposition depends on the material characteristics of the tested object, conductivity,

frequency of the applied current, and physical dimensions of the test object and probe coil.

The reduction in magnetic flux due to the opposing fields causes a change in coil impedance, which is monitored. When the eddy current flow is disturbed, as, for example, by a defect, a resultant change in the magnetic field occurs, which in turn affects the coil impedance as described above.

From another viewpoint, when a probe is in air, the coil has a fixed inductance and resistance. When the probe approaches a conducting surface, the apparent resistance increases (due to losses in the conductor) and the inductance decreases due to the oppo,ing fields. Inductive reactance and resistance combine to yield impedance, a vector quantity (or phasor). Reactance is usually assigned as the ordinate (vertical axis) and resistance as the abscissa
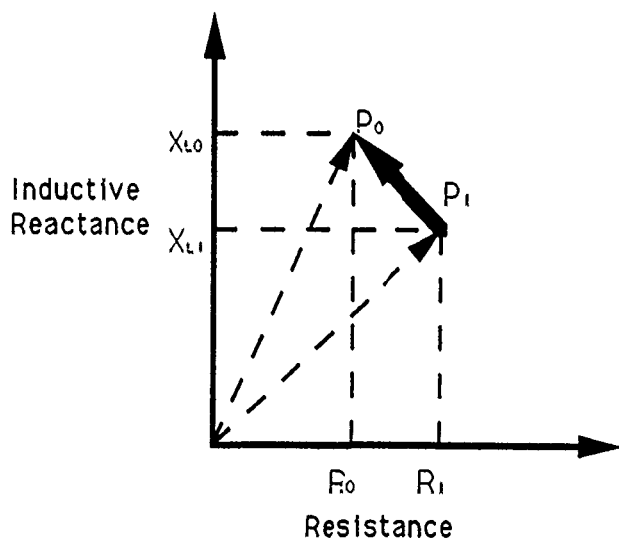
Fig. 1. Impedance plane diagram.   Fig. 2. AC bridge circuit.
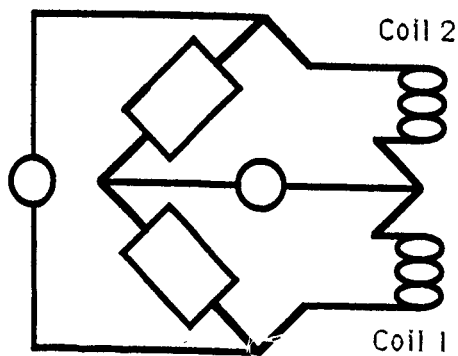
(horizontal axis) in an impedance plane diagram. The impedance plane diagram is normally used as the display in presenting the eddy current test information.

In a passive probe (coil), the probe impedance, whose dynamic characteristics were described earlier, changes due to the presence of a test object (and any alterations therein). These changes are detected in a bridge circuit, translated into a voltage, and then displayed on a CRT screen. This display is the representation of the (complex) impedance plane.

Please note that Figure 1 is consistent with the coil impedance having a larger resistive component and smaller reactive component as the probe approaches the test object.

Impedance can be written in its complex form as $Z = R + jX$, where R represents the resistive component and X is the inductive reactance component. A dependence of impedance on frequency can be uncovered if we expand X as $X = \omega L = 2\pi f L$, where L is coil inductance and f is the frequency of operation of the excitation current.

Frequency of operation is one of the most important parameters in eddy current testing, and is one of the few that the inspector has control over. The more choices the inspector has in picking a frequency for the test (i.e. wider instrument bandwidth), the more powerful the instrument, in general.

In the development of analytical models for eddy current characterization of defects (voids) the integral equation approach has been widely used. A major drawback in using the integral equation approach as the heart of a defect characterization scheme lies with the basic dipole assumption; how can such a model be used as the basis of the inverse problem solution when one has

already assumed a priori that the defect has a spheroidal shape? But this is what is done. Work should be continued in this area for arbitrarily shaped defects.

As an example of the approach taken in formulating a finite element analysis, consider a two dimensional region with z directed current density vector $\mathbf{J}$. For this geometry, the magnetic vector potential $\mathbf{A}$ is directed parallel to $\mathbf{J}$, and is described by:

$$\frac{\partial}{\partial x}\left[\frac{\partial \mathbf{A}}{\partial x}\right] + \frac{\partial}{\partial y}\left[v\frac{\partial \mathbf{A}}{\partial y}\right] = -\mathbf{J} + j\omega\sigma\mathbf{A} \tag{1}$$

where sinusoidal excitation of angular frequency ω is assumed.

The principles of variational calculus lend a solution to (1) by minimizing

$$F = \iint\limits_{R}\left[\tfrac{1}{2}v\{\left[\frac{\partial \mathbf{A}}{\partial x}\right]^2 + \left[\frac{\partial \mathbf{A}}{\partial y}\right]^2\} + \frac{j\omega\sigma}{2}\mathbf{A}^2 - \mathbf{J}\cdot\mathbf{A}\right]dx\,dy \tag{2}$$

where R is the bounded region of interest. F is called the energy functional.

Analytical models for layered media revolve around finding a solution for the magnetic vector potential, $\mathbf{A}$, as well. For a circular coil electrically close to a plane material medium (dielectric constant ε_1, conductivity σ_1, and permeability μ_1) at z=0 and coil axis in the z direction only the A_ϕ (cylindrical coordinates ρ, ϕ, z) component has any significance [Ref. 7].

By solving the wave equation and implementing the boundary conditions we find, for z>0

$$A_\phi = \frac{\mu_0 Ia}{2}\int_0^\infty J_1(\varepsilon p)J_1(\varepsilon a)e^{-\lambda_0(z-d)}\frac{\varepsilon d\varepsilon}{\lambda_0} \tag{3}$$

$$+ \frac{\mu_0 Ia}{2}\int_0^\infty J_1(\varepsilon p)J_1(\varepsilon a)e^{-\lambda_0(z+h)}\{\frac{\mu_1\lambda_0 - \mu_0\lambda_1}{\mu_1\lambda_0 + \mu_0\lambda_1}\}\frac{\varepsilon d\varepsilon}{\lambda_0}$$

where $J_1(\varepsilon p)$ refers to the first order Bessel function of the first kind, d is the

coil/material separation, a is the coil radius, and λ_i refers to the wavelength in region i (i = 1 or 0 here). The term is brackets is analagous to the reflection coefficient expression used in describing electromagnetic wave interaction with materials.

If the medium above is replaced by a layered media of thickness, t, ε_1, μ_1, σ_1 and ε_2, μ_2, σ_2 describing its material properties, we obtain A_ϕ as in (3) upon replacing the term is brackets with

$$\Gamma = \frac{\Gamma_{01} + \Gamma_{12}\, e^{-2\lambda_1 t}}{1 + \Gamma_{01}\Gamma_{12}\, e^{-2\lambda_1 t}}$$

where

$$\Gamma_{01} = \frac{\mu_1\lambda_0 - \mu_0\lambda_1}{\mu_1\lambda_0 + \mu_0\lambda_1}; \quad \Gamma_{12} = \frac{\mu_2\lambda_1 - \mu_1\lambda_2}{\mu_2\lambda_1 + \mu_1\lambda_2}$$

Upon making this substitution one obtains the solution of the magnetic potential.

For defect characterization work, a model is required which allows for a variety of defect shapes, test geometries and excitation condition. Both finite-difference and finite element techniques appear to have this flexibility. Even with these methods, published work in the literature, by and large, is restricted to characterization of defects which have well defined symmetries in nonmagnetic materials. Additional work must be done to extend the techniques to to three dimensions as well as defects of arbitrary shapes and to take account of nonlinear material properties so that defects which include hysteresis effects in ferromagnetic materials can be studied. As with any computer-based modeling, finite element analysis techniques give results whose accuracy is very much a function of the quality of the input data.

IV. ACTUAL TEST SYSTEMS

a.) Experiments were performed using the Nortec NDT-16 and the Hocking AV100SE eddy current testers, and the HP4192A Impedance Analyzer. Samples of Rhobidium and stainless steel with machined notches of depths 0.2mm, 0.4mm, and 1.0mm were used, as well as an unknown sample with a barely visible crack. These were analyzed to varying degrees by the aforementioned devices. There were some limitations to the test procedure that we could not undertake due to lack of calibration standards for use in eddy current testing in the MLLP section. See section **V. RECOMMENDATIONS**, in this report.

b.) There was some difficulty in learning to use the Nortec NDT-16 eddy current tester, as there was no documentation available on the instrument. By speaking with individuals with some experience in eddy current testing and by using other testers, a knowledge of this particular device was obtained.

By placing the probe on the test object and nulling the bridge, a new reference point (or origin) for the impedance plane diagram is made. Now, as the surface of the test object is scanned, changes in its material properties are monitored, and the signal from these changes is with respect to the reference point. This is vital in that the lift off signal (the signal change between probe in air and probe on the test object) dominates the defect signals. So the probe is placed on the material, nulled, and then the lift off signal is directed horizontally (by

convention), through the use of a phase rotation control. This provides some discrimination against the lift off signal. Now the smaller signals are distinguished from the lift off signal by both amplitude (smaller) and phase. Skin depth is inversely proportional to the square root of frequency and conductivity. The higher the frequency, the smaller the distance the currents penetrate into the material (for a constant conductivity). Currents at greater depths below the material surface flow at greater phase angles that lag behind the currents nearer the surface. Therefore, the signals generated from the lower frequency, deeper penetrating currents lag behind signals generated by the higher frequency currents. This can be observed in Fig. 4 and 5. A stainless steel sample was scanned at constant gain for frequency, f=500kHz (Fig. 4) and f=1MHz (Fig. 5). Note that the liftoff signal is in the horizontal direction and that despite changes in amplitude for the three crack signals (amplitude proportional to crack depth), an obvious phase difference is observed.

There are some trends present in the figures which remain consistent for any sample (a rubidium sample machined exactly as the stainless steel sample was also tested). They are: signal amplitude increases with frequency and crack depth, and phase difference from liftoff signal increases with frequency and crack depth.

Important limitations exist in using the Nortec NDT-16. The most important of these is the frequency bandwidth, and the fact that only 3 discrete frequencies are available. It would be greatly desirable to have continuous frequency settings, as well as the capability to use frequencies well below the lower limit on this instrument of 500kHz. This is a necessity in having the ability to detect subsurface defects.
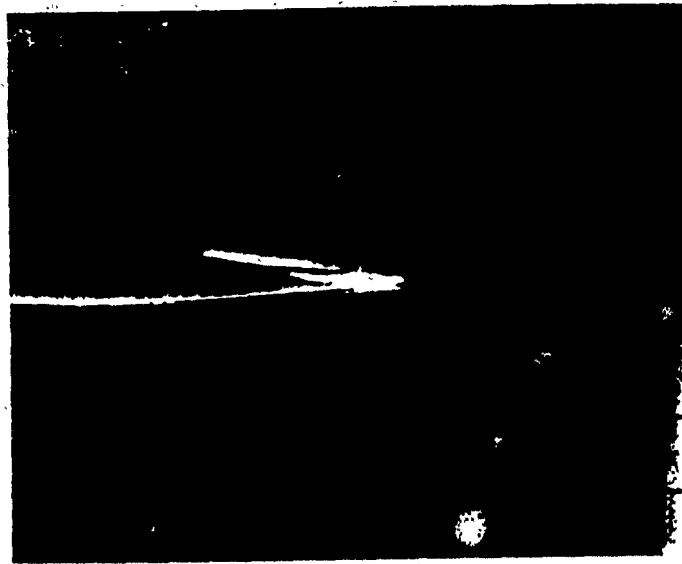
FIG. 4. Eddy current test signal for stainless steel
calibration standard taken from the Nortec NDT-16
at 500kHz.
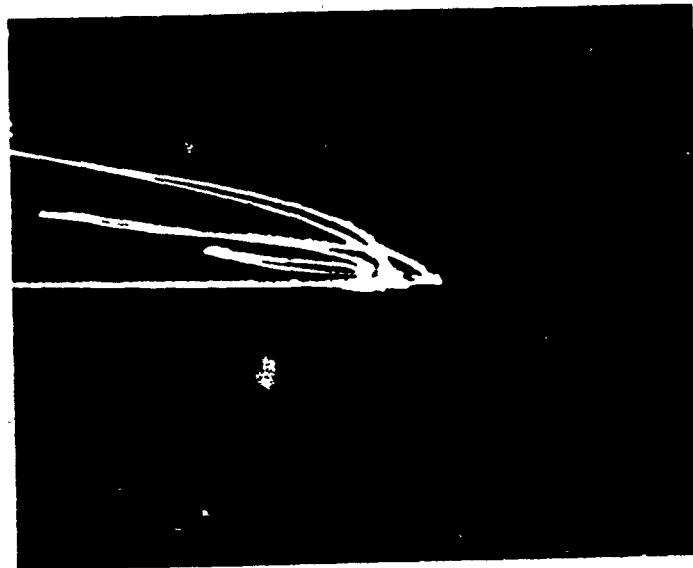
FIG. 4. Eddy current test signal for stainless steel
calibration standard taken from the Nortec NDT-16
at 1MHz.

For the detection of subsurface cracks, low frequency techniques need to be implemented. In attempting to inspect second and third layer materials, greater success will be achieved if the outer layers are, in the case of a change in material with layer, lower in conductivity. In lowering frequency, a necessary increase in probe diameter is required. This results in an increase in the size of detectable cracks.

Due to a lack of proper calibration samples, no scientific data was obtained for subsurface cracks for this report. However, using the Hocking AV100SE instrument, and simulating a crack by placing two thick (electrically) copper sheets adjacent to one another and covering them with a sheet of metal composite, the area between the copper sheets was an observable signal.

The HP Model 4192A LF Impedance Analyzer is a fully automatic, high performance test instrument designed to measure a range of impedance parameters as well as gain, phase, and group delay. Other features of the instrument pertinent to the project are: 1.) frequency of the oscillator output can be varied from 5Hz to 13MHz, 2.) oscillator level is variable from 5mV to 1.1V (rms) with 1mV resolution, 3.) measurement range of gain/loss of the test channel with reference to the reference channel is -100dB to +100dB with 0.001dB maximum resolution and 0.02dB to 0.09dB basic accuracy, and 4.) measurement range of phase is -180 to +180 degrees with 0.01 degrees resolution and 0.1 degrees to 0.2 degrees basic accuracy. The 4192A provides HP-IB interface capability and this feature makes it possible to integrate the 4192A into an automated test system which reduces time and cost of testing.
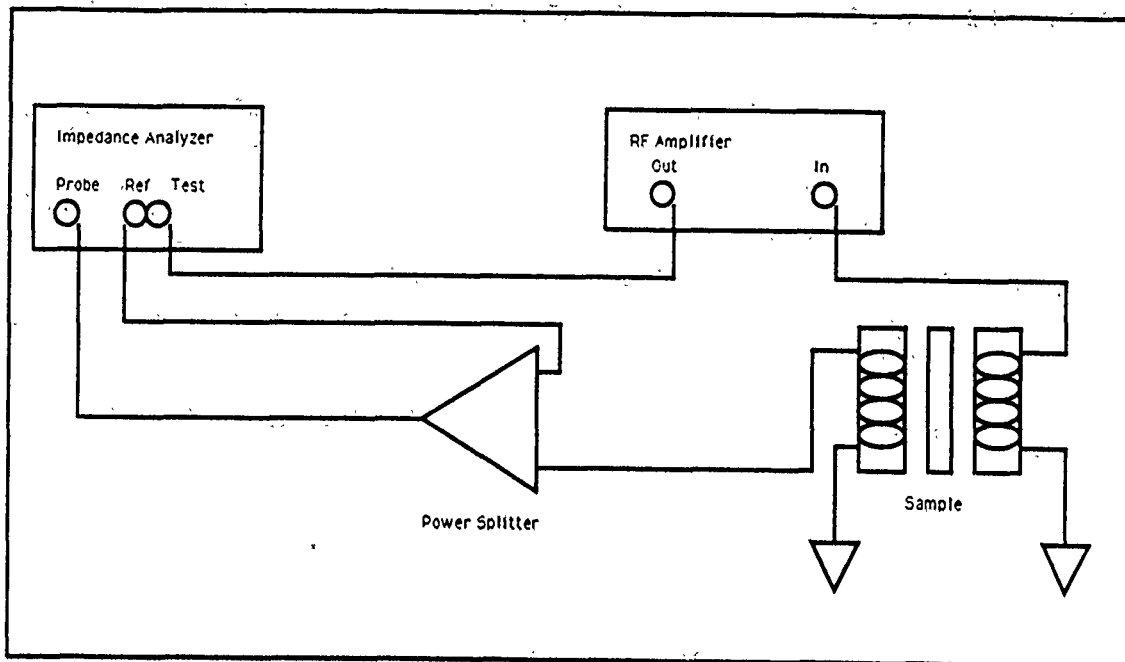
Fig. 3. Eddy current testing using the HP 4192A Impedance Analyzer.

In our tests the HP4192A was modified using an external power amplifier and specially designed transmit and receive probes for the eddy current measurements. A sample of aluminum foil, 25.4 microns thick with a hole of 1.07mm diameter was used in this setup (Fig. 3) with an excitation frequency of 1MHz. A 9.8dB change in the receiver probe signal magnitude was observed as the sample was passed over the probes. It is to be noted that operating the coils at their resonant frequency makes eddy current measurements difficult. Good sensitivity was achieved in measuring the coil impedance by the impedance analyzer. The resonant condition occurred well below 1MHz. This method shows good potential for automation, and we plan to continue the study at the University of Kentucky.

A method similar to that of Waidelich [Ref.7], with some modifications, was investigated. This is a pulse technique and capable of detecting defects in nonmagnetic metals to a depth of 0.5cm or better. A brief description of the experimental setup used follows: the output of the HP8111A pulse generator

(prf: 1kHz). was amplified by a broadband power amplifier and fed to a transmitting probe coil. The transmitting coil consisted of a 600 turn coil of copper wire wound on a straight ferrite core of 3/16 inch diameter cross section. Likewise the receiving probe coil was made of 2000 turns of wire wound on a similar ferrite rod. The output of the receiver probe was amplified and displaced on an oscilloscope. Both the probes were shielded axially to prevent leakage of stray flux. Two of the important parameters are the height of the peak voltage of the displayed pulse and the time delay from the beginning of the pulse to the peak. For aluminum sheets these two quantities were measured for thicknesses of 0.2 to 10 cms. It was found that the peak height varies as $1/t^{3.5}$ where t is the thickness of the specimen and the time delay of the output pulse peak varies as $t^{1.8}$. By moving the probe coils over the surface of the sheet and observing the behavior of the tail of the output pulse, the position and depth of the defects can be ascertained. This technique has potential and it is proposed to continue these measurements at the University of Kentucky. Preliminary tests on holes drilled in one aluminum sheet and sandwiched between two other sheets seem encouraging.

V. RECOMMENDATIONS

There is at present a lack of calibration standards for doing eddy current testing at the MLLP branch. EDM notches or machined slits should be obtained for the materials for which there is interest in inspecting. These calibration standards should duplicate test material in geometry as well as in electrical and magnetic properties. Due to the failure to characterize in detail the flaw signal in eddy current analysis of materials, eddy current testing is not an absolute method.

The Nortec NDT-16 eddy current instrument currently in the MLLP branch has serious limitations in probe excitation frequency as well as probe types and sizes. There are instruments available that will go much lower in frequency. This capability is a necessity, especially when materials classified as good conductors are being evaluated for flaws. It becomes imperative if multi-layered structures are to be analyzed. Probe (coil) size offers another choice to the operator of test equipment, and a variety of coil diameters should be made accessible.

It is our feeling that complete characterization of defects will combine knowledge obtained from analytical and numerical models integrated with a data base of signals obtained from calibration standards. Neural networks could very well play the role of integrator. Calibration standards should include signals from both machined defects as well as actual, typical defects of the material to be inspected.

In conclusion, eddy current testing has proven to be reasonably successful in the past. Further improvements in methodology and test equipment would make the technique more versatile in flaw detection in metallic materials.

REFERENCES

1. RUDLIN, J.R., "A Beginner's Guide to-Eddy Current Testing," <u>British Journal of NDT</u>, University College Underwater NDE Centre, June 1989, pp. 314-320.

2. Hagemair, D.J., "Eddy Current Impedance Plane Analysis," presented to 1983 Air Transport Association Non-Destructive Testing Forum, Atlanta, Georgia, September 1982, pp. 1-5.

3. Dodd, C.V., Whitaker, L.M., and Deeds, W.E., "An Accurate Laboratory Test System Using Commercial Equipment for Eddy Current Measurements," <u>Materials Evaluation</u>, November 1988, pp. 1569-1574.

4. Libby H.L., <u>Introduction to Electromagnetic Nondestructive Test Methods</u>, Richland, Washington, Wiley-Interscience Publishers, 1971.

5. Cecco, V.S., and Van Drunen, G.,"Recognizing the Scope of Eddy Current Testing," in <u>Research Techniques in Nondestructive Testing Vol. VIII</u> Sharpe, R.S., Ed., Academic Press Inc., London, 1985, pp. 270-301.

6. Auld, B.A., Muennemann, F.G., and Riazat, M., "Eddy Current Modelling," in <u>Research Techniques in Nondestructive Testing Vol. VII</u>, Sharpe, R.S., Ed., Academic Press Inc., London, 1984, pp. 37-76.

7. Waidelich, D.L., "Pulsed Eddy Current Testing of Steel Sheets," in <u>Eddy Current Characterization of Materials and Structures</u>, Birnbaum/ Free, Ed., ASTM Special Technical Publication 722, Philadelphia, 1981, pp. 367-373.

8. Palanisamy, R., and Lord, W., "Finite Element Analysis of Eddy Current Phenomena," <u>Materials Evaluation</u>, October 1980, pp. 39-43.

9. Cheng, D., "The Reflected Impedance of a Circular Coil in the Proximity of a Semi-Infinite Medium," PhD Dissertation, University of Missouri, January 1964.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Preparation and Characterization of Polypeptide Thin Films

Prepared by:	Joseph B. Lambert
Academic Rank:	Professor
Department and	Chemistry
University:	Northwestern University
Research Location:	WRDC/MLPJ (Electromagn. Mat. Div.) Wright-Patterson AFB Dayton, OH 45433-6539
USAF Researcher:	Robert L. Crane
Date:	June 15, 1990
Contract:	F49620-88-C-0053

Preparation and Characterization of Polypeptide Thin Films

by

Joseph B. Lambert

ABSTRACT

Polypeptides based on the glutamic acid backbone have nonlinear optical properties. In order to test practical applications of such materials, thin films have been cast by use of spin coating. Films were prepared for poly(benzyl-L-glutamate) (PBLG) and for poly(N-(p-trans-azobenzene)-L-glutamide) (PALG). These films were characterized by polarized microscopy, spectroscopic ellipsometry, and Fourier transform infrared spectroscopy.

Acknowledgments

This work was sponsored by the Air Force Systems Command and the Air Force Office of Scientific Research and was administered by Universal Energy Systems, to whom the author is grateful. This work was carried out in the Biotechnology Laboratory of the MLPJ branch of the Electromagnetic Materials Division of the Materials Laboratory of Wright-Patterson Air Force Base. The author is particularly indebted to Dr. Robert Crane, Dr. Thomas Cooper, Dr. Zbigniew Tokarski, and Mr. Timothy Bunning for guidance and assistance in this work.

I. INTRODUCTION

Organic nonlinear optical materials are of interest to the Air Force for a wide variety of applications, including optical computing, optical storage, wave guiding, and optical sensor/vision protection. The major effort in recent years has been on highly conjugated carbon chains such as polyacetylenes. The discovery that polysilanes have NLO properties (Miller and Michl, 1989) revealed the existence of an entirely new NLO material based on saturated chains of silicon atoms. Electron delocalization in these materials is based on $\sigma \rightarrow \sigma*$ conjugation rather than the traditional $\pi \rightarrow \pi*$ conjugation.

For the past ten years our research at Northwestern has centered around the preparation and investigation of organic molecules containing silicon, germanium, or tin that possess novel electronic properties. We have compared the abilities of these three atoms to delocalize charge through $\sigma \rightarrow \pi$ conjugation (Lambert et al., 1989) and found that silicon and germanium have similar abilities but that tin has a vastly enhanced ability. Consequently, we became interested in the possibility that polystannes could have enhanced NLO properties compared with those of polysilanes.

Polysilanes have many outstanding properties for optical uses, including fast response times, low visible absorptivity and scattering, good thermal stability, and good processibility. Moreover, their $\chi^{(3)}$ are in the range of the better NLO conjugated carbon chains. We expect polystannanes to possess all these positive

properties and in addition have greatly enhanced values of $\chi^{(3)}$ because of the much higher polarizability of tin.

To learn about the preparation and processing of NLO polymers, I felt that experience in the Materials Laboratory of Wright-Patterson Air Force Base would be invaluable, as that laboratory has a long history of investigation of NLO materials. Consequently, I requested this laboratory in order to gain exposure to materials methodology and to work with NLO polymeric materials.

II. OBJECTIVES OF THE RESEARCH EFFORT

The major objectives of this work were to obtain a thorough grounding in the theory of nonlinear optics, to learn how to manipulate long chain NLO polymers, and to understand what the significant problems in NLO chemistry are today. In order to achieve these objectives, a protocol was selected to prepare and characterize a family of high polymers with NLO properties, the derivatives of polyglutamic acid. Levine and Bethea (1976) have measured the second-order and third-order nonlinear optical properties of the α-helix form of poly(benzyl-L-glutamate) (PBLG) with molecular weight about 550 kilodalton. Suitable alteration of the side chain might improve the NLO properties, so Dr. Thomas Cooper of the MLPJ laboratory has been engaged in a synthetic project to prepare such derivatives. Past work has concentrated on studying the NLO properties of these materials in solution. It was our specific objective to determine whether homogeneous thin films, free of solvent, could be cast and characterized. Consequently, we developed methods using the spin coating technique to prepare such thin films of PBLG and of the azo derivative PALG whose structure is shown on the next page. The PBLG films were thoroughly characterized by several techniques and the more fragile PALG films by infrared spectroscopy.

Azo-modified polymer
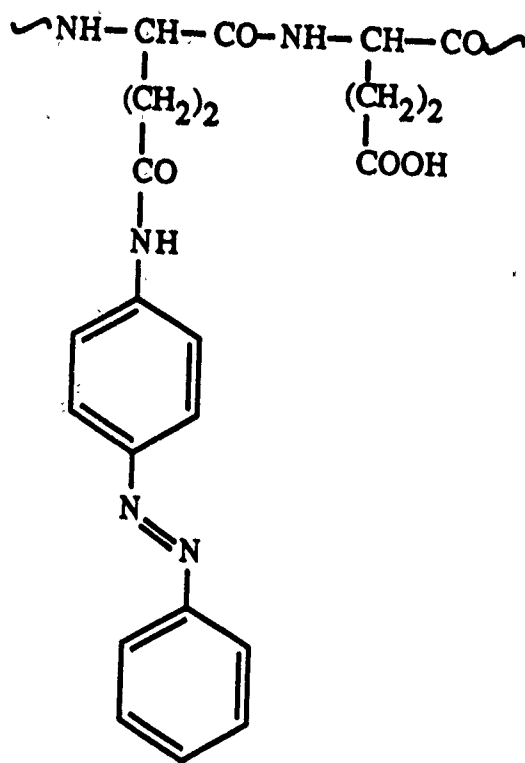
III.

a. Prior to this work there was no operative spin coater in this laboratory. The available equipment was therefore set up and put into working condition. All spin coating was carried out on a Solitec 5110-TC instrument. Figure 1 on the following page shows the calibration that was carried out between settings and the actual spindle speed.

Several solvents were examined for dissolution of PBLG and PALG. For optimal results, the solid polymer must be converted to a somewhat viscous, entirely homogeneous solution. PBLG was dissolved successfully in 1,2-dichloroethane (DCE), dimethyl-formamide (DMF), 1,4-dioxane (DO), and chloroform. Thin films initially were cast on standard glass microscope slides. Both DCE and DO gave excellent, clear, homogeneous thin films of considerable strength. The chloroform solution was too viscous and did not produce a uniform film. DMF spread out evenly but gave a film with a spotty appearance. For comparison, a film also was prepared on a glass slide by mechanical shearing with another glass slide of PBLG dissolved in DCE. Samples also were cast on silicon or germanium wafers. Two cycles were used on the spin coater, a slower spread cycle and a faster spin cycle. The speed of the spread cycle typically was 80-120 krpm for 4-10 s, and that of the spin cycle was 200-500 krpm for 20-30 sec. Actual selection of conditions determined the thickness and quality of the film.

Attempts were made to dissolve PALG in acetone, DCE, and DO, but homogeneous, highly colored solutions were obtained only with DO.

Biotechnology Laboratory Spin Coater

Speed = -0.6519 + 0.01088 * Setting
r: 0.9995971749

Spindle Speed (KRPM)
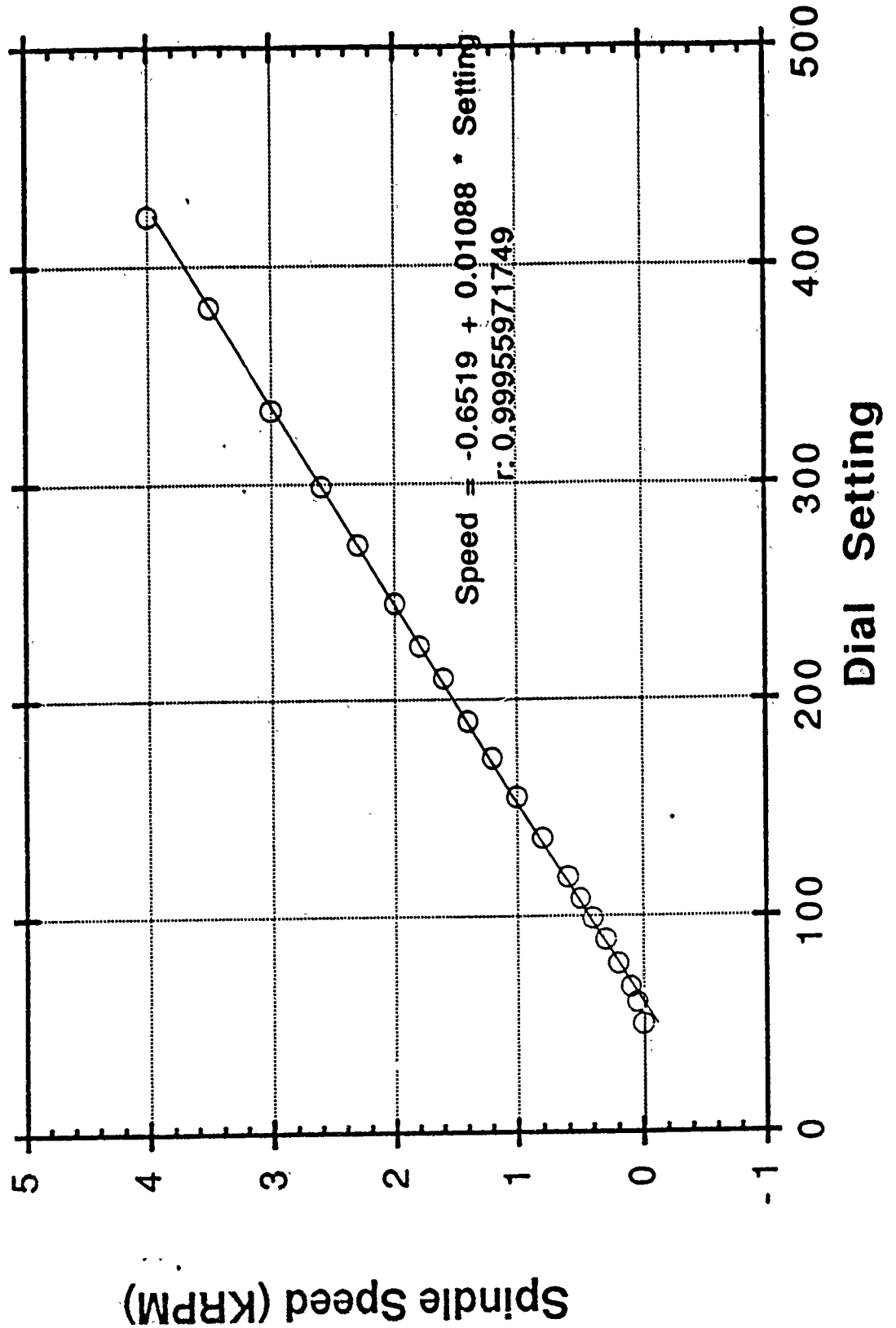
Dial Setting

Figure 1. Spindle speed as a function of dial setting.

114-9

b. Film thickness was measured approximately on a Dektak II surface profile instrument. Figure 2 on the next two pages shows the resulting profiles from films of PBLG cast on glass. The sample obtained by mechanical shearing gave a rather thick and very uneven film, with a thickness varying between 50,000 and 150,000 Å, with a mean thickness of about 100,000 Å (see A). The film from DCE gave a relatively homogeneous film about 30,000 Å thick (B). That from DO was homogeneous and thinner (C). The slide from DMF (D) gave a very spotty appearance that could not be considered a film.

c. The films on glass were examined by polarized microscopy on a Nikon Type 109 microscope. The figure below shows the appearance of the PBLG film out of DCE at 400x. The more viscous material from chloroform exhibited birefringence.
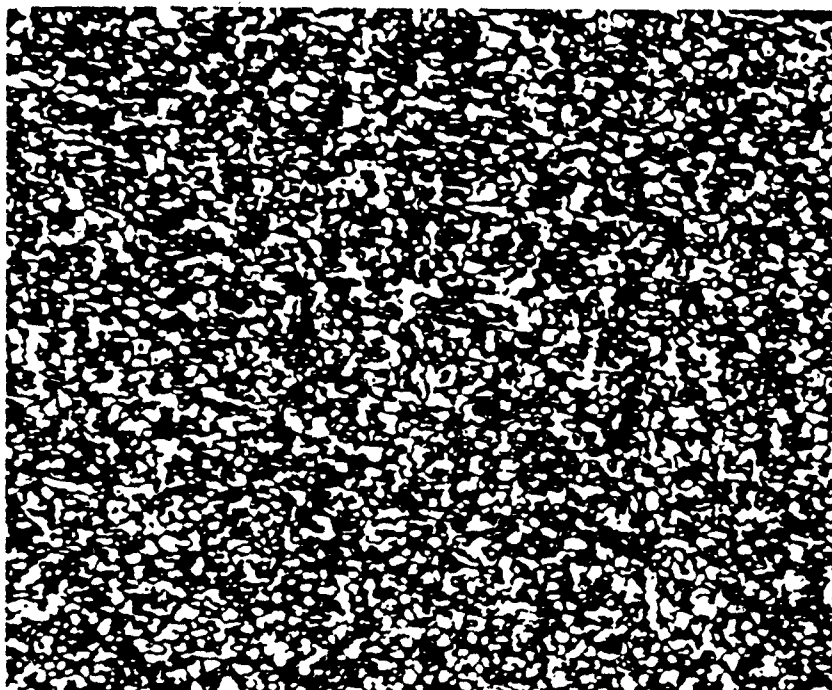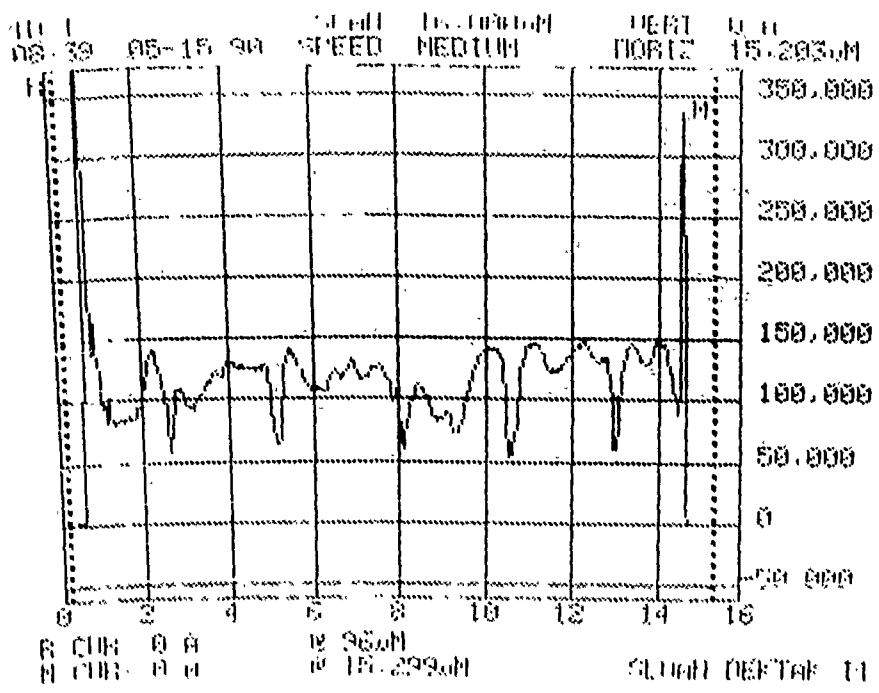
Figure 2A. Surface profile of PBLG mechanically sheared from 1,2-dichloroethane.
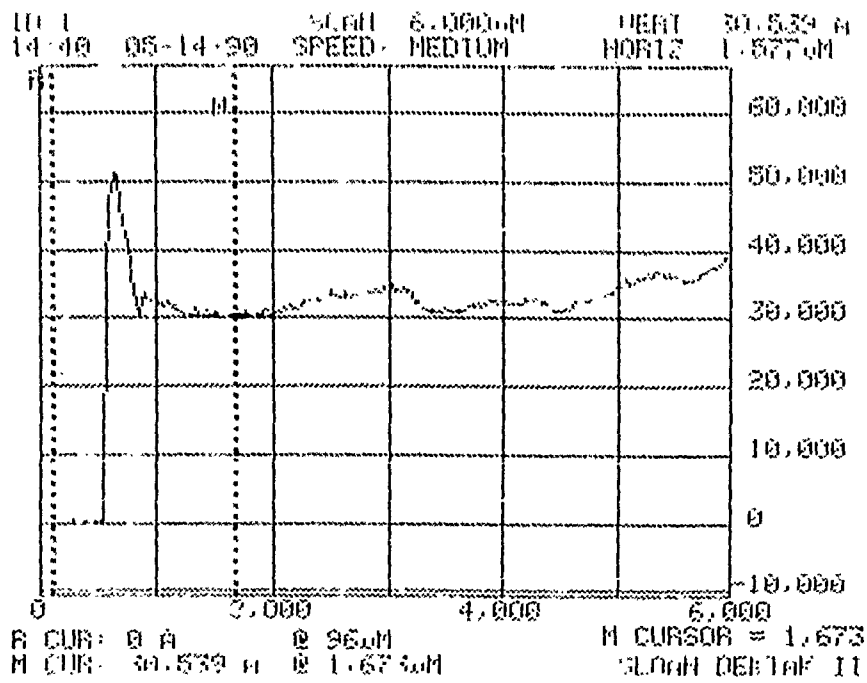
Figure 2B. Surface profile of PBLG cast from 1,2-dichloroethane by spin coating.

114-11
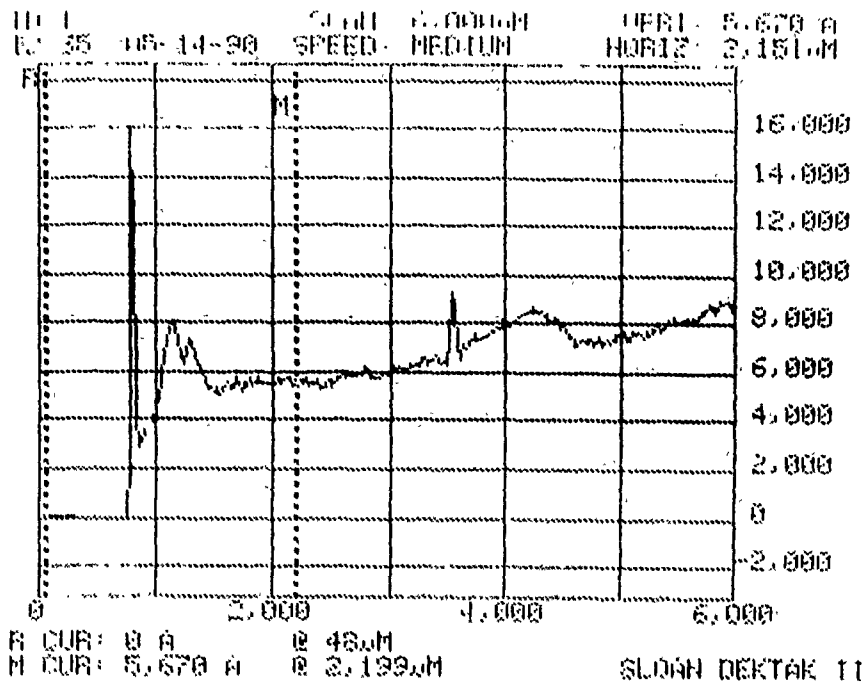
Figure 2C. Surface profile of PBLG cast from 1,4-dioxane by spin coating.
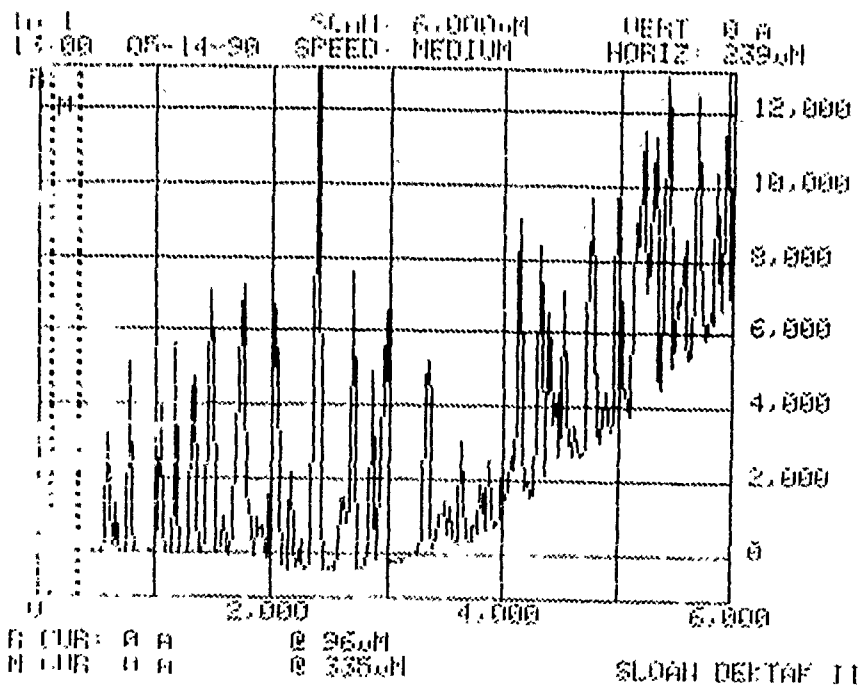
Figure 2D. Surface profile of PBLG cast from N,N-dimethylformamide by spin coating.

d. In order to carry out spectroscopic studies without interference from the glass substrate, films were cast on silicon and germanium wafers. Samples cast from DCE or DO solutions on silicon were examined by spectroscopic ellipsometry on the Rudolph s2000 spectrometer. Figures 3 and 4 on the next two pages give the resulting spectra. The interference pattern derives from the thickness of the film, and other features depend on the refractive index, as well as the thickness. These spectra are undergoing further analysis by Capt. J. Targove of this division.

e. Infrared spectra were recorded on samples cast on germanium or silicon with the Perkin-Elmer model 1725-X FT-IR spectrometer. Figure 5 is of PBLG cast from DO. Features causing rolling of the baseline are due to absence of background correction for the germanium disc. Figure 6 is of PALG cast from DO.

The major features of the spectrum of PBLG include the NH stretch at 3291 cm^{-1} and two carbonyl stretches at 1735 (ester) and 1653 (amide I) cm^{-1}. The latter feature is indicative of an α-helix structure.

The azo-modified derivative also has a strong NH stretch but a more complex carbonyl region. The small peak at 1735 cm^{-1} is probably from the carboxylic acid and the major peak at 1654 is the amide I band.
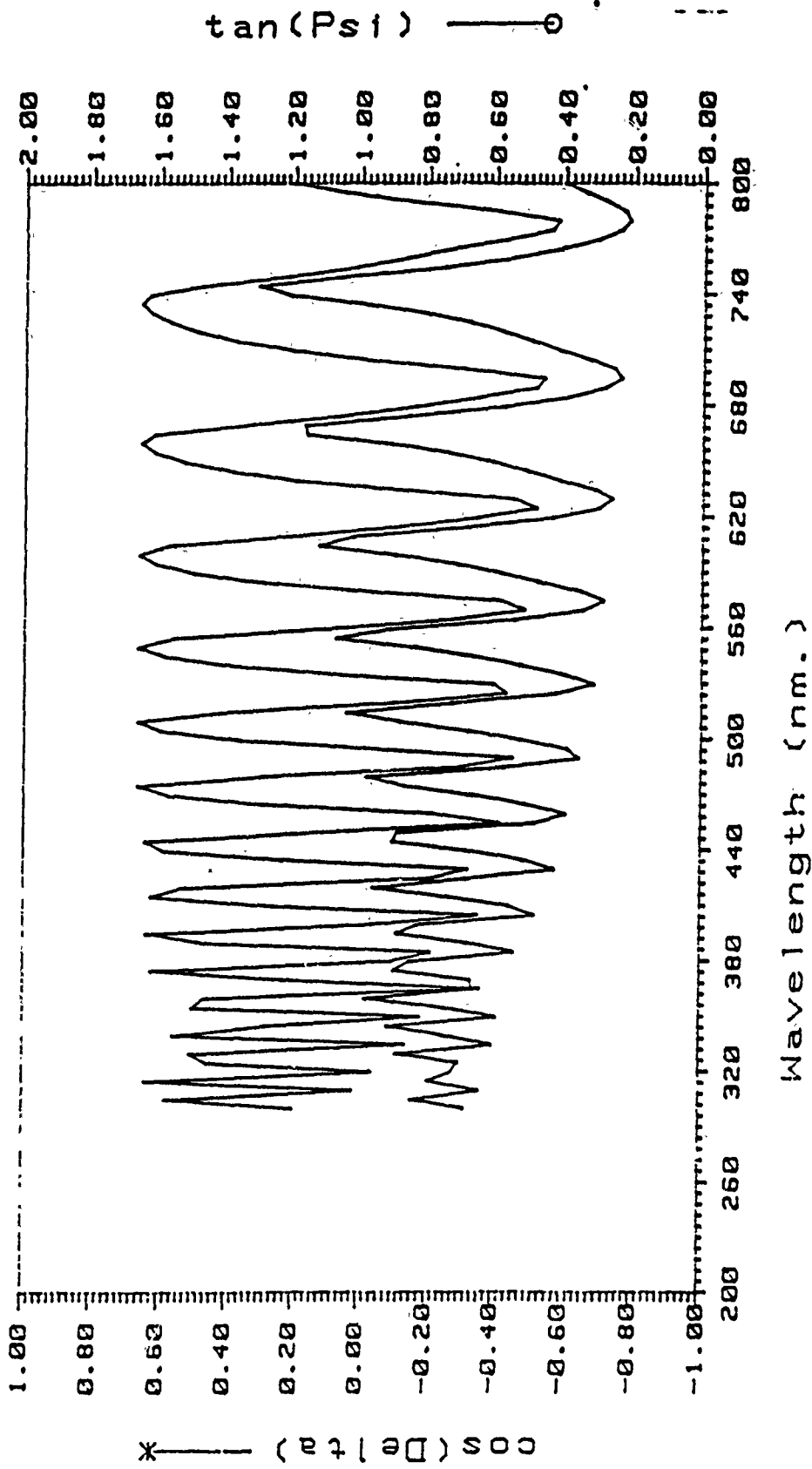
Figure 3. Spectroscopic ellipsometry of PBLG cast from 1,2-dichloroethane by spin coating.

114-14
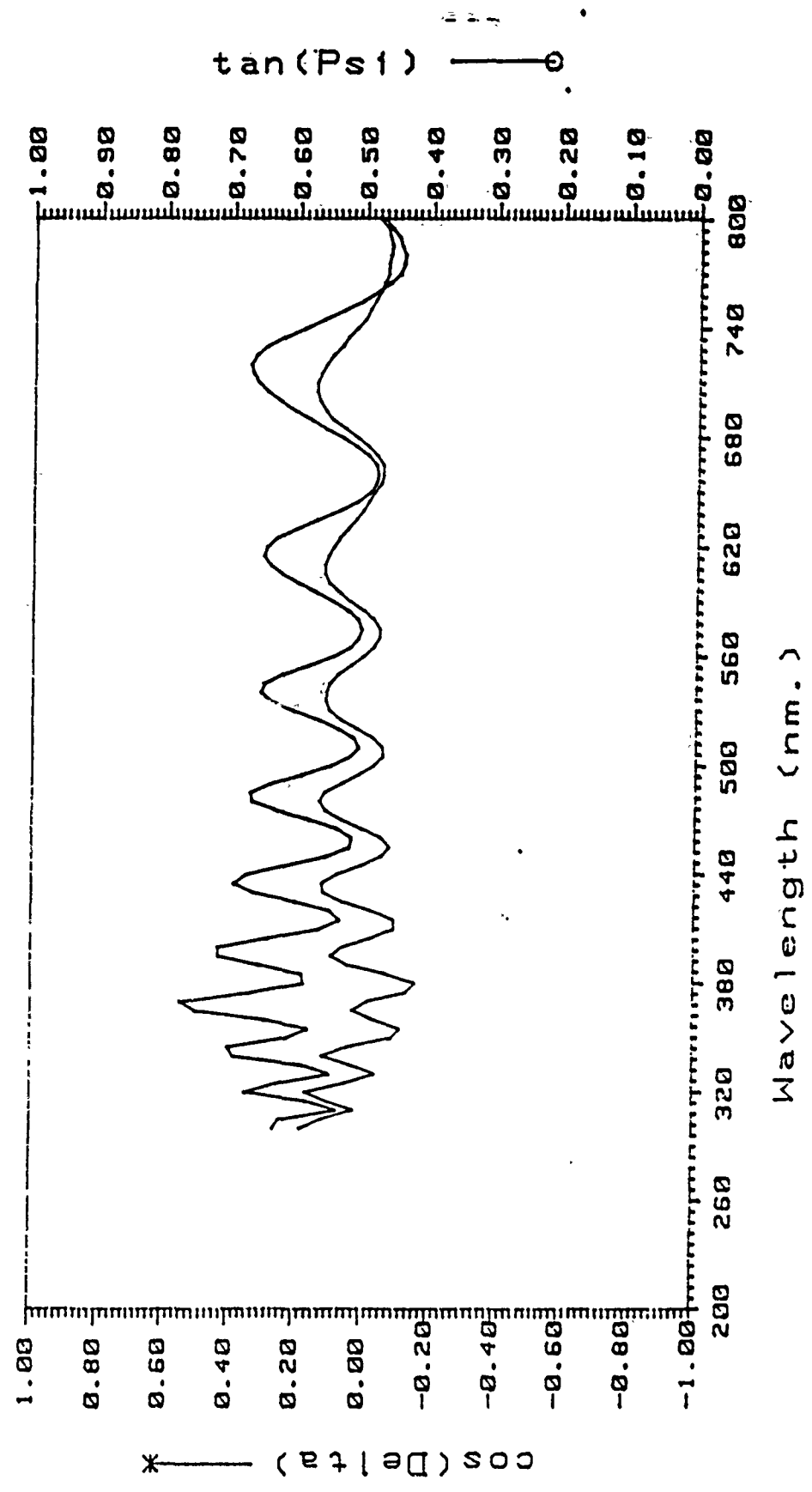
Figure 4. Spectroscopic ellipsometry of PBLG cast from 1,4-dioxane by spin coating.
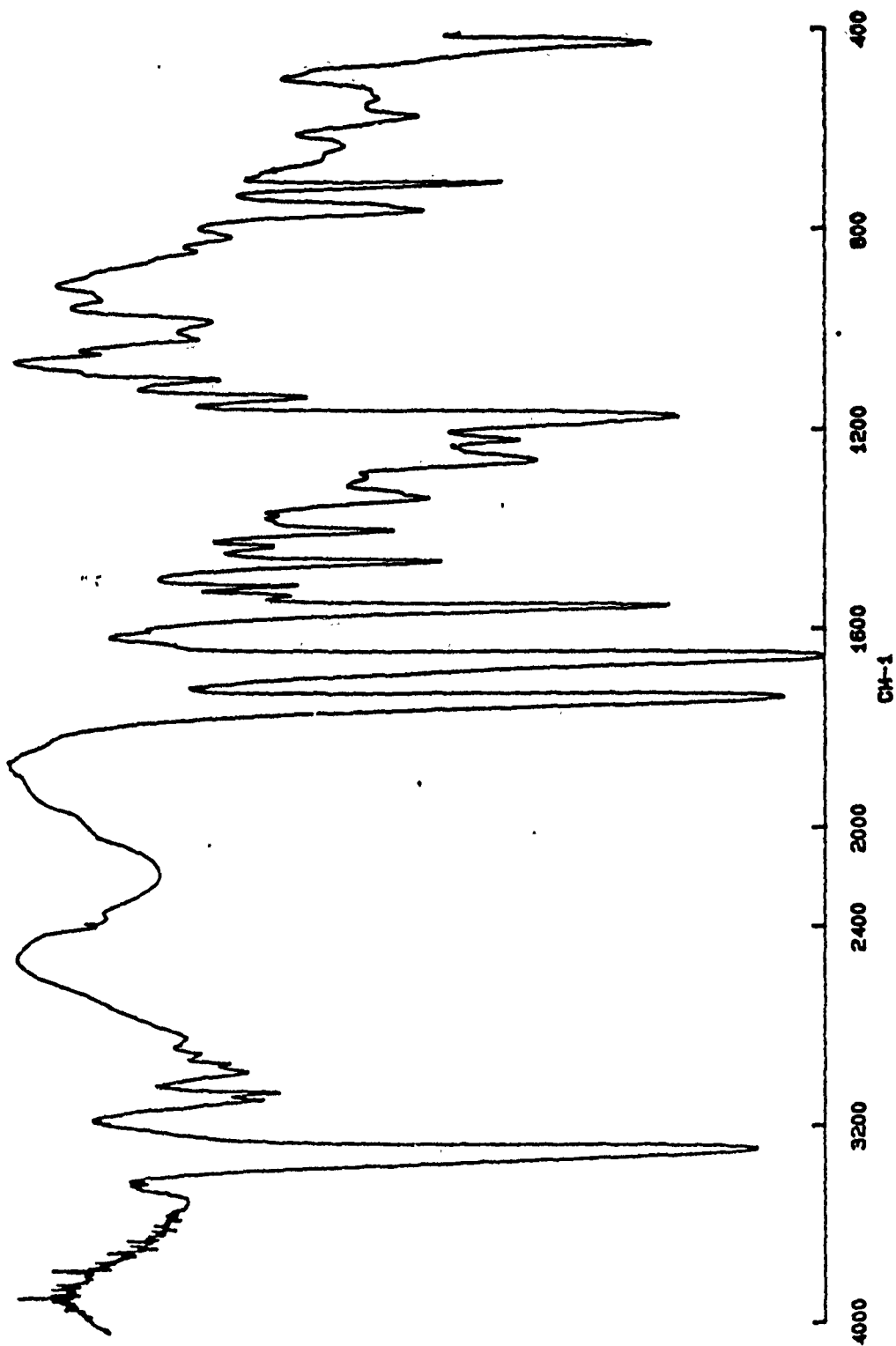
Figure 5.   Infrared spectrum of PBLG cast from 1,4-dioxane by spin coating.

114-16
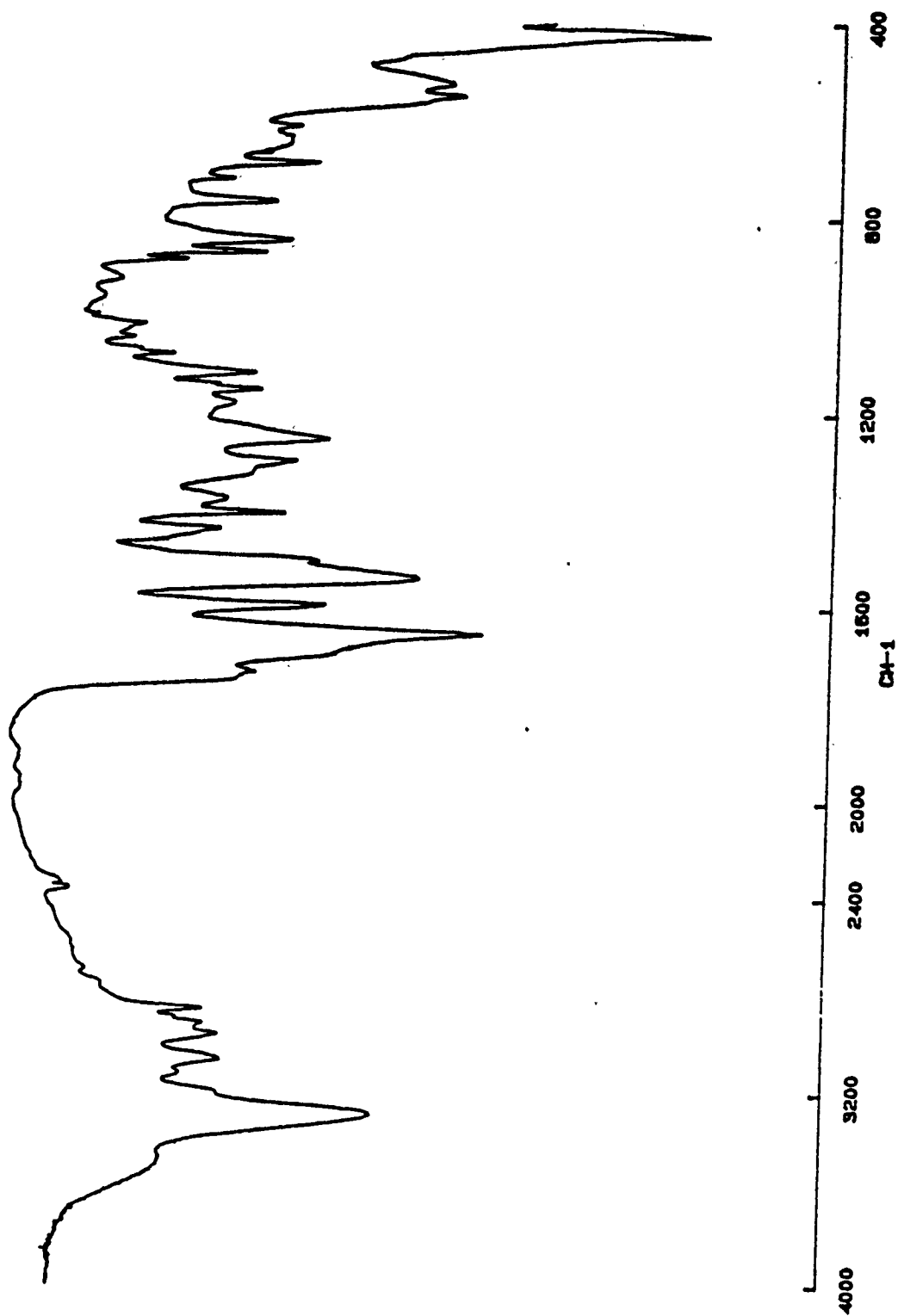
Figure 6.   Infrared spectrum of azo-modified polyglutamic acid (PALG) cast from 1,4-dioxane by spin coating.

114-17

IV. RECOMMENDATIONS

This work has demonstrated that homogeneous thin films of high polymer peptides derived from polyglutamic acid may be cast by spin coating techniques. Given a compatible substrate, these films may be characterized by a variety of techniques, including surface profiling, polarized microscopy, spectroscopic ellipsometry, and infrared spectroscopy. These films have a reasonably homogeneous thickness, which can be controlled by the spin coater speed. Consequently, their processibility properties appear to be excellent.

REFERENCES

Lambert. J. B., Wang, G.-t., Teramura, D. H., Interaction of the carbon-germanium or carbon-tin bond with positive charge on a β carbon, J. Org. Chem., 1988, Vol. 53, pp. 5422-5428.

Levine, B. F., Bethea, C. G., Hyperpolarizability of a polypeptide α-helix: poly-γ-benzyl-L-glutamate, J. Chem. Phys., 1976, Vol. 65, pp. 1989-1993.

Miller, R. D., Michl, J., Polysilane high polymers, Chem. Rev., 1989, Vol. 89, pp. 359-1410.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Chemical Induced Grain Boundary Migration in Al_2O_3

Prepared by:	Gary L. Leatherman
Academic Rank:	Assistant Professor
Department and	Mechanical Engineering
University:	Worcester Polytechnic Institute
Research Location:	WRDC/MLLM Wright-Patterson AFB Ohio 45433
USAF Researcher:	Randy Hay
Date:	November 26, 1990
Contract No:	F49620-88-C-0053

Chemically Induced Grain Boundary Migration in Al_2O_3

by

Gary L. Leatherman

ABSTRACT

Chemically induced grain boundary migration (CIGM) was observed in bulk polycrystalline alumina. The presence of gallia coupled with a bismuth oxide flux was able to induce the migration of grain boundaries in the alumina. Use of the same system for inducing grain boundary motion in alumina fibers with a "bamboo" microstructure proved unsuccesful due to experimental difficulties in observing the effect. Additional attempts were made to observe CIGM in these fibers using sols of Ga_2O_3, Fe_2O_3, and Cr_2O_3 doped with bismuth oxide. The microstructure of the fibers prevented in the time frame of the program the conclusive identification of CIGM.

Acknowledgements

I wish to thank the Air Force Office of Scientific Research for sponsoring this work. Universal Energy Systems was also of great help in the non-technical aspects of the Summer Fellow Program.

I. INTRODUCTION

Grain boundary migration which occurs in the absence of known driving forces in which the chemical composition changes are observed in the volume of the crystal swept by the migrating boundary is known as chemically induced grain boundary migration (CIGM) [1] or as diffusion induced grain boundary migration (DIGM) [2]. The driving force for this phenomenum is believed to be related to coherency strain energy anisotropies which develop along the grain boundary do to the presence of solute atoms. [3] Prior to the start of this work CIGM had not been observed in aluminum oxide.

Alumina fibers are currently being considered as the reinforcing element in oxide ceramic/ceramic composites. Although polycrystalline alumina fibers are quite inexpensive, their mechanical properties are not suitable for high performance composite applications. These fibers, after thermal treatment which occurs either in processing of the composite or in service, develop a segmented ("bamboo") grain structure. Once this structure is obtained there is no further driving force for conventional grain growth. The grain boundaries are quite weak; failure originates at these points. This limits the effective length of the fiber to the "bamboo" segment length which is not long enough to provide effective reinforcement to the matrix. Single crystal saphire fibers overcome this problem. However, they are extremely expensive. Current costs are of the order of hundreds of thousands of dollars per kilogram.

II. OBJECTIVES OF THE RESEARCH EFFORT

The development of a polycrystalline alumina fiber with sufficiently long bamboo segment length could be a low cost alternative to single crystal saphire fiber. CIGM may provide a mechanism by which further grain growth can be occur in alumina fiber. Thus providing a suitable segment length for the fibers resulting in adequate composite properties. This project examines CIGM in alumina. The first objective was to observe CIGM in bulk polycrystalline alumina. After the presence of CIGM in bulk alumina was established, the next step was to determine if CIGM occurs in bamboo structure alumina fiber.

III. RESULTS OF RESEARCH EFFORT

Specimens of polycrystalline alumina were placed in a bed of gallia powder doped with 10 weight percent Bi_2O_3. The specimens were heated to 1320°C and held for 24 hours. Tapered sections of the surface exposed to the powder bed were porduced using standard petrographic techniques. Figure 1 shows a typical section of an alumina specimen which underwent the above thermal treatment. It is typical of a CIGM microstructure. This is strong evidence that CIGM does occur in alumina. In addition, during the course of this work researchers in Korea published results also showing that CIGM occurs in alumina[4]. Thus the first objective of the program was reached.

Attempts to observe CIGM in alumina fibers using an identical approach were unsuccesful. The bismuth oxide flux created a film on the fibers that prevented observation of the microstructure. An alternative approach was attempted using sol-gel techniques. Sols based on Ga_2O_3, Fe_2O_3, and Cr_2O_3 were developed. Alumina fibers were then coated with these sols and thermally treated at 1320°C for 24 hours. Although the microstructure of these fibers could be observed under polarized light in transmission, direct evidence of net grain growth do to CIGM was not obtained. The large distribution in the grain size of the fibers before CIGM treatment made any before and after observations difficult to interpret without detailed statistical analysis of both the before and after grain size distributions. There was not sufficient time in the program to do this sort of analysis. Longitudinal cross sections of the fibers were examined by scanning electron microscopy (SEM) for compositional gradients typical of CIGM. Difficulties in producing quality samples as well as equipment problems and the time limit of the program led to inconclusive results.

IV. RECOMMENDATIONS

Although there is conclusive evidence for CIGM in alumina, there is not yet sufficient information on CIGM's presence or extent in alumina fibers. The following work should be done to answer this question. First, detailed grain counting for complete characterization of the grain size distribution in alumina fibers should be conducted before and after CIGM exposure. This should determine if on average there is any increase in bamboo

segment length due to grain boundary migration. Second, longitudinal cross section of fibers that have been exposed to the CIGM environment at a variety of depths should be produced and carefully examined using an electron microprobe. This should determine if CIGM is occuring in alumina fibers and if it is a useful process by which to induce grain growth in alumina fibers.

REFERENCES

1. Hillert, M. and Purdy, G.R., "Chemically Induced Grain Boundary Migration," _Acta Met._ 1978, Vol. 26, pp. 333-340.

2. Balluffi, R.W. and Cahn, J.W., "Mechanism for Diffusion Induced Grain Boundary Migration," _Acta Met._ 1981, Vol. 29, pp. 493-500.

3. Hillert, M., "On the Driving Force for Diffusion Induced Grain Boundary Migration," _Scripta Met._ 1983, Vol. 17, pp. 237-240.

4. Lee, H. and Kang, S., "Chemically Induced Grain Boundary Migration and Recrystallization in Al_2O_3," _Acta Met._ 1990, Vol. 39, pp. 1307-1312.
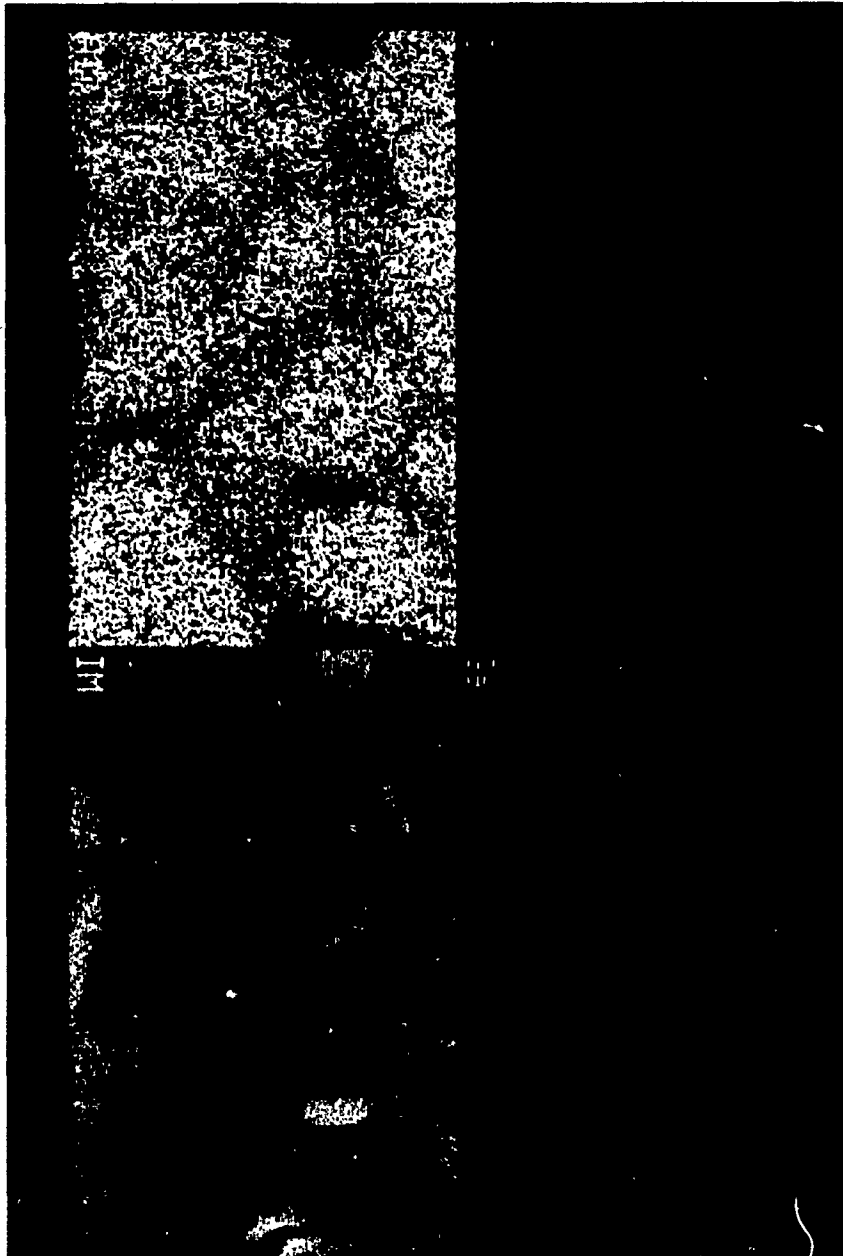
Figure 1: S.E.M. image and O, Al, and Ga E.D.S. maps of polycrystalline alumina sample exhibiting gallia driven chemically induced grain boundary migration.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

On the Use of QPA (Qualitative Process Automation) for Batch Reactor Control

Prepared by:	Won-Kyoo Lee, Ph.D.
Academic Rank:	Associate Professor
Department and	Department of Chemical Engineering
University:	The Ohio State University
Research Location:	WRDC/MLIM Wright-Patterson AFB WPAFB, OH 45433-6533
USAF Researcher:	Maj Steven R. LeClair
Date:	21 Aug 90
Contract No;	F49620-88-C-0053

On the Use of QPA (Qualitative Process Automation) for Batch Reactor Control

by

Won-Kyoo Lee

ABSTRACT

Control of batch reactors and the self-directed process control system, QPA, were reviewed to determine if the QPA system could be used for intelligent control of batch reactors. The control of batch reactors has been formulated as optimal control problems, with the solution being an open-loop temperature trajectory. However, this optimal temperature profile is based on very complicated, but still incomplete mathematical models to account for the unique nonlinear and time-varying dynamics of batch reactors. This means that the greatest remaining challenge in controlling batch reactors is to develop a totally adaptive control strategy that can result in the optimal operation using a minimum of mathematical models. In this regard, the capability of the QPA control is expected to be more beneficial for batch reactors, especially in the presence of process changes, and the dynamic, nonlinear nature of the batch reactors. It is suggested that the QPA system be tested to further demonstrate its concept and consequently extend its applicability by being applied to an experimental unit.

ACKNOWLEDGEMENTS

I wish to thank the Air Force Systems Command and The Air Force Office of Scientific Research for sponsorship of this research. Universal Energy Systems must be mentioned for their concern and help to me in all administrative and directional aspects of this program.

My experience was eye-opener, and subsequently I was able to broaden my background by exposing myself to on-going manufacturing research projects. Maj Steve LeClair provided me with support, and enjoyable research environment.

I. INTRODUCTION:

Batch processes have gained their importance in recent years as many chemical industries seek the manufacture of low-volume, high-value-added chemicals. In particular, batch reactors are extensively used in producing specialty chemicals, polymers, new drugs, etc., because of their great flexibility and rapid response to changing market conditions. For batch reactors, the high valued final product makes determination of an optimal control profile important, especially for maximization of some function of the composition (e.g., productivity, yield, or selectivity). These problems have been formulated as optimal control problems, with the solution being an open-loop temperature or flow-rate trajectory. However, these optimal control profiles are difficult to obtain because of the absence or excessive development cost of adequate models. Another factor that complicates the determination of these optimal control profiles is the presence of constraints on both the control and state profiles.

The Manufacturing Research Group of the Materials Laboratory at Wright Research and Development Center, Wright-Patterson Air Force Base, is concerned with the development of intelligent (self-directing and self-improving) systems for real-time control of advanced material processes. As a result, a self-directed process control philosophy, referred to as QPA (Qualitative Process Automation), was developed for the control of autoclave curing process of composite materials by making use of, not only qualitative physics, but also concurrent expert systems cooperating together. More specifically, the QPA system is capable of automating, not only the control of the process, but also the development of an optimized cycle. Further, because QPA develops the processing knowledge adaptively in situ, this capability is expected to be more beneficial for batch chemical reactors in the presence of process changes and the dynamic, nonlinear behavior of the batch reactors.

My research interests have been in the areas of adaptive control, on-line optimizing control, model-based control, statistical process control, neural networks, genetic algorithm and expert system framework for process control system synthesis, with a balance between theory and application. My work on adaptive strategies for the automatic startup and control of batch processes contributed to my assignment to the Manufacturing Research Group of the Materials Laboratory.

II.   OBJECTIVES OF THE RESEARCH EFFORT:

The increasing interest in the manufacture of low-volume specialty chemicals in batch reactors will continue to present many process control challenges. First, the frequent product and process changes associated with a batch chemical manufacturing facility requires an optimal startup policy for smooth and rapid transition to the desired operating conditions with little or no overshoot. The second type of problem is to control batch reactors to optimize composition functions such as productivity. These problems have been formulated as optimal control problems, with the solution being an open-loop temperature or flow-rate trajectory. Techniques are available to solve for the optimal control profiles, once good process and kinetics models are known. However, these optimal solutions are based on very complicated, but still incomplete mathematical models to account for the unique nonlinear and time-varying dynamics of batch processes. This means that in controlling batch reactors, often the greatest challenge is the development of a totally adaptive control strategy that requires a minimum of mathematical models and consequently can be applied to many batch processes.

My assignment as a participant in the 1990 Summer Faculty Research Program (SFRP) was to examine the attributes specific to batch reactor control and then determine if the self-directed process control system, referred to as QPA, researched and developed at the Materials Laboratory could be used for intelligent control of batch reactors.

III.   BATCH REACTOR CONTROL:

In general batch reactors are characterized by both very different manufacturing environment and different dynamic behavior. For example, the multiproduct environment of batch reactors requires that ingredients, control loop setpoints, and tuning parameters must be changed frequently. The frequent product and process changes demand good dynamic response over the entire operating range of the controlled variable for startup and shutdown regulatory control. Also, the wide operating ranges and nonstationary behavior that cause difficult sensor problems in batch reactors influence control system design. The controller design is further complicated by asymmetric penalties, such as in composition control where the formation of unwanted byproducts is irreversible. In the control of batch reactors, there are two related but distinct challenges.

(a) The first is the control of batch reactors with highly exothermic reactions. Safety-related concerns such as adiabatic runaway are primary, followed by productivity concern if the heat-release rate is far from constant.

(b) The second control challenge is that of controlling the batch reactor to optimize a composition function such as yield or selectivity.

There has been considerable interest in the past on the control of batch reactors. However, most of the previous work has focused on either the determination of the optimal temperature trajectory based on reaction kinetics alone or on the design of control systems for tracking this predetermined temperature profile. Also, in practice the operation of a batch reactor is formulated in terms of a temperature trajectory because temperature is the most readily available output. The trajectory is specified as a sequence of steps, consisting of three parts: (a) startup, where the reactor contents are brought from the initial charging conditions to the desired operating level; (b) maintenance of the desired nominal operating conditions for as long as it is beneficial to do so; (c) termination of the reaction according to

either optimality or product specifications considerations. Thus, the discussion of batch reactor control is divided into three parts. The first discusses the automatic startup and isothermal control of batch reactors. The second and third discuss the optimal control of batch reactors to maximize some function of composition.

A.   Automatic Startup and Isothermal Batch Reactor Control

For exothermic batch reactors that undergo frequent startups and shutdowns, the startup phase constitutes an important part of the batch cycle in enhancing productivity. It is, therefore, most often desirable to provide a smooth and rapid transition to the desired operating condition with little or no overshoot. This can be achieved by applying time-optimal control as an optimal startup policy. The time-optimal control involves switching of the control input from one constraint to the other at predetermined times. For example, the control input needs to be switched only once to achieve the most rapid transition for second-order, single-input/single-output systems, and the switching time can be calculated a priori or online from a second-order process model. Alternatively, PID controllers can also be used to follow a pre-determined time-optimal trajectory.

However, both time-optimal control with predetermined switching time and PID controller have difficulty in accommodating changing process characteristics that could be experienced from batch to batch or product to product. These process changes and the dynamic, nonlinear nature of batch reactors require recalculation of the switching times, the time-optimal trajectory or new settings of PID controller for satisfactory results. Moreover, for batch reactors with unknown or poorly understood dynamics, neither of these techniques can provide satisfactory control. To alleviate the problems with varying process dynamics, a model-predictive time-optimal control method and an adaptive time-optimal control algorithm were proposed for the startup control problem, respectively. Another

adaptive control techniques using a linear dynamic model have been applied to the startup problems for batch reactors since these adaptive control schemes provide systematic, flexible approaches for dealing with uncertainties, nonlinearities, and time-varying process parameters. In particular, the self-tuning controller (STC) of Clarke and Gawthrop was applied to control of batch reactors for the manufacture of synthetic latex of a variety of grades. The general control objective was to produce polymer of acceptable quality in the shortest practicable time. For this purpose, the batch temperature setpoint was predetermined for each grade, and a control scheme was used to follow the predetermined batch temperature profile as close as possible without overshoot.

Recently, an adaptive strategy based on the prediction error model was proposed to achieve desired temperature profiles during startup of a continuous fluidized sand bath, and then later extended to the automatic startup and control of a batch heating process and a simulated exothermic batch reactor. Prediction error models contain information about the quality of the prediction, and thus provide better predictions of the future output variable, which is used for on-line calculation of the startup switching time. For regulation after the setpoint is reached, a self-tuning controller with the PID structure was applied to obtaining PID controller settings, and a long-range predictive control strategy was also proposed by incorporating the prediction error model in an extended horizon approach. Results have shown the feasibility of the adaptive approach based on prediction error models for the rapid startup and control of exothermic batch reactors. However, there are limitations and potential operating problems associated with current adaptive control schemes. The biggest drawback of adaptive control techniques is the inherent linear nature of an assumed dynamic model. Moreover, the range of uncertainty may be substantially greater than can be tolerated by existing algorithms for adaptive systems.

B.   Nonisothermal Batch Reactor Control

Although the previous discussion of regulatory temperature control has focused on startup and isothermal reactor operations, nonisothermal temperature trajectories are a common means of optimizing the throughput of many batch reactors. This is because the productivity is limited by the maximum heat-generation rate, and this typically occurs for only a short time. For exothermic reactions, the temperature trajectories are often calculated to maintain a constant reaction rate, which is within a specified safety limit of the maximum heat-removal capacity of the reactor. Determination of the constant-rate temperature trajectories requires a mathematical rate expression, and further the solution of variational optimization problem or experimental determination is required when rate equations are complex. These temperature profiles are then implemented as a series of open-loop setpoint changes in the reactor temperature controller. A major problem with this open-loop strategy is that induction effects in the chemical reactions introduce uncertainties in the timing of the setpoint changes.

Alternative approach to maintaining constant-rate operation in the face of unknown or uncertain kinetic is to obtain a real-time measurement of the reaction rate. However, on-line analytical instruments that can measure properties of interest are often not available. Since almost all reactors are or can be well instrumented with temperature and flow measurements, these measurements can be combined with a set of dynamic energy balance equations for the reactor in order to monitor in real time the instantaneous heat release due to reaction. Form this the instantaneous rate of reaction can be evaluated, and subsequently one can obtain an estimate of the conversion of material in the reactor at any instant of time. These quantities are then very useful in monitoring the reactor for potentially dangerous conditions, in improving temperature control, and in improving the control of composition and other quality variables.

C.   Control of Composition in Batch Reactors

The objective here is to manipulate the temperature in order to optimize some function of the composition (e.g., productivity, yield, or selectivity). These problems have been formulated as optimal control problems, with the solution being an open-loop temperature trajectory. Techniques are in hand to solve for the optimal temperature trajectory once good process and kinetic models are known. However, these optimal control profiles are difficult to obtain because of the absence or excessive development cost of adequate mathematical models, especially for cases where the reaction kinetics and mechanisms are unknown or poorly known. Again, a major problem with implementing this open-loop strategy is that variations in the process variables and parameters during each cycle and/or from cycle to cycle require recalculation of the optimal control trajectory, as with the control of nonisothermal batch reactors.

IV.   JUSTIFICATION OF QPA FOR USE IN BATCH REACTOR CONTROL:

As discussed previously, the optimal operation of batch reactors has been formulated as optimal control problems, with the solution being an optimal temperature profile based on mathematical models. And then, control systems have been designed to guide the batch reactors along this trajectory. There is a major problem with this open-loop strategy in that mathematical solutions are limited in accommodating changing process conditions and/or variables which can not be accounted for. It is thus clear that there is a great need to develop an on-line control algorithm that adaptively results in optimum operation in the presence of process changes and the dynamic, nonlinear nature of batch reactors, by using a minimum of mathematical models.

QPA is an intelligent process control system based on "goal-driven" control philosophy, which is capable of developing the processing knowledge adaptively in situ, generating an optimized cycle, and automating the control of the process to achieve the desired product goal. These capabilities are made possible by the use of qualitative physics because of the reduced computational load on the control system and the ability to accommodate both uncertain and nonlinear process characteristics through qualitative relationships. More specifically, the QPA system autonomously interprets heterogeneous sensor data and represents that data qualitatively using process-specific knowledge coupled with expert heuristics, and reasons about the data to construct a processing plan in real time. In summary, QPA offers the capability which are considered to be desired features of a totally adaptive control system for batch reactors.

A prototype QPA system for autoclave cure of graphite-epoxy composites has been demonstrated to successfully control the autoclave process for making advanced composite materials. In addition, this QPA system has demonstrated dramatic improvements in material processing times as a result of its ability to adapt on-line to material processing behavior and to generate tailored cure cycles. It should be noted that the autoclave is a batch reactor which requires a startup period for heating of the matrix resin to start flow and chemical reaction, followed by a cure period for completing of the reaction and a cooldown period for removing of the cured laminate. The control objective was to complete the process as rapidly as possible without damage by the potentially excessive heat generated by the reaction itself. As described above, the autoclave curing of graphite-epoxy exhibits the same process characteristics and thus has the same control problem domain issues as addressed in controlling the exothemic batch reactors.

V. RECOMMENDATIONS:

Real-time control of batch reactors is difficult by conventional control theory because of the absence or excessive development cost of adequate mathematical models to account for uncertainties, nonlinearities and time-varying process parameters. Another factor that complicates the use of conventional control theory is the existence of transitions between processing stages with different kinds of phenomena and physical objectives. The capability of the QPA system and its proven application to the autoclave curing of composite materials being considered, control of batch reactors is expected to greatly benefit from QPA.

A follow-on research is suggested on the use of QPA for the control of batch reactors to further demonstrate the concept of QPA and extend its applications. The QPA will be tested by being applied to an experimental batch reactor, possibly a batch polymerization reactor, under the sponsorship of an AFOSR Mini Grant. To this end, the QPA software will be obtained from Universal Technology Corporation, Dayton, Ohio, and effort will be made to seek an industrial partner.

REFERENCES

Clarke, D.W., and P.J. Gawthrop. Self-Tuning Controller. Proc. IEE Part D. 1975, Vol. 122, pp. 929-934.

Clarke, D.W., and P.J. Gawthrop. Self-Tuning Control. Proc. IEE Part D. 1979, Vol. 126, pp. 633-640.

Cuthrell, J.E., and L.T. Biegler. Simultaneous Optimization and Solution Methods for Batch Reactor Control Profiles. Computers Chem. Engng. 1989, Vol. 13, pp. 49-62.

De Keyser, R.M.C., and A.R. Van Cauwenberghe. A Self-Tuning Multistep Predictor Application. Automatica. 1981, Vol. 17, pp. 167-174.

Hodgson, A.J.F., and D.W. Clarke. Self-Tuning Applied to Batch Reactors. Proc. IEE Vocation Sch. Ind. Digital Control System, IEE, London, 1984.

Hsu, E.H., S. Bacher, and A. Kaufman. A Self-Adapting Time-Optimal Control Algorithm for Second-Order Processes. AIChE J. 1972, Vol. 18, pp. 133-139.

Juba, M.R., and J.W. Hamer. Progress and Challenges in Batch Process Control. Proc. of the Third International Conference on Chemical Process control, Asilomar, CA, 1986.

Koppel, L.B., and L.R. Latour. Time Optimal Control of Second-Order Overdamped Systems with Transportation Lag. Ind. Engng. Chem. Fundam. 1965, Vol. 4, pp. 463-471.

Latour, P.R., L.B. Koppel, and D.R. Coughanowr. Time-Optimum Control of Chemical Porocesses for Set Point Changes. Ind. Engng. Chem. Process Des. Dev. 1967, Vol. 6, pp. 452-460.

Latour, P.R., L.B. Koppel, and D.R. Coughanowr. Feedback Time-Optimum Process Controllers. Ind. Engng. Chem. Process Des. Dev. 1968, Vol. 7, pp. 345-353.

LeClair, S.R., and F.L. Abrams. Qualitative Process Automation. Int. J. Computer Integrated Manufacturing. 1989, Vol. 2, pp. 205-211.

LeClair, S.R., F.L. Abrams, and R.F. Matejka. Qualitative Process Automation: Self-Directed Manufacture of Composite Materials. AI EDAM. 1989, Vol. 3, pp. 125-136

Lewin, D.R., and R. Lavie. Designing and Implementing Trajectories in an Exothermic Batch Chemical Reactor. Ind. Eng. Chem. Res. 1990, Vol. 29, pp. 89-96.

MacGregor, J.F. On-Line Reactor Energy Balances via Kalman Filtering. Preprints of 6th International IFAC/IFIP/IMEKO Conf. on Instrumentation and Automation in the Paper, Rubber, Plastics, and Polymerization Industries, Akron, OH, October 1986, pp. 27-31.

Manzini, R.A., and E.A. Roehl. Flexible Control of an Organic Matrix Composite Cure Process Using Object-Oriented Control Concepts. Proc. 1990 American Control Conference, San Diego, CA, May 1990, pp. 1980-1985.

Mellichamp, D.A. Model Predictive Time-Optimal Control of Second-Order Processes. Ind. Eng. Chem. Process Des. Dev. 1970, Vol. 9, pp. 494.

Merkle, J.E., and Won-Kyoo Lee. Adaptive Strategies For Automatic Startup and Temperature Control of a Batch Process. Computers Chem. Engng. 1989, Vol. 13, pp. 87-103.

Ozkan, A.T., and Won-Kyoo Lee. Adaptive Strategies Applied to Automatic Startup and Control of an Exothermic Batch Reactor. Proc. 1989 American Control Conference, Pittsburge, PA, June 1989, pp. 1834-1839.

Pardee, W.J., M.A. Shaff, and B. Hayes-Roth. Intelligent Control of Complex Materials Processes. AI EDAM. 1990, Vol. 4, pp. 55-65.

Phillips, S., D.E. Seborg, and K.J. Legal. Adaptive Control Strategies or Achieving Desired Temperature Profiles during Process Startup. 1. Model Development and Simulation Studies. Ind. Eng. Chem. Research. 1988, Vol. 27, pp. 1434-1443.

Phillips, S., D.E. Seborg, and K.J. Legal. Adaptive Control Strategies for Achieving Temperature Profiles during Start-up. 2. Experimental Application. Ind. Eng. Chem. Research. 1988, Vol. 27, pp. 1444.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

conducted by the

Universal Energy Systems, Inc.

<u>FINAL REPORT</u>

<u>ULTRASONIC TECHNIQUES FOR AUTOMATED DETECTION OF FATIGUE MICROCRACK INITIATION AND OPENING BEHAVIOR</u>

Prepared by: Michael T. Resch, Ph.D.

Academic Rank: Assistant Professor

Department and   Engineering Mechanics

University: University of Nebraska-Lincoln

Research Location: WRDC/MLLN
                    Building 655, Room 023
                    Wright-Patterson AFB, OH 45433

USAF Researcher: Theodore Nicholas, Ph.D.

<u>ULTRASONIC TECHNIQUES FOR AUTOMATED DETECTION OF</u>

<u>FATIGUE MICROCRACK INITIATION AND OPENING BEHAVIOR</u>

by

Michael T. Resch, Ph.D.

ABSTRACT

A surface acoustic wave non-destructive evaluation technique was used to detect the natural nucleation of surface microcracks in highly stressed regions of hourglass shaped aluminum specimens during fatigue cycling. The experimental procedure involved excitation of Rayleigh waves on the surface of each specimen and observation of the presence of a specular reflection from the nucleating crack superimposed on nonspecular reflections from microstructural features surrounding the flaw. Contacting wedge transducers were used to excite the incident waves and to detect the reflected wave signals. The effectiveness of a split-spectrum processing algorithm to improve the minimum detectable crack size of isolated cracks in the scattering field was demonstrated. Additionally, measurements of crack opening behavior were performed both acoustically and with the laser interference displacement gage. Initial results indicate that the acoustic technique is more sensitive to small traction forces on adjacent crack faces than is the laser interference technique.

Acknowledgements

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. Universal Energy Systems must be mentioned for their concern and help to me in all administrative and directional aspects of this program.

My experience was rewarding and enriching because of many different influences. Both Ted Nicholas and Jay Jira of the high temperature metals and ceramics group went out of their way to ensure that I was supplied with all the instrumentation and technical help necessary to address my stated research goals and objectives. Special thanks are due to Prasanna Karpur for sharing his expertise in signal processing of the acquired signals. All split-spectrum processing of the data obtained in these experiments was performed courtesy of Dr. Karpur utilizing his personal software. Finally, the comeraderie and experimental expertise of Rick Kleismet of U.D.R.I. was invaluable in the preparation of fatigue specimens and operation of the laser interference displacement gage equipment.

I. INTRODUCTION:

The reason for developing new quantitative nondestructive evaluation techniques to measure the size and opening behavior for surface microcracks is that in the so called small crack size regime small cracks have been observed to grow at rates which are orders of magnitude higher than large sized cracks when subjected to identical magnitudes of crack driving force. Quantitative measurements of crack depth below the surface for surface microcracks facilitate the evaluation of crack growth rate v. the number of cycles. Nondestructive measurements of crack opening behavior are especially important here because many current theories which address the issue of why small cracks grow faster postulate that small cracks have less closure than large cracks, resulting in a higher driving force for growth for small cracks.

The High Temperature Metals and Ceramics Branch of the Wright Research and Development Center is particularly concerned with the creation of practical techniques to facilitate automation of fatigue testing of materials important to the mission of the Air Force both under laboratory conditions and on structures operating in the field. The expertise of this branch in the development of innovative and useful techniques to characterize the kinetics of growth of microscopic surface fatigue cracks in metals, ceramics, and composites is well documented in the international literature.

My research interests have been in the area of development of surface

acoustic wave scattering techniques to measure the size, growth rate, and opening behavior of small surface cracks [1,2]. Therefore, a project concerning evaluation of the feasiblity of using this technique in automated fatigue crack experiments seems a logical extension of the technology.

II. OBJECTIVES OF THE RESEARCH EFFORT:

Currently, nondestructive detection of initiation of surface microcracks in hourglass shaped specimens now in use for fatigue crack initiation studies in the High Temperature Metals and Ceramics Branch of the Wright Research and Development Center (WRDC) is accomplished by optical scanning at high magnification in a standard metallographic microscope. This process is exceedingly time consuming, requiring periodic inspections of the specimen performed by physically taking the specimen out of the servohydraulic machine, scanning many square millimeters of surface area in the high stress region, and replacing and realigning the specimen in the hydraulically actuated grips before resuming the test. Additionally, during crack mouth opening displacement measurements of already nucleated fatigue microcracks with a Laser Interference Displacement Gage (LIDG), there is uncertainty concerning the exact amount of tensile stress to be applied to the specimen to cause the adjacent crack faces of the microcrack to become traction free.

Surface Acoustic Wave (SAW) scattering has the potential to contribute to this effort in two ways. First, a beam of surface acoustic waves

directed toward the high stress region of a specimen interrogates the total area on one side of the specimen where microcrack initiation technique is expected to occur. Measurement of the reflection of the waves is accomplished in a few seconds during a single measurement without removing the specimen from the testing machine. Surface microcracks as small as 50 micrometers deep have been detected using this technique [1,2]. Second, measurement of the reflected amplitude of surface acoustic waves from a surface microcrack as a function of the amplitude of applied tensile stress normal to the plane of the adjacent crack faces has demonstrated that ultrasonic measurements of opening behavior correlate well with measurements of Crack Mouth Opening Displacement (CMOD) and Crack Tip Opening Displacement (CTOD) performed under stress in a scanning electron microscope [2].

The chief physical phenomenon which stands as an obstacle to the utilization of the SAW scattering technique in these experiments is that the metallurgical features in complex alloys contribute to the scattered signal which also contains the reflected echo from a nucleating fatigue crack. The primary goal has been to acquire scattered signals from the high stress region of hourglass shaped fatigue specimens at periodic intervals, and examine the acquired waveforms for evidence of earliest possible detectability of reflections from surface microcracks when the amplitude of the reflection is obscured by microstructural scattering. The secondary goal is to compare measurements of crack opening behavior using the SAW scattering technique with measurements obtained using the LIDG technique.

III. NEW HOURGLASS SHAPED SPECIMENS

a.   Design and fabricate new hourglass shaped specimens similar to specimens now in use in the high temperature metals and ceramics laboratory which are suitable for; electropolishing, gripping by self aligning hydraulic grips in use at WRDC, and which allow use of new surface acoustic wave wedge transducers developed by this author at the University of Nebraska-Lincoln.

b.   Specimens were produced with a configuration used by Larsen et al.[4], except with a major cross section of 12.5 mm by 12.5 mm, and a reduced section of 3 mm by 12.5 mm, giving a geometrical stress concentration factor of 1.2 . Additionally, the specimens are initially machined with 37.5 mm by 50 mm rectangular ends suitable for insertion in standard compact tension fracture mechanics clevis grips using 12.5 mm diameter pins. This allows fatigue initiation experiments to be conducted at locations which do not yet possess self aligning hydraulic grips. The enlarged areas may be cut off subsequent to fatigue initiation experiments to facilitate insertion in hydrauic testing machines which feature a laser interference displacement gage for additional measurements concerning crack opening behavior.

IV. SWITCHING SYSTEM FOR TWO DUAL-ELEMENT ARRAYS

a.   Build switching box to accomodate two dual-element transducers to enable alternate surface acoustic wave scanning of both sides of the

high stress areas on hourglass shaped specimens during fatigue cycling.

b.   A switch box was designed and fabricated to allow manual switching between two sets of dual element surface acoustic wave contacting wedge transducers using a double pole, double throw manual switch.   This enabled both sides of a fatigue specimen to be interrogated at periodic intervals during fatigue cycling without disturbing the contact between the transducers and the specimen.

V. NATURAL INITIATION OF FATIGUE MICROCRACKS

a.   Produce naturally initiated cracks during fatigue cycling on hourglass specimens of aluminum using the SAW scattering technique as the primary feedback mechanism for detecting microcrack initiation.

b.   An hourglass shaped specimen is metallographically prepared using conventional wet grinding. diamond polishing, and electropolishing techniques in the high stress region to minimize the effect of fabrication on surface roughness and residual stress. The signal due to backscattered Rayleigh waves from the high stress region is acquired with a digitizing oscilloscope during application of a tensile stress of 275 MPa . Then at 2000 cycle intervals of applied fluctuating stress (maximum stress equal to 275 MPa with a stress ratio of 0.1) this measurement is repeated. This procedure continues until at least one crack is produced in the high stress region which just exceeds the maximum amplitude of the microstructurally induced interference pattern.

A crack is relatively easy to detect in the interference pattern because its reflection is only visible with an applied tensile stress applied to the specimen which is large enough to completely separate adjacent crack faces.

VI. WAVEFORM TRANSFER SOFTWARE

a. Obtain digitized waveforms from crack initiation area at regular intervals during cycling and transfer them to the hard disk of an IBM PC-AT for subsequent transfer to the VAX system using PROCOMM software.

b. A program was written in BASICA language to transfer the signal obtained from a Hewlett Packard 54201A digitizing oscilloscope to the screen of a CGA monitor and write the 1001 data points in the signal to the hard disk utilizing a HP-IB card designed for the IBM PC-AT. Each data file was then transferred to the VAX computer system for subsequent split spectrum processing.

VII. SPLIT SPECTRUM PROCESSING

a. Use Split Spectrum Processing (SSP, a nonlinear digital signal processing technique as implemented by Dr. P. Karpur of UDRI) to determine the absolute minimum detectable size and the number of cycles at which this size is detected [5,6].

b.   After a single crack is detected in the scattered field, the specimen is removed from the testing machine and examined at high magnification in a metallographic microscope. The length of the crack at the surface is measured using a calibrated eyepiece. This single measurement of crack size is then used to obtain the scattering parameters necessary for scaling the size of the crack from ultrasonic scattering theory [1]. Split spectrum processing is performed on the set of all waveforms obtained during periodic examination of the specimen to remove the nonspecular reflections of microstructural origin. This process reveals the size of the specular reflections from the crack obtained from each cycling interval during the experiment. The minimum detectable crack size for a naturally initiated surface microcrack in 2024 aluminum using this technique was determined to be 20 micrometers [3].   This is substantially smaller than the microstructurally limited minimum detectable crack size of 80 micrometers previously reported in the literature for this material [1].

VIII. SAW NDE SOFTWARE FOR OPENING MEASUREMENTS

a.   Develop software for computerized data acquisition of Surface Acoustic Wave NonDestructive Evaluation (SAW NDE) information concerning the amplitude of the reflected echo from the crack as a function of applied stress during application of stress to the specimen utilizing a servohydraulic testing machine.

b.   A program was written in BASICA language to transfer the amplitude

of a time gated signal containing the reflection from an isolated

surface microcrack obtained from a Hewlett Packard 54201A digitizing

oscilloscope utilizing a HP-IB   card designed for the IBM PC-AT.

Additionally, a DATA TRANSLATION 2818 analog to digital conversion

system was used to detect the signal from the load cell of the

servohydraulic testing machine. Under manual control of the set point

of the servohydraulic system in load control, applied force v. crack

amplitude could be acquired automatically point by point, and plotted

graphically on the CGA screen of the PC-AT.   The results of a typical

scan from an isolated surface microcrack are displayed in Figure 1. The

nonlinear portion of the curve from zero force up to the point at which

the plot becomes vertical reveals the presence of surface tractions

between adjacent crack faces. The vertical portion reveals the point at

which surface tractions disappear, and the crack is fully open.

IX.   SAW NDE v. LIDG MEASUREMENTS OF CRACK OPENING

a.   Obtain force v. CMOD for naturally nucleated surface microcracks in

specimens of aluminum   using the Laser Interference Displacement Gage

(LIDG) at identical levels of force as the ultrasonic measurments and

compare the experimental results.

b.   Force v. CMOD for several naturally initiated surface microcracks

was accomplished using the LIDG apparatus available on machine #1 in the

high temperature metals and ceramics laboratory at WRDC.   A typical
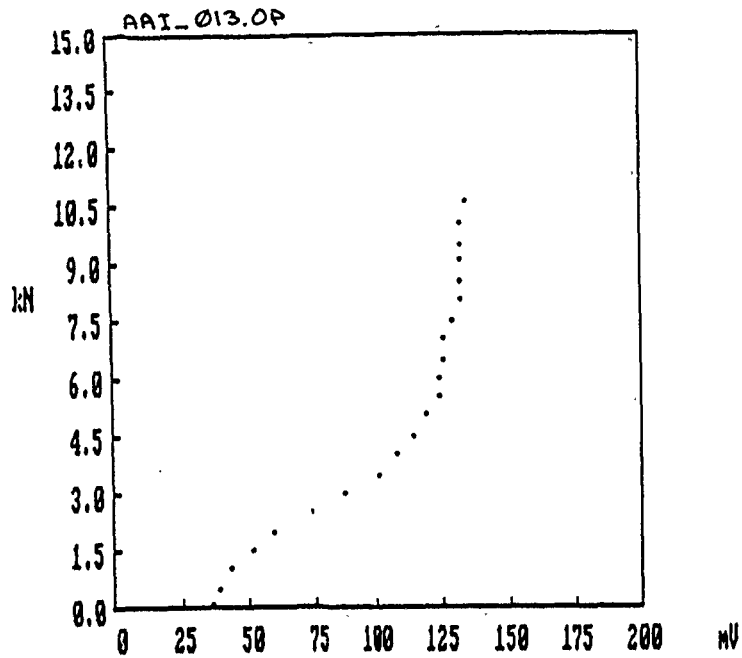
Figure 1.   Plot of crack reflection amplitude v. force applied normal to
the crack face obtained from the SAW NDE scattering technique.
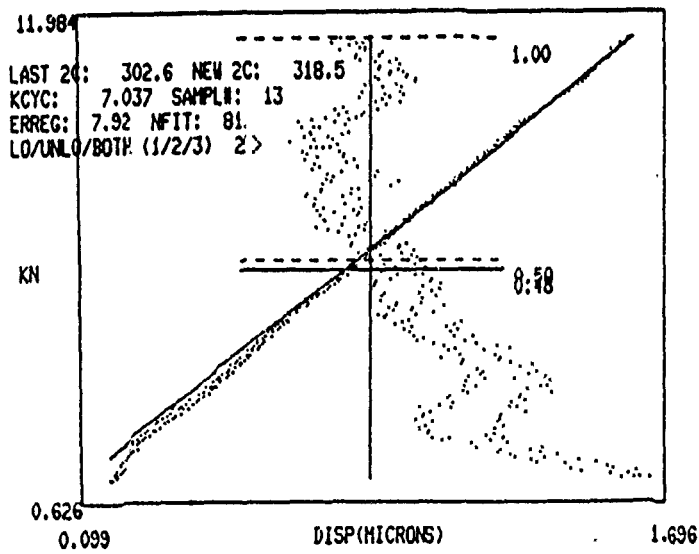
Figure 2.   Plot of crack mouth opening displacement v. force applied
normal to the crack face obtained from the LiDG technique.

result of this type of measurement obtained for the crack which was also ultrasonically measured is shown in Figure 2. An important feature is the observed linear relationship between CMOD and applied stress which is distinctly nonlinear in the ultrasonic measurement of opening behavior of this crack shown in Figure 1.

X. RECOMMENDATIONS:

a. Periodic measurements of surface acoustic wave scattering from the high stress region of fatigue specimens have demonstrated that the split spectrum processing technique is capable of significantly decreasing the minimum detectable crack size of naturally initiated surface microcracks. Comparison of ultrasonic measurements of naturally initiated surface microcracks with laser measurements reveals that the ultrasonic technique appears to be more sensitive to the existence of small tractions on adjacent crack faces than the laser technique. Consequently, these results demonstrate that the surface acoustic wave scattering technique provides useful information during automated microcrack initiation experiments, and may be easily integrated into the testing program of the high temperature metals and ceramics laboratory for further research and development.

b. Development of the technique is limited by the lack of contacting wedge transducers small enough for a dual element array to fit on the standard sized 7.5 mm wide specimens used in the high temperature metals and ceramics laboratory. Consequently, design and fabrication of new

miniature dual element contacting wedge transducer arrays is necessary for implementation of the technique during automated detection of microcrack initiation on standard sized specimens. This will involve experimental research on the physics of wave generation in commercially available piezoelectric ceramics in order to characterize the maximum allowable power levels which can be achieved, in addition to the application of the principles of physical acoustics in order to establish the optimum level of miniaturization possible using available materials.

REFERENCES

1   Resch, M.T., Nelson, D.V., Yuce, H.H., and Ramusat, G.F., "A Surface Acoustic Wave Technique for Monitoring the Growth Behavior of Small Surface Fatigue Cracks", Journal of Nondestructive Evaluation, Vol. 5, No. 1, 1985, pp. 1-7.

2   Resch, M.T., Nelson, D.V., Yuce, H.H. and London, B.D., "Use of Nondestructive Evaluation Techniques in Studies of Small Fatigue Cracks", Basic Questions in Fatigue: Volume I, ASTM STP 924, J.T. Fong and R.J. Fields, eds., American Society for Testing and Materials, Philadelphia, 1988, pp. 323-336.

3   Karpur, Prasanna, and Resch, Michael T., "Improved Detectability of Fatigue Microcracks by Split Spectrum Processing of Backscattered Rayleigh Waves", <u>Review of Progress in Nondestructive Evaluation</u>, Vol. 10, D.O. Thompson and D.E. Chimenti, eds., Plenum Press, New York, 1991, in press.

4   Larsen, J.M., Jira, J.R., and Weerasooriya, T., "Crack Opening Displacement Measurements on Small Cracks in Fatigue", Fracture Mechanics: Eighteenth Symposium, ASTM STP 945, D.T. Read and R.P. Reed, Eds., American Society for Testing and Materials, Philadelphia, 1988, pp. 896-912.

5   Karpur P., Shankar, P.M., Rose, J.L. and Newhouse, V.L., "Split Spectrum Processing: Optimizing the Processing Parameters Using Minimization", <u>Ultrasonics</u>, Butterworth Scientific, Vol. 25, July 1987, pp. 204-208.

6   Karpur P., Shankar, J.L., Rose, J.L. and Newhouse, V.L., "Split Spectrum Processing: Determination of the Available Bandwidth for Spectral Splitting", <u>Ultrasonics</u>, Butterworth Scientific, Vol. 26, July 1988, pp. 204-209.

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

NMR and IR Investigations of Conformational Dynamics and Surface Interactions of Perfluoropolyalkylethers

Prepared by:	Martin Schwartz
Academic Rank:	Professor
Department and	Chemistry Department
University:	University of North Texas
Research Location:	Materials Laboratory WRDC/MLBT Wright-Patterson AFB Ohio 45433-6533
USAF Researcher:	Dr. Kent J. Eisentraut (Technical Focal Point) and Dr. Harvey L. Paige
Date:	September 13, 1990
Contract No:	F49620-88-C-0053

NMR and IR Investigations of Conformational Dynamics and Surface Interactions of Perfluoropolyalkylethers

by

Martin Schwartz

ABSTRACT

Fluorine-19 NMR spin-lattice (T_1) relaxation times were measured for several perfluoropolyalkylethers (PFPAE's). Derived rotational correlation times (τ_c) revealed that perfluoromethylene (CF_2) groups adjacent to $-OC_2F_4O-$ chain segments rotate more slowly than those attached to $-OCF_2O-$ fragments. The decreased chain mobility was investigated using molecular mechanics to model bond rotations in linear PFPAE's. The calculations revealed that $-OC_2F_4O-$ units introduce steric repulsions which severely restrict rotation about neighboring C-O bonds. These results can be used to explain the generally observed correlation between C:O ratios and fluid viscosities in perfluoroethers.

Preliminary semi-empirical quantum mechanical calculations of conformational energies and potential barriers have been performed for several perfluorocompounds. The results will be compared with those from ab initio computations and, when available, to experimental data. The ultimate goal of these studies is to develop realistic conformational potential energy functions, which will permit the prediction of static and dynamic fluid properties and, therefore, aid in the design of new PFPAE lubricants.

The application of infrared microscopy to characterize the interactions of fluid additives with metal surfaces was investigated. Several problems in the acquisition of reliable spectral data were noted. It is recommended that further tests be performed in conjunction with XPS experiments to provide a definitive assessment of the utility of IR spectroscopy in the study of chemisorbed additives.

ACKNOWLEDGEMENTS

I. INTRODUCTION

Perfluoropolyalkylethers (PFPAE's) are a class of liquids possessing many physical and chemical properties desirable in a liquid phase lubricant,[1,2] including a wide liquid range, excellent thermal and oxidative stability, a high viscosity index, good lubricity and shear stability, and, importantly, they are almost completely non-flammable. No currently available commercial PFPAE lubricants, however, are capable of operation at the temperature extremes (-54 °C to +371 °C), in an oxidative environment, required for lubrication of a high performance gas turbine jet engine. The Air Force has begun a multi-year interdisciplinary program (IHPTET)[3] to study the physical and chemical properties of PFPAE's in order to develop suitable high temperature jet engine lubricants and additives.

I have extensive experience in the application of Nuclear Magnetic Resonance (NMR) and Fourier Transform Infrared (FTIR) spectroscopy to study the reorientational dynamics and intermolecular interactions of molecules in the liquid phase. Both NMR and FTIR spectroscopy are potentially quite useful for the investigation of the liquid phase structure of lubricants and their interactions with surfaces. Therefore, I was invited to participate in the high temperature lubricant research program at the Materials Laboratory at Wright-Patterson AFB this summer.

II. OBJECTIVES OF THE RESEARCH EFFORT

During my pre-summer visit to the Materials Laboratory, Dr. Kent Eisentraut (my technical focal point), Ed Snyder, and I determined that I would pursue two research goals this summer: (1) explore the application of Fluorine-19 NMR spectroscopy to study the liquid phase three dimensional structure (conformation) and polymer chain dynamics of PFPAE lubricants and (2) assess the utility of FTIR microscopy to profile the interactions of lubricants and additives with metal

surfaces. The NMR spectra were to be acquired by the graduate students in my research group at the University of North Texas on samples provided by the Materials Laboratory. The latter FTIR investigations were to be performed on site in collaboration with Dr. Jim Liang, a research chemist with the University of Dayton Research Institute, who is permanently assigned to the Materials Laboratory.

During the course of the summer, it became clear that quantum mechanical and classical molecular modelling, using programs available at the Materials Laboratory offered an excellent approach, complementary to NMR, to investigate the conformation and chain dynamics in liquid lubricants. Work in this area was begun in the latter half of the summer in collaboration with Dr. Harvey Paige of the Materials Laboratory.

III. NMR RELAXATION AND CONFORMATIONAL DYNAMICS

<u>Introduction</u>   The measurement of Nuclear Magnetic Resonance spin-lattice relaxation times (T_1's) is a well established technique to probe both the rates and mechanism of molecular reorientation in liquids.[4] The method has also been used quite profitably to characterize the conformational mobility of flexible chain polymers.[5,6] Almost all of the studies to date have been on carbon-13 or proton relaxation, commonly in hydrocarbon, halocarbon or alkylether polymer chains. During the summer, we investigated the applicability of fluorine-19 NMR relaxation times to characterize the chain reorientational dynamics in PFPAE's of varying molecular structure. We have also begun studies on the utility of a new two dimensional NMR technique (2D-NOESY spectroscopy) to furnish information on the equilibrium conformational behavior of these molecules.

<u>Experiments</u>   Fluorine-19 NMR measurements were performed by students in my research group at the University of North Texas on a Varian VXR-300 Fourier Transform NMR Spectrometer operating at

118-5

$v_o(^{19}F)=282$ MHz ($B_o=70.5$ kG). Spin-lattice (T_1) relaxation times were determined using the standard Inversion Recovery pulse sequence,[7] $(180°-\tau-90°-Acq.)_n$, with 10-12 τ values plus $\tau\to\infty$. T_1 was calculated from the peak intensities by a non-linear fit to the three parameter magnetization equation.[8]

Measurements were performed on the following liquids (at ambient temperature, 25 °C): (1) Perfluoropoly(ethylene oxide), $CF_3O[CF_2CF_2O]_nCF_3$ (ML088-50); (2) Perfluoropoly(triethylene glycol), $R_fO[(CF2CF2O)_3CF_2O]_nR_f$,$R_f = CF_3$, C_2F_5 (ML088-131); (3) Krytox-AC, $C_3F_7O[CF(CF_3)CF_2O]_nC_2F_5$ (ML071-6); (4) Fomblin-Z, $R_fO[CF_2O]_m[C_2F_4O]_n[C_3F_6O]_qR_f$, $R_f=CF_3$, C_2F_5 (ML078-80). In addition, two dimensional COSY[9,10,11,12] and NOESY[13],[12,14] spectra were acquired for Krytox-AC and Fomblin-Z using standard techniques.[12,15]

Results The ^{19}F NMR results for the five compounds above are presented in Tables 1-4 of Appendix A. As expected, the NMR spectra of Krytox and Fomblin are more complex than those of the model compounds (ML088-50, ML088-100, ML088-131). A partial set of assignments for these two lubricants are given in the tables. We have acquired 2D COSY spectra for these liquids which, together with NMR spectra of a second molecular weight fraction (to permit assignment of end groups), to be obtained this fall, will permit us to complete their assignments.

The predominant relaxation mechanism for fluorine nuclei in large molecules is via dipolar interactions with neighboring ^{19}F spins. Therefore it is possible, using standard relations,[16] to calculate a quantity termed the correlation time, τ_c, which, for a CF_2 unit in a perfluoroether, represents (approximately) the time for the vector between the two fluorine nuclei to rotate by one radian (57°). Thus, τ_c, is a measure of the degree of chain mobility (flexibility) in the region of the CF_2 group.

Several trends have emerged from the data acquired this summer. In all cases, rotational times for CF_3O- end groups are very short,

verifying that there is no significant barrier to rotation of this group about the C-O axis. Reorientation of the perfluoromethyl group in CF_3CF_2O- units is somewhat slower (longer τ_c), as a result of the three-fold barrier to rotation about the C-C bond. One finds, too, from the tables that rotational times (τ_c) of CF_2 groups within three atoms of the chain ends are comparatively short, a feature generally observed in polymer dynamics,[5,6] since there is a lowered frictional torque retarding bond rotation in this region of the chain.

The most significant trend in the data analyzed to date may be seen in the correlation times of Perfluoropoly(triethylene glycol) (A.2) and Fomblin (A.5). Without exception, perfluoromethylene (CF_2) rotational correlation times are shorter in $-\underline{CF_2}OCF_2O-$ units than in $-\underline{CF_2}OC_2F_4O-$ fragments on the chain. In order to understand this observation, we utilized a molecular mechanics modeling program (CHEM-X) to simulate rotation about the two central C-O bonds in a $-OCF_2CF_2-O-CF_2-O-$ segment of a perfluoropolyether chain. From the resulting potential energy contour plot, we observed that rotation of the right hand $-OCF_2O-$ fragment about the first C-O bond is severely restricted due to steric repulsions between this CF_2 group and the $-CF_2CF_2-$ unit on the left hand side of the chain segment. In contrast, the three-fold barrier to rotation of the left hand $-OCF_2CF_2O-$ fragment about the second C-O bond is relatively low, resulting in more rapid rotation and, thus, shorter correlation times for this group. We note that the lowered rotational mobility caused by the presence of $-OC_2F_4O-$ (and $-OC_3F_6O-$) segments in the polymer chain can explain the generally observed positive correlation between C:O ratios and viscosities in PFPAE's.

NMR relaxation times measure the rate of change of a polymer's conformation via rotation about its various bonds. The equilibrium (or average) conformation also exerts a major influence on the molecule's bulk and molecular properties. Recently, two dimensional NOESY

experiments have been applied to study proton...proton distances and, hence, equilibrium structures in proteins[17,18] and other polymers.[19] We are trying to extend this new experimental technique to obtain quantitative fluorine···fluorine distances (which are a sensitive function of the average dihedral angles about the C-C and C-O bonds) in PFPAE's. During the summer, we have acquired initial fluorine NOESY spectra in Fomblin-Z. Further refinement of the spectral acquisition parameters, necessary to obtain quantitative intensities and F···F distances, is currently in progress.

IV. MOLECULAR MODELLING OF CHAIN CONFORMATION AND MOBILITY

<u>Introduction</u>   In order to understand and, eventually, to predict the equilibrium conformations and chain mobilities in PFPAE's, it is essential to have an accurate knowledge of the potential energy as a function of rotation angle about various single bonds in these fluids. Once the potential energy functions have been accurately characterized, it is possible to utilize well-established statistical mechanical methods[20] to calculate both static and dynamic properties as a function of temperature in the liquid phase.

During the latter half of the summer, I began collaborative studies with Dr. Harvey Paige of the Materials Laboratory to investigate the applicability of <u>ab initio</u>[21] and semi-empirical[22] quantum mechanical techniques and of classical molecular mechanics to determine conformational potential energy functions in PFPAE's.

<u>Calculations</u> <u>and</u> <u>Results</u>   As discussed in a previous section, the classical molecular mechanics model was employed successfully to explain the trend in rotational correlation times obtained from NMR relaxation times in Fomblin-Z. During the summer, we also utilized this simple model to generate energy contour maps for rotation about the two central C-O bonds in chain segments of a linear ($-OCF_2-O-CF_2O-$) and branched ($-OCF_2-O-CF(CF_3)O-$) perfluoroether. The results revealed that the

branching perfluoromethyl group in the latter segment induces steric repulsions which severely restrict the rotation about the right hand $O-CF(CF_3)$ bond. The resulting reduced chain mobility explains, qualitative, the general observation that branched PFPAE's exhibit an order of magnitude greater viscosity than linear perfluoroethers of the same molecular weight.

We also began preliminary calculations on the conformational energies of several simple perfluorocompounds (Appendix B), using three of the semi-empirical quantum mechanics programs (MNDO, AM1 and PM3) in the MOPAC[22] molecular orbital package. The three methods differ primarily in their integral parametrization by comparison to experimental results. As a consequence of their comparative scarcity, there is relatively little data on perfluorocompounds, with which to parametrize the semi-empirical Hamiltonian. Our goal in this portion of the project is to determine which of the three Hamiltonians yields conformational energies in closest agreement with ab initio calculations and, where available, with experimental conformational energy differences, potential barriers and molecular dipole moments.

Perfluorodimethylether (CF_3OCF_3) is one molecule whose structural parameters have been determined experimentally,[23] and for which ab initio calculations have been reported in the literature.[24] As seen in Table B.1, the PM3 Hamiltonian yields bond lengths and angles that, generally, are in superior agreement with experiment than those from either AM1 or MNDO, or, interestingly, from the ab initio calculation using a minimal basis set (STO-3G). The results from PM3 agree reasonably well, too, with parameters calculated with the extended 4-31G basis functions.

The hypothetical molecule, tetrafluoroethane-1,2-diol ($HOCF_2CF_2OH$), like perfluoroethers, contains polar C-F and C-O bonds, whose interactions influence the relative energies of equilibrium conformers and potential barriers. We have used the three

semi-empirical Hamiltonians (PM3, AM1 and MNDO) and classical molecular mechanics (MM2) to calculate the energy and dihedral angle of the equilibrium gauche conformer (E_G) and of the two barriers (E_{GG} and E_{GT}), relative to $E_T=0$. One sees from Table B.2 that the energies from the various methods differ substantially, by as much as a factor of 3-4. Work is in progress to determine which of these methods yields results in closest agreement with an extended basis set _ab initio_ calculation.

The conformational potential energy function has also been calculated for perfluorobutane ($CF_3CF_2CF_2CF_3$) using the semi-empirical methods. As found above, the three Hamiltonians yield markedly different results (Table B.3). Here, too, _ab initio_ calculations are in progress. We also intend to obtain perfluorobutane and measure the relative equilibrium energies of the gauche and trans conformers experimentally (by temperature dependent IR intensity ratios) as well as its gas phase dipole moment. A comparison with the predictions of the various methods will provide important evidence on their relative merits.

V.   IR MICROSCOPY AND SURFACE INTERACTIONS

Introduction   To improve their performance under extreme conditions, most lubricants require the addition of small amounts of various additives. One current theory of the action of these additives is that they chemisorb on an engine's metal surfaces, passivating the metal from corrosion, oxidation and wear. FTIR microscopy[25] has proven to be a useful new technique to study the adsorption of molecules on surfaces.[26] Since one can monitor a region as small as 10-20 microns in diameter, it is potentially possible obtain a detailed profile of additive/surface interactions on scarred or corroded regions of metal. Too, the infrared spectra may be able to provide more detailed information on the structure of chemisorbed species than is possible from most other surface analysis techniques. One goal of my summer

research at the Materials Laboratory was to assess the capability of FTIR microscopy to study chemisorption of additives on metal surfaces that had been exposed to four ball wear and oxidation/corrosion tests. This work was performed in collaboration with Dr. Jim Liang of the University of Dayton Research Institute.

Experiments   IR spectra were obtained on a Perkin-Elmer 1750 Fourier Transform Infrared Spectrometer, equipped with a Spectra Tech IR-Plan Infrared Microscope (Model no. 0043-033) and MCT detector. The following tests were run: (1) Four ball wear test on 5P4E (5-phenyl-4-ether) lubricant without additive, (2) Four ball wear test on 5P4E with 5% TCP (tricresyl phosphate) antiwear additive, (3) washed balls from earlier four ball wear test, (4) Oxidation/Corrosion test on a branched PFPAE (MLO89-293), and (5) Reflection tests of unused lubricants on metal coupons.

Results   Three problems were encountered in the application of IR microscopy to study additive/metal interactions: (1) similarity of lubricant and additive IR spectra, (2) spectral anomalies caused by reflection off uneven metal surfaces, and (3) sensitivity; each is addressed in turn.

(1) Similarity of lubricant and additive spectra. Standard organic additives lack sufficient solubility to be useful in improving the performance of the new high temperature lubricants. Therefore, it has been necessary to synthesize new additives with structures very similar to that of the lubricant in order to enhance solubility. As a consequence, one usually finds that the IR of the additives and base fluids overlap in all spectral regions. Thus, it is quite difficult to monitor changes in the additive's spectrum if any lubricant remains on the metal.

(2) Spectral anomalies. It is well known that IR reflection spectra of films on metal surfaces depend critically upon the film thickness and reflection angle;[27,28] neither thickness nor angle can be controlled

on scarred or corroded metal surfaces. As a result, we observed that
the relative peak intensities and positions of adsorbed film spectra
varied irregularly from one experiment to the next, even when unused
(and thus unaltered) lubricants were applied to clean balls or metal
coupons. Therefore, one must be extremely cautious in ascribing any
variations in the IR spectrum to chemical modification resulting from
surface adsorption.

(3) Sensitivity. Ideally, one may eliminate the above problems by
washing the metal surface to remove any residual lubricant film on the
surface, leaving only the chemisorbed additive. This procedure has been
reported to be successful in obtaining the XPS spectrum of adsorbed
species. A number of experiments were performed on washed balls and OC
coupons. Unfortunately, in none of the tests were we able to observe
any signal due to chemisorbed species.

VI. RECOMMENDATIONS

NMR Relaxation   As outlined above, the results obtained to date
demonstrate that Fluorine-19 NMR T_1's can provide valuable information
on the effect of the local molecular environment on the chain mobility
in PFPAE's. Investigations of additional model compounds and of the
effect of medium (temperature, solvent, polymer mixtures...) are
currently underway. In addition, we intend to begin measuring carbon-13
T_1's and spin-spin (T_2) and rotating frame ($T_{1\rho}$) relaxation times to
obtain complementary information on the chain dynamics. I am currently
preparing a research proposal for an RIP grant to request funding for
this continuing research.

Molecular Modelling   Although calculated parameters obtained from the
PM3 semi-empirical Hamiltonian were in good agreement with the
experimental values, there is insufficient data to ascertain whether
this method is generally superior to AM1 and MNDO. As further _ab initio_
and experimental results are accumulated, it may prove appropriate to

develop a new semi-empirical Hamiltonian designed specifically for the perfluoroethers.[29]

The simple molecular mechanics model has proven useful to ascertain qualitative trends in conformational energies of perfluoroethers (vide supra). However, the results bear no quantitative significance. As reliable quantum mechanical and experimental results become available, it will be possible to introduce realistic torsional potentials and bond stretching and bending force constants into a modified molecular mechanics model parametrized specifically for PFPAE's. It will then become feasible to utilize this very fast computational procedure to perform molecular dynamics calculations on the PFPAE's. The results can be used to predict rotational correlation times, which can be compared to those measured experimentally from NMR relaxation to determine the adequacy of the modified classical model. Once accurate potential functions have been developed, they may be used to calculate both static and dynamic properties of these fluids, thus aiding in the design of new PFPAE lubricants.

IR Microscopy   It is apparent from the problems described in Section V that reliable IR spectra of chemisorbed additives cannot be determined in the presence of residual lubricant fluids due to their spectral similarities and the anomalous intensities and peak positions of reflectance spectra from liquid films. As noted, when the balls (from four ball experiments) or OC coupons were washed, no IR spectra of adsorbed species were detected. However, rather than resulting from a lack of sufficient sensitivity, it is possible these metals contained no significant amounts of the additive. We (Dr. Liang and I) recommend that when XPS studies on these systems are begun in the coming year, that IR spectra be recorded at the same time. A comparison of results from the two experiments should provide a definitive test of the utility of infrared microscopy to characterize the interactions of additives with metal surfaces.

REFERENCES

1. C. E. Snyder, Jr. and R. E. Dolle, Jr., <u>ASLE Trans.</u> 19, 171 (1975)

2. C. E. Snyder, Jr., L. J. Gschwender and C. Tamborski, <u>Lubr. Eng.</u> 37, 344 (1981)

3. Integrated High Performance Turbine Engine Technology (IHPTET) Program

4. R. T. Boeré and R. G. Kidd, <u>Ann. Rep. NMR Spectrosc.</u> 13, 319 (1982)

5. F. Heatley, <u>Prog. in NMR Spectrosc.</u> 13, 47 (1979)

6. F. Heatley, <u>Ann. Rep. NMR Spectrosc.</u> 17, 179 (1986)

7. M. L. Martin, G. J. Martin and J.-J. Delpuech, <u>Practical NMR Spectroscopy</u>, Heyden, London (1980), Chap. 6

8. A. A. Rodriguez, S. J. H. Chen and M. Schwartz, <u>J. Magn. Reson.</u> 74, 114 (1987)

9. COSY is the acronym for Correlation Spectroscopy

10. W. P. Aue, E. Bartholdi and R. R. Ernst, <u>J. Chem. Phys.</u> 64, 2229 (1976)

11. A. Bax and R. Freeman, <u>J. Magn. Reson.</u> 42, 181 (1981)

12. R. Benn and H. Günther, <u>Angew. Chem. Int. Ed. Engl.</u> 22, 350 (1983)

13. NOESY is the acronym for Nuclear Overhauser Enhancement Spectroscopy

14. B. H. Meier and R. R. Ernst, <u>J. Am. Chem. Soc.</u> 101, 6441 (1979)

15. G. Wider, S. Macura, A. Kumar, R. R. Ernst and K. Wüthrich, <u>J. Magn. Reson.</u> 56, 207 (1984)

16. E. D. Becker, <u>High Resolution NMR: Theory and Applications</u>, 2nd. Ed., Academic, New York (1980), Chap. 9

17. A. Kumar, G. Wagner, R. R. Ernst and K. Wüthrich, <u>J. Am. Chem. Soc.</u> 103, 3654 (1981)

18. M. Weiss, D. Patel, R. Sauer and M. Karplus, <u>Proc. Natl. Acad. Sci. USA</u> 81, 130 (1984)

19. P. A. Mirau, F. A. Bovey, A. E. Tonelli and S. A. Heffner, <u>Macromolecules</u> 20, 1701 (1987)

20. P. J. Flory, <u>Statistical Mechanics of Chain Molecules</u>, Wiley-Interscience, New York (1969)

21. W. J. Hehre, L. Radom, P. von R. Schleyer and J. A. Pople, Ab Initio Molecular Orbital Theory, Wiley, New York (1986)

22. J. J. P. Stewart, J. Computer-Aided Molec. Design 4, 1 (1990)

23. J. Pacansky and B. Liu, J. Phys. Chem. 89, 1883 (1985)

24. A. H. Lowrey, C. George, P. D'Antonio and J. Karle, J. Mol. Struct. 63, 243 (1980)

25. R. G. Messerschmidt and M. A. Harthcock, Infrared Microspectrometry: Theory and Applications, Marcel Dekker, New York (1988)

26. J. F. Rabolt, M. Jurich and J. D. Swalen, Appl. Spectrosc. 39, 269 (1985)

27. D. L. Allara, A. Baca and C. A. Pryde, Macromolecules 11, 1215 (1978)

28. V. J. Novotny, I. Hussia, J.-M. Turlet and M. R. Philpott, J. Chem. Phys. 90, 5861 (1989)

29. Dr. James J. P. Stewart, Project Scientist at the Frank J. Seiler Research Laboratory, was a principal participant in the development of the MOPAC molecular orbital package, and has, himself, developed the newer PM3 extension. Dr. Stewart, in correspondence with Dr. Harvey Paige, has indicated his willingness to develop a new semi-empirical parameter set specifically for perfluorocompounds, as ab initio and experimental results become available.

APPENDIX A. NMR RELAXATION AND CORRELATION TIMES

A.1 Perfluoropoly(ethylene oxide) , $CF_3O[CF_2CF_2O]_nCF_3$ (ML098-50)

Peak	Delta*	Assignment	T_1	τ_c
1	73.6 ppm	$CF_3OCF_2CF_2O$	0.80 s	150 ps
2	75.8	OCF_2CF_2O	0.47	250
3	108.3	$CF_3OCF_2CF_2O$	1.09	60

*Chemical shifts in this and following tables are measured in ppm downfield from hexafluorobenzene.

A.2 Perfluoropoly(triethylene glycol) , $R_fO[(CF2CF2O)_3CF_2O]_nR_f$, $R_f = CF_3$, C_2F_5 (ML088-131)

Peak	Delta	Assignment	T_1	τ_c
1	73.7 ppm	$CF_3OCF_2CF_2O$	0.81 s	140 ps
2	74.0	$OCF_2CF_2OCF_2O$	0.63	180
3	75.7	$OC_2F_4OCF_2CF_2OC_2F_4O$	0.59	200
4	76.8	CF_3CF_2O	0.60	110
5	108.2	$CF_3OC_2F_4O$	1.21	60
6	112.6	$OC_2F_4OCF_2OC_2F_4O$	0.67	210

A.3 Krytox-AC , $C_3F_7O[CF(CF_3)CF_2O]_nC_2F_5$ (ML071-6)

Peak	Delta	Assignment	T_1	τ_c
1	19.7	$OCF_2CF(CF_3)O$	0.20	130 ps
2	20.0	$OCF_2CF(CF_3)O$	0.20	130
3	20.3	$OCF_2CF(CF_3)O$	0.20	130
4	34.4	$CF_3CF_2CF_2O$	0.36	280
5	75.6	OCF_2CF_3	0.50	130
6	76.8	$CF_3CF_2CF_2O$	0.49	240
7	82.4	$CF_3CF_2CF_2O$	0.52	120
8	83.8	$OCF_2CF(CF_3)O*$	0.40	290
9	84.2	$OCF_2CF(CF_3)O$	0.38	170
10	84.3	$OCF_2CF(CF_3)O*$	0.38	310
11	84.8	$OCF_2CF(CF_3)O*$	0.40	290
12	85.3	$OCF_2CF(CF_3)O*$	0.37	320

*The CF_2 fluorines in Krytox are non-equivalent since they are adjacent to an asymmetric carbon. Therefore, they split into a quartet with a large geminal coupling constant

A.4 Fomblin-Z , $R_f O[CF_2O]_m[C_2F_4O]_n[C_3F_6O]_q$, $R_f = CF_3$, C_2F_5 (ML078-80)

Peak	Delta	Assignment	T_1	τ_c
1	35.2 ppm	$OCF_2CF_2CF_2O$	0.52 s	200 ps
2	74.2	$OCF_2CF_2OCF_2O$	0.63	190
3	75.9	$OCF_2CF_2OC_2F_4O$	0.51	230
4	79.5	$OCF_2CF_2CF_2OCF_2O$	0.63	190
5	81.1	$OCF_2CF_2CF_2OC_2F_4O$	0.50	230
6	106.9	CF_3OCF_2O	1.08	70
7	108.6	$CF_3OC_2F_4O$	1.19	60
8	109.5	$OCF_2OCF_2OCF_2O$	0.95	140
9	111.3	$OCF_2OCF_2OC_2F_4O$	0.78	180
10	111.5	CF_3OCF_2O	0.79	170
11	112.9	$OC_2F_4OCF_2OC_2F_4O$	0.66	210

APPENDIX B. CALCULATED CONFORMATIONAL ENERGIES

B.1   Perfluorodimethylether (CF_3OCF_3): Comparison between theory and experiment

Param	Expt.[a]	4-31G[b]	STO-3G[b]	PM3[c]	AM1[c]	MNDO[c]
C-O[d]	1.369	1.369	1.431	1.392	1.406	1.406
C-F[d]	1.327	1.340[g]	1.366[g]	1.334[g]	1.353[g]	1.344[g]
C-O-C[e]	119.1	125.9	113.2	118.9	121.3	130.4
F-C-O-C[f]	166.0	162.2	162.0	161.5	165.8	165.2

a) Ref. 23
b) Ref. 24
c) This work
d) Bond length (in Angstroms)
e) Bond angle (in degrees)
f) Dihedral angle (in degrees)
g) Average value

B.2   Tetrafluoroethane-1,2-diol ($HOCF_2CF_2OH$)[a]

Method	E_G^b	Φ_G^b	E_{GG}^c	Φ_{GG}^c	E_{GT}^d	Φ_{GT}^d
PM3	0.24	62	2.98	0	2.25	119
AM1	0.59	80	1.83	0	0.69	112
MNDO	0.53	64	4.15	0	3.21	120
MM2	0.27	62	6.41	2	5.13	120

a) All energies given in kcal/mol, referenced to $E_T=0$
b) Energy and dihedral angle of gauche (G+ and G⁻) equilibrium conformations
c) Energy and dihedral angle of barrier between G+ and G⁻ conformations
d) Energy and dihedral angle of barrier between G+ (or G⁻) and T conformations

B.3 Semi-Empirical Conformational Energies of Perfluorobutane[a]

ϕ[b]	PM3	AM1	MNDO
0°	5.89 kcal	4.60 kcal	8.01 kcal
10	6.47	4.52	7.54
20	6.74	3.94	6.41
30	2.72	3.31	4.99
40	1.17	2.62	3.60
50	-0.36	2.00	2.57
60	-0.08	1.53	2.11
70	1.29	1.28	2.19
80	1.62	1.16	2.37
90	0.93	0.81	2.13
100	0.25	0.62	2.06
110	0.98	0.59	2.23
120	1.43	0.56	2.27
130	1.42	0.47	2.00
140	1.24	0.31	1.45
150	0.47	0.14	0.80
160	-0.23	0.01	0.28
170	0.07	-0.02	0.04
180	0.00	0.00	0.00

a) All energies in kcal/mol, and referenced to $E(180°)=0$

b) Dihedral angle between C_1 and C_4

1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

MODELING OF CASTING SOLIDIFICATION

Prepared by: Hai-Lung Tsai

Academic Rank: Assistant Professor

Department and   Mech. and Aero. Engr. and Engr. Mech.

  University: University of Missouri-Rolla

Research Location: WRDC/MLLM

                      Wright-Patterson AFB, OH 45433

USAF Researcher: James C. Malas

Date: September 5, 1990

Contract No: F49620-88-C-0053

MODELING OF CASTING SOLIDIFICATION

by

Hai-Lung Tsai

ABSTRACT

A general purpose finite element computer program, CAST3, for modeling casting solidification was evaluated from both the user's and the technical aspects. The CAST3 code was developed by the Universal Energy Systems, Inc. under the sponsorship of the Air Force. Although several commercial packages are available, it was found that CAST3 is the only software dedicated to the casting solidification modeling. As a result, CAST3 code has a superior capability in handling the casting-mold interfacial thermal resistance and the time stepping algorithm, which make the program computationally more efficient than any other available codes (to the knowledge of the author). An excellent start has been made by CAST3 in achieving the goal of developing an ideal casting design package for the Air Force. However, the present version of CAST3 is not yet completed for being able to simulate some casting problems. Therefore, recommendations are made for improving and expanding CAST3, so that the code can be used in the foundry industry as a powerful design tool for obtaining high quality casting parts.

I. INTRODUCTION

Metal casting foundry is a basic industry which provides either raw materials or finished parts for many other manufacturing industries. Hence, the casting production technology in the foundry will affect directly or indirectly the cost as well as the quality of many downstream products. During the past years, the production of American castings has been decreased significantly, due to the loss of market to foreign countries where the labor is much cheaper [1]. One of the major issues for the American foundry to survive and maintain the leading edge in the global competition is the introduction of modern technology to this industry for obtaining castings of better quality at a lower cost. To achieve this, the use of computer modeling as well as CAD/CAM in the foundry are required.

To date, several commercial finite element packages are available for modeling casting solidification, which include ABAQUS, ADINA, ANSYS, NASTRAN, NISA, MARK, etc. In general, these programs were developed for the general purpose use in many areas, and they are not dedicated to the metal casting simulation. As a result, special features unique to casting solidification, such as the latent heat release and the metal-mold interfacial thermal resistance are not well-handled in these programs, which crucially affect the required computer time. Consequently, for any of the above-mentioned code, it needs a tremendous computer time to simulate a real casting process, which hinders the usage of computer modeling in the

foundry industry.

During the past years, Drs. Mark Samonds and David Waite of
the Universal Energy Systems, Inc. have developed a general
purpose finite element code, CAST3, dedicated to the modeling
of casting solidification, under the sponsorship of the Air
Force. It has been demonstrated by several foundries that the
basic portion (heat conduction) of CAST3 is very useful in pre-
dicting the transient temperature distributions in the casting
and the mold, so that any possible hot spots in the casting
could be avoided by changing some casting conditions. However,
so far a rigorous evaluation on CAST3, from either a user's
point of view or the technical aspect, is not available. The
author has worked on the casting solidification modeling for
about 10 years, including the use of several above-mentioned
commercial codes and many years of actual foundry experience.
In addition, he has developed part of NISA code, and is famil-
iar with both finite element and finite difference methods.
Hence, the author is one of the very few candidates who are
qualified for evaluating CAST3 code.

II. OBJECTIVES OF THE RESEARCH EFFORT

My assignment as a participant in the 1990 Summer Faculty Research Program was to evaluate the code CAST3, and then to make recommendations on the direction for further research in the modeling of casting solidification. The Air Force was particularly interested in the D'arcy flow portion in the CAST3. Hence, it was determined to run CAST3 for a unidirectional macrosegregation problem, so that CAST3 can be validated by comparing the computational results with the available experimental data.

The CAST3 User Manual was studied in detail and a one dimensional solid model was established using PATRAN for simulating the macrosegregation problem. As many places in the manual were not clearly described, the CAST3 program developer, Dr. Mark Samonds, was called to discuss the technical contents of CAST3. In order to be able to exchange ideas more convenient, Mark was invited to the laboratory. The discussion has gone through every phase of the program CAST3. It was found that the D'arcy flow portion in CAST3 was developed under the assumption that a continuous feeding of molten metal is provided, which is inconsistent with an actual macrosegregation experiment, in which there is no additional feeding. It was also found that the proposed macrosegregation problem has been simulated using CAST3 by Dr. David Waite, and the results were available. Although the modeling results appeared reasonable based on the assumption made when the program was developed,

they were completely different from the experimental data. It was concluded that a major modification in CAST3 is required in order to be able to predict the macrosegregation phenomenon observed in experiments. As a result, the idea of using CAST3 to simulate the unidirectional macrosegregation problem was abandoned, and efforts were focused on identifying the merits and limitations of CAST3, as well as the directions for possible improvements on CAST3.

III. MERITS AND LIMITATIONS OF CAST3

The most distinctive difference between CAST3 and other
commercial programs is that CAST3 was developed solely for
modeling casting solidification. Hence, special features
unique to casting solidification are well-handled in CAST3.
The merits of CAST3 are summarized as follows:

MERITS:

1. Method in handling metal-mold interfacial problem

The technique used in CAST3 to handle the casting-mold
interfacial thermal resistance is excellent. In casting model-
ing, the domain consists of both the casting and the mold,
which is a conjugate problem. It is noted that usually the
thermophysical properties of the casting are very different
from those of the mold. For example, the thermal diffusivity
of aluminum based alloys is about 100 times that of a silica
sand mold. Especially, there is a phase change in the casting
which introduces nonlinearity in the numerical solution. The
technique used by CAST3 can have a complete different grid sys-
tem and time stepping scheme for each of the casting and mold
domains, so that the computational efficiency can be increased.
The author uses a similar technique to handle the metal-mold
interface, and it is called the "two-domain method" [2].

2. Method in handling latent heat release

In CAST3, the enthalpy method is employed for handling the
latent heat release during alloy solidification. Enthalpy

method has been the most popular technique to deal with latent heat, even if a fluid flow presents in the casting. The relationship between the casting solid fraction and temperature can be assumed [3], or directly input from the alloy phase diagram. This provides a more realistic way for implementing the latent heat in the program.

3. Time stepping algorithm

In the time domain, depending upon the characteristics of the problem, CAST3 uses two- or three-level predictor-corrector time stepping scheme, which gives the user a flexibility. The scheme can be fully implicit, fully explicit, or in between of them. Proper selection of the time stepping scheme can substantially save the computational time. For example, due to the large difference in material properties between the casting and mold, a fully implicit scheme can be used for the casting domain to guarantee the numerical convergence, while a fully explicit scheme can be employed in the mold to reduce the computer time.

4. Computational efficiency

By using CAST3, significant computational cost can be reduced due to the techniques employed in handling the metal-mold interfacial thermal resistance and the time stepping algorithm, as mentioned above. In addition, CAST3 uses the Conjugate Gradient method to solve the assembled matrix, which could be more efficient than the traditional Gaussian elimination technique. This is particularly important for a large, sparce

equation systems, which usually occur in a finite element analysis. It is expected that CAST3 code can save more than 50% of computational time, as compared with commercial codes such as ABAQUS and ANSYS, with which the author is familiar.

5. Capability in calculating radiation view factor

CAST3 has the capability to calculate the changes of view factor as a function of time, so that the radiation heat transfer can be properly accounted. This special feature is particularly useful for the simulation of single crystal growth during the pulling process.

LIMITATIONS:

1. Heat conduction only

The major limitation of CAST3 is that only heat conduction in the casting can be handled, but not the fluid flow problem. The D'arcy flow portion in CAST3 was developed under oversimplified assumptions, and it is generally not applicable to a real casting process. As a result, several casting defects associated with the fluid flow in the casting cannot be predicted by CAST3. These include the formations of microsegregation, macrosegregation, and porosity in the casting. The fluid flow in the casting can be caused by thermal and/or solutal gradients, as well as shrinkage [4].

2. Mold is instantaneously filled

Although theoretically CAST3 allows arbitrary initial conditions for the casting and the mold, uniform initial tempera-

tures are usually assumed due to the unknown initial conditions. In order to obtain the initial temperature and/or velocity distributions, the mold filling process must be simulated. In addition, the study of mold filling is important for many casting processes, for example, the thin-section casting, die casting, and lost foam casting.

IV. RECOMMENDATIONS FOR IMPROVING CAST3

From the above discussion, it can be concluded that CAST3 is an excellent program for modeling casting solidification, if only heat conduction in the casting is considered. The program can be employed to determine the constant temperature curves, so that any possible hot spots in the casting can be detected. Then, by changing the casting condition a favorable solidification pattern can be achieved and, consequently, hot spots and the associated casting defects can be eliminated.

However, several improvements in CAST3 are required in order to make CAST3 a true tool for predicting and solving several casting problems. A brief discussion of some areas for possibly including in the future version of CAST3 will be given in the following. However, it is noted that several topics to be discussed are still under intensive investigation and, perhaps, are not matured enough to be implemented in a casting program such as CAST3.

1. More user-friendly

The CAST3 User Manual needs to add several well-defined and worked-out examples to illustrate the procedure, step by step, how to use the program. The theoretical development and the user's manual could be separated into two volumes. There are many misprints in the manual. It is recommended that the terminologies used in the manual should be consistent with the tradition. For example, when there is a fluid flow in the

casting, the heat transfer occurred is called convection, but not conduction as used in the manual. Also, several data input formats are not convenient as comparing with the typical forms available in the literature. In conclusion, additional efforts need to be made to enhance the user-friendly of CAST3.

2. Mold filling process

The simulation of mold filling process is important not only it provides the initial conditions for the subsequent casting solidification modeling, but also many casting defects are determined by the filling process. For example, in die casting, lost foam casting, and injecting molding, the quality of casting parts is nearly determined by the mold filling process. The most popular technique to handle the free surface during filling process is described in a program, SOLA-VOF, developed by the Los Alamos Scientific Laboratory. However, the original program can handle only the fluid flow portion, but not the heat transfer. Additional capability needs to be developed if solidification occurs during the filling process.

3. Fluid flow problem

It is well known that the fluid flow in the casting, due to either the forced convection, natural convection, shrinkage, or their combination, can significantly affect the casting quality. Hence, a casting solidification model must be able to handle the fluid flow in the casting. The difficulty in handling the fluid flow in the casting stems from the fact that a solidifying casting consists of the solid, liquid, and mushy

regions. In addition to the unknown shapes of the interface between different phases, there is a latent heat release. The state-of-the-art technique uses a single set of differential equations throughout the entire casting domain [4].

4. Method for handling mushy region

In the mushy region, both the solid and liquid phases coexist, hence a special technique need to be developed to handle the fluid flow and heat transfer in the mushy region. It is noted that the fluid flow in the mushy region for a solidifying casting is quite different from the traditional two-phase flow typically found in the chemical or nuclear engineering. The state-of-the-art technique is to use the continuum model for the mushy region, which is based on the volume averaging scheme [4].

5. Gas- and shrinkage-caused porosities

Casting porosity has been one of the oldest and most concerned problems in the foundry industry, in particular, for the aluminium based alloys. The existence of micro pores significant decreases the strength of the casting parts, especially under the environment at high temperature and high pressure. The pores in the casting can be caused by either the dissolved gases in molten metal, the shrinkage, or their combination. Intensive research on the formation of porosity in casting is still undertaken both in the academia and the industry [5].

6. Macrosegregation

For large casting parts, a severe large-scale macrosegregation has been frequently found. It is noted that once the macrosegregation is created in the casting, there is no means to eliminate it, except just discard the casting. Macrosegregation in the casting is caused by the fluid flow both in the liquid phase and the mushy region. Although some over-simplified models on macrosegregation were published in the early 1960's, a comprehensive mathematical model is still not available to date.

7. Shrinkage-induced domain change and fluid flow

Although the fluid flow caused by shrinkage is expected to be very small, it has been found that some casting defects such as segregation and porosity are mainly owing to the shrinkage induced fluid flow [6]. This is understandable, since when the solid fraction in the casting exceeds 50%, the fluid flow in the mushy zone due to natural convection is almost diminished. However, the fluid flow caused by shrinkage is always present, due to the density difference between the solid and the liquid phases.

8. Actual latent heat release

The use of phase diagram to account for the latent heat release is possible if there are only two major constitutes in the alloy. Also, phase diagram is obtained under the assumption that an equilibrium exists. In any casting solidification process, some degree of undercooling is always existent in the

solidifying metal. Hence, the rate of actual latent heat release should be determined by experiments, so that it can be directly implemented in the modeling [7].

It is noted that the above-mentioned topics by no means is a complete list for CAST3 to improve. For example, the stress built up in the casting during solidification is very important for determining the hot tear and casting distortion, which, in turn, can affect the gap formation between the casting and the mold.

V. CONCLUSIONS

The computer program, CAST3, developed by the Universal
Energy Systems, Inc., was evaluated during the Summer Faculty
Research Program. CAST3 code is a unique computer program
dedicated to modeling casting solidification. The code has
many superior capabilities than any existing commercial codes.
However, additional improvements and expansion of CAST3 are
required. It is felt that CAST3 has made an excellent start to
become a sophisticated code, which can be truly used in the
foundry for designing casting processes.

In view of the increasing demands for high quality casting
parts, there is a strong need for a mathematical model, which
can predict and, subsequently, eliminate possible casting
defects. Hence, it is suggested that the Air Force should con-
tinue to support the research and development in the area of
casting solidification modeling.

ACKNOWLEDGEMENTS

REFERENCES

1. *The Census of World Casting Production*, Published Yearly Since 1966 by the American Foundrymen's Society, Inc., Des Plains, IL.

2. Chen, J. H. and Tsai, H. L., "An Efficient and Accurate Numerical Algorithm for Multi-Dimensional Modeling of Casting Solidification - Part I. Control Volume Method," *AFS Transactions*, Vol. 98, 1990.

3. Chen, J. H. and Tsai, H. L., "Comparison on Different Modes of Latent Heat Release for Modeling Casting Solidification," *AFS Transactions*, Vol. 98, 1990.

4. Chiang, K. C. and Tsai, H. L., "Shrinkage-Induced Fluid Flow and Domain Change in Two-Dimensional Alloy Solidification," to appear.

5. Chiou, I. J. and Tsai, H. L., "Modeling on Porosity Formation in Castings," *AFS Transactions*, Vol. 98, 1990.

6. Chiang, K. C. and Tsai, H. L., "Interaction Between Shrinkage-Induced Fluid Flow and Natural Convection During Alloy Solidification," to appear.

7. Su, C. H. and Tsai, H. L., "A direct Method to Include Latent Heat Effect for Modeling Casting Solidification," to appear.